Fatigue in rheumatoid arthritis: from patient experience to measurement

Stephanie Nikolaus

Fatigue in rheumatoid arthritis: from patient experience to measurement

Stephanie Nikolaus

Thesis, University of Twente, 2012
ISBN/EAN 9789461082763
© Stephanie Nikolaus

Lay-out and printed by Gildeprint Drukkerijen, Enschede, the Netherlands

The studies presented in this thesis were performed at the department of Psychology, Health & Technology of the University of Twente (Enschede) and the Arthritis Center Twente at Medisch Spectrum Twente hospital (Enschede) and the hospitals of "Ziekenhuisgroep Twente" (Almelo and Hengelo), the Netherlands. The rheumatology research program of PHT is financially supported by the Dutch Arthritis Foundation (Reumafonds).
This project is financially supported by Stichting Reumaonderzoek Twente and the Institute of Behavioural Research.
Publication of this thesis was financially supported by the Dutch Arthritis Foundation (Reumafonds).

FATIGUE IN RHEUMATOID ARTHRITIS: FROM PATIENT EXPERIENCE TO MEASUREMENT

PROEFSCHRIFT

ter verkrijging van
de graad van doctor aan de Universiteit Twente,
op gezag van de rector magnificus,
prof.dr. H. Brinksma,
volgens besluit van het College voor Promoties
in het openbaar te verdedigen
op vrijdag 16 maart 2012 om 14.45 uur

door

Stephanie Nikolaus
geboren op 2 november 1980
te Kleef, Duitsland

Dit proefschrift is goedgekeurd door de promotor Prof. dr. M.A.F.J. van de Laar en assistent-promotoren Dr. C. Bode en Dr. E. Taal.

Samenstelling promotiecommissie:

Promotor:	Prof. dr. M.A.F.J. van de Laar
Assistent-promotoren:	Dr. C. Bode Dr. E. Taal
Leden:	Prof. dr. E.T. Bohlmeijer (University of Twente) Prof. dr. J. Dekker (Free University Amsterdam) Prof. dr. R. Geenen (University Utrecht) Prof. dr. C.A.W. Glas (University of Twente) Prof. dr. S. Hewlett (University of West England, Bristol UK) Prof. dr. P.L.C.M. van Riel (Radboud University, Nijmegen)

Contents

Chapter 1	General introduction	9
Chapter 2	Fatigue and factors related to fatigue in rheumatoid arthritis: a systematic review	21
Chapter 3	Measuring fatigue in rheumatoid arthritis	69
Chapter 4	New insights into the experience of fatigue among patients with rheumatoid arthritis: a qualitative study	75
Chapter 5	Four different patterns of fatigue in rheumatoid arthritis patients: results of a Q-sort study	89
Chapter 6	Selection of items for a computer-adaptive test to measure fatigue in patients with rheumatoid arthritis: a Delphi approach	115
Chapter 7	Which dimensions of fatigue should be measured in patients with rheumatoid arthritis? A Delphi study	137
Chapter 8	Expert's evaluations of fatigue questionnaires used in rheumatoid arthritis – A Delphi study among patients, nurses and rheumatologists in the Netherlands	149
Chapter 9	Calibration of a multidimensional item bank to measure fatigue in rheumatoid arthritis patients	163
Chapter 10	Summary and general discussion	197
	Summary in Dutch (Samenvatting)	207
	Acknowledgements (Dankwoord)	215
	Curriculum vitae	219
	List of publications	221

Chapter 1

General introduction

When you enter the term “human fatigue” into Google approximately 131.000.000 hits emerge within 0.17 seconds. This enormous amount of information indicates that fatigue is a topic that has considerable relevance for people. Basically, being tired is nothing to worry about. Everybody knows it as signal that body and mind need sleep or relaxation for recovery. Usually fatigue diminishes after taking adequate rest, body and mind are refreshed again. The Oxford dictionary defines fatigue as “a feeling of being extremely tired, usually because of hard work or exercise”.[1] However, for most persons with a chronic somatic condition such as rheumatoid arthritis (RA), the situation is different.

Aim of this thesis was to explore how fatigue is experienced in patients with RA as only very little is known about this phenomenon. Furthermore it was investigated how fatigue in RA can be measured adequately. The topic is thus twofold; meaning and measurement of fatigue in RA.

Rheumatoid arthritis and the experience of fatigue

RA is a chronic auto-immune disease that is characterized by inflammation of the joints.[2] Typical symptoms are pain, tender and swollen joints, stiffness, functional limitations and fatigue. Between 40% and 80% of the patients report substantial fatigue.[3-8] In fact, they mention fatigue as one of their most bothersome problems with RA.[9] Patients describe their fatigue as annoying, multidimensional symptom with far-reaching consequences for daily life.[10-13] It is different from usual fatigue or tiredness because it is more extreme and no longer earned what makes it unpredictable. Thus general descriptions and lay knowledge do not reflect the experience of patients with RA. Consequently, a different definition of fatigue is needed than for healthy people. However, an internationally accepted definition of fatigue in RA does not exist.[14] We do not yet know for sure whether there are differences between RA fatigue and fatigue in other somatic diseases. According to Hewlett there are similar multidimensional components and consequences of fatigue but the diseases differ significantly with regard to the attribution and interpretation of fatigue.[11] In cancer for example, fatigue is attributed to chemotherapy and therefore can be predicted and prepared for.[15] In MS, fatigue is perceived as exacerbating existing symptoms, whereas in RA existing symptoms are perceived to exacerbate fatigue.[16] Based on these differences we might conclude that a definition of fatigue in RA seems necessary and useful.

A definition without referring to previous activity is provided by Dittner et al.[17] They state that fatigue is an essentially subjective experience that can be described as “extreme and persistent tiredness, weakness or exhaustion—mental, physical or both”.[17] In our

interview study[13] that is reported in chapter 4, a woman described her experience of fatigue as follows:

"Yes, well you feel like you're carrying a load of flour on your back. That's what it feels like, that there's always something weighing your shoulders down, that you've got to lug that around with you."
(female patient with RA, 58 years old)

Although this quotation is representative for many persons with RA, not all of them experience fatigue in the same way. Nearly all patients mentioned a negative impact of fatigue on their lives,[13] but the severity varied. The quotation

"I think that if I hadn't been so tired, I would really have liked to have had children, but it's not possible. So, yes, the tiredness does stand in the way of a lot of things. My relationship, too. I think the tiredness is also the reason why that didn't work out."
(female patient with RA, 39 years old)

reflects a more serious impact than the following:

"I know that it [the fatigue] will pass quickly, so I'm not too bothered about it. And sometimes I think it might have to do with getting older."
(male patient with RA, 68 years old)

The causes for such inter-individual differences in fatigue among patients with RA are not yet fully understood. The same applies to intra-individual differences as variations in severity, frequency and duration of fatigue. A participant of the interview study[13] stated:

"It is very variable, the fatigue. At one time you have more energy than the other time but the energy will never be enough. The causes are not always clear, why you have more energy at one moment than at another."
(female patient with RA, 39 years old)

The perspective of patients

Most patients try to manage their fatigue by trial and error and report to receive no adequate medical or para-medical support.[18] The perception of patients regarding RA outcomes and their health state often differs from those of physicians and other health care providers, so it is of special importance to include the patients´ perspective in clinical care and research.[19] The limited attention that fatigue gets from clinicians might be explained by the complexity of the phenomenon, the lack of knowledge about its origins and the absence of treatment options. Furthermore it is not immediately clear what a patient means when he or she talks about fatigue. The lay use of the word "fatigue" lacks a clear meaning and can refer to different manifestations as exhaustion, muscle weakness, lack of energy, depletion of resources, apathy or depression.[20] Moreover, patients often use cues instead of talking about fatigue directly.[21] The perspective of patients is so important because fatigue is basically a subjective experience. It is not possible to gain knowledge about fatigue in RA without asking it to a patient; self-report is essential. This is vividly demonstrated in the following quotation from interview material:[13]

"Acceptation is the most difficult thing with fatigue; you have constantly to slow down. And that is fatiguing too. You cannot make clear to another person how it feels. A broken finger you can point at and explain that you have pain. That is visible and other people believe you. But when you say that you are tired, it is difficult to understand because it is not so easy to display. That makes me thinking: Well, I better do not tell them. When you take the whole package then the fatigue is the most annoying. You can bear down pain and you can talk about it more easily. But fatigue, that is not tangible."
(female patient with RA, 65 years old)

About ten years ago, the issue of fatigue was raised by patients themselves on OMERACT meetings (international group of experts on outcome measures in rheumatology) and thoroughly discussed.[22] Thereupon, fatigue was recommended to be included in the core set of outcome variables in clinical trials.[23]

The challenge in the assessment of fatigue is that there is no objective way to measure it. Research on fatigue will therefore always be indispensably connected to the perspective of the patient.[24] Measurement of fatigue is also related to the interpretation of words used in questionnaires. Cultural and linguistic aspects should be taken into account when differences between populations are intended to be examined.[25]

A precondition for internationally useable measurement instruments is that the experience of fatigue is comparable in different countries. The cross-cultural meaning of fatigue in patients with RA seems to be relatively uniform since interview studies on the experience of fatigue in the United States, the United Kingdom and the Netherlands revealed largely overlapping results.[10-13] However, it is not yet known how fatigue is experienced by patients with RA in non-western countries. The previously mentioned studies were conducted in relative similar cultural environments. It is of special importance to conduct research in different countries to find out more about the causes and impact of fatigue.

Multicausality and impact of fatigue

Knowledge about the mechanisms of fatigue in RA is still lacking. It is often assumed that levels of fatigue are highly associated with inflammatory processes. Anyway, conflicting evidence exists about the relation between fatigue and inflammatory markers of RA,[26] and fatigue turned out to be relatively stable over time,[27-30] also in phases of remission. Modern medical treatment has hardly beneficial impact on fatigue in RA,[31] as discussed in chapter 3 of this thesis.[32] Hewlett et al.[33] proposed a conceptual model reflecting interactions between disease processes (RA dimension), thoughts, feelings and behaviours (cognitive/ behavioural dimension) and personal issues in the life of a patient (personal dimension). However, a theoretically supported model, explaining the multicausal pathways of fatigue does not yet exist. To gain an overview about various factors associated with fatigue, the available scientific knowledge and empirical results were summarized in a systematic review. This review is described in chapter 2.[34]

Medical treatment for fatigue in RA, as described in chapter 3, indicate a relation between fatigue and inflammatory processes. In contrast, biopsychosocial models also include psychological and social aspects. We do not yet know which pathways are important for fatigue in RA. By conducting the systematic review (chapter 2), we intended to broaden the view from inflammatory markers to other aspects that are possibly related to fatigue. Fatigue is not only an annoying symptom for patients; it has relevance for society either. The societal impact of fatigue in RA becomes clear when examining the rates of work disability and its drivers. Fatigue turned out to be related to the work ability of employees with RA.[35] Work absenteeism in combination with the functional restrictions associated with the disease can lead to early retirement or work disability. Fatigue is also associated with participation restrictions in patients with RA;[36] they were impeded in social activities or hobbies by fatigue. Its impact can affect all areas of a patient's life but it is not yet

possible to provide an overview about statistically supported causes and consequences of fatigue in RA. Therefore it is of special importance to measure fatigue adequately.

Measurement of fatigue

The precise measurement of fatigue is essential for the evaluation of potential treatment effects.[37] It is also important for getting insight into causes and consequences of fatigue. Moreover, adequate measurement of fatigue is needed for diagnosis and screening purposes and facilitates the communication about fatigue between patient and professional.

Several uni- and multidimensional scales were developed to assess fatigue in clinical practice and research. Whereas unidimensional questionnaires are usually brief and provide a single score, multidimensional scales comprise a larger number of items and provide more detailed information that can give insight into different profiles and underlying mechanisms of fatigue.[17]

Fatigue measures with single item scales, such as visual analogue scales (VAS) or numerical rating scales (NRS), might have some value,[38] but do not correspond with the multidimensional character of fatigue as reported by patients.[10-13] In line with patients' experiences measurement of fatigue should be multidimensional, although it is not yet clear which and how many dimensions should be assessed.

Of the four multi-item fatigue questionnaires with reasonable evidence for validity in RA,[14] only the Multidimensional Assessment of Fatigue scale (MAF)[39] comprises several dimensions: severity, distress, timing and interference. For the Functional Assessment of Chronic Illness Therapy Fatigue Scale (FACIT-F),[40] separate scores for experience and impact can be calculated although it is usually applied unidimensionally. Also the Short Form 36 subscale vitality (SF-36)[41] and the Profile of Mood States subscale fatigue/inertia (POMS)[42] only have one dimension. Nicklin et al.[43] demonstrated that none of these four scales covers patient-reported concepts of fatigue comprehensively. Patients described fatigue in terms of frequency, duration, energy, sleep, cognition, coping, emotion, impact, social life, planning, relationships, and quality of life.[11,43] Moreover, none of these four instruments met all criteria for validity in patients with RA, so still further validation in this patient group is needed.[14] The main problem concerned content validity; a limitation of these traditional instruments is that the perspective of RA patients was not included during their development. In chapter 8, we describe how the items from traditional fatigue questionnaires are evaluated by patients, rheumatologists and nurses.[44]

Recently, the Bristol RA Fatigue Multi-Dimensional Questionnaire (BRAF-MDQ)[45] was developed from the patient's perspective and evaluated in a British RA population. With its four dimensions (physical, living, cognition, emotion), a score for each can be

calculated. The BRAF-MDQ is a promising instrument as it is based on the perspective of patients. However it is, even so as the other mentioned fatigue scales, a fixed-length questionnaire, meaning that each patient has to fill in the same items in the same order. This has the disadvantage that patients might be confronted with questions that do not apply to their individual level of fatigue.

In contrast, computerized adaptive testing (CAT) provides the possibility to measure patient reported outcomes with few items.[46] Items are respectively selected from an item bank based on a patient's previous answer so that precise measurement at individual level with few items becomes possible.[47] For the computerized selection of the best matching items, a large item pool is needed that contains more items than are presented to a patient. Before a computer-adaptive test (CAT) can be developed, such an item pool has to be scaled according to item response theory (IRT). With IRT, item parameters as the difficulty level can be assessed for each item independently.[47] This information is required to ideally match the items to the patient's individual level and for inter-individual comparisons on the measured construct even if patients filled in different items. Primarily, CATs were used for ability and achievement testing, but the interest in computerized adaptive testing for health-related measures is growing.[48] In this thesis, the steps to develop a calibrated item bank for the consecutive construction of a computerized adaptive test (CAT) are described (chapter 4-9).

Structure of this thesis

The first study that is presented in this thesis gives an overview of the available scientific knowledge about factors that are associated with fatigue in form of a systematic review (chapter 2). In the following chapter, the relevance of precise measurement of fatigue for the evaluation of potential treatment effects is discussed.

To gain more insight in the experience of fatigue from the patients´ perspective in depth-interviews were held with RA patients. This interview study is reported in chapter 4. After conducting this interview study we executed a Q-sort study to find out whether there are certain groups of patients who have a common perspective on the experience of fatigue (chapter 5).[49] Based on these two studies, and already existing fatigue questionnaires, we constructed a preliminary item pool for a comprehensive measurement. For the development of content valid items, it is essential to include the experience of patients and professionals in the field.[50,51] Only patients can report on the subjective experience of fatigue while clinicians have the most experience with the outward manifestation of fatigue.[52] Therefore in a Delphi study an expert panel of patients, rheumatologists and nurses evaluated the preliminary item pool as reported in chapters 6-8.[53,54,44] Result was a qualitatively evaluated item pool consisting of 245 items spread among 12 dimensions of

fatigue. In order to construct an item pool for the CAT fatigue in RA, its dimensionality structure was examined by factor analyses and the item pool was calibrated according to item response theory (IRT) (chapter 9).[55]

Finally, the thesis provides insights into meaning and measurement of fatigue in RA. First, an overview about factors related to fatigue and the impact of fatigue was provided in form of a systematic review. Then the meaning of fatigue for individual patients as well as groups of patients was investigated. The final product of the thesis is an initially calibrated item pool for the comprehensive and patient-friendly measurement of fatigue in RA. In its development the patients' perspective and modern psychometrics were combined to form the basis for the construction of a CAT fatigue in RA in a consecutive project.

REFERENCES

1. Hornby AS, Wehmeier S, McIntosh C, Turnbull J, Ashy M, eds. Oxford Advanced Learner's Dictionary. Oxford University Press, 2007:559.
2. Kvien TK, Scherer HU, Burmester G-H. Rheumatoid arthritis. In: Bijlsma JWJ ed. EULAR compendium on rheumatic diseases. London: BMJ Publishing Group and European League Against Rheumatism,2009:61-80.
3. Belza BL, Henke CJ, Yelin EH, Epstein WV, Gilliss CL. Correlates of fatigue in older adults with rheumatoid arthritis. Nurs Res 1993;42:93–9.
4. Wolfe F, Hawley DJ, Wilson K. The prevalence and meaning of fatigue in rheumatic disease. J Rheumatol 1996;23:1407-1417.
5. Katz PP. The stresses of rheumatoid arthritis: Appraisals of perceived impact and coping efficacy. Arthritis Care Res 1998;11:9-22.
6. Pollard LC, Choy EH, Gonzalez J, Koshaba B, Scott DL. Fatigue in rheumatoid arthritis reflects pain, not disease activity. Rheumatology 2006;45:885-889.
7. Repping-Wuts H, Fransen J, van Achterberg T, Bleijenberg G, van Riel P. Persistent severe fatigue in patients with rheumatoid arthritis. J Clinical Nurs 2007;16:377-383.
8. Van Hoogmoed D, Fransen J, Bleijenberg G, van Riel P. Physical and psychosocial correlates of severe fatigue in rheumatoid arthritis. Rheumatology 2010;49:1294-1302.
9. Hewlett S, Carr M, Ryan S, Kirwan J, Richards P, Carr A, et al. Outcomes generated by patients with rheumatoid arthritis: how important are they? Musculoskel Care 2005;3:131–42.
10. Belza Tack B. Fatigue in rheumatoid arthritis. Conditions, strategies, and consequences. Arthritis Care Res 1990;3:65–70.
11. Hewlett S, Cockshott Z, Byron M, Kitchen K, Tipler S, Pope D, et al. Patients' perceptions of fatigue in rheumatoid arthritis: Overwhelming, uncontrollable, ignored. Arthritis Care Res 2005;53:697–702.
12. Repping-Wuts H, Uitterhoeve R, van Riel P, van Achterberg T. Fatigue as experienced by patients with rheumatoid arthritis (RA): A qualitative study. Int J Nurs Stud 2008;45:995–1002.
13. Nikolaus S, Bode C, Taal E, van de Laar MAFJ. New insights into the experience of fatigue among patients with rheumatoid arthritis: A qualitative study. Ann Rheum Dis 2010;69:895–897.
14. Hewlett S, Hehir M, Kirwan JR. Measuring fatigue in rheumatoid arthritis: A systematic review of scales in use. Arthritis Rheum 2007;57:429–439.
15. Magnusson K, Moller A, Ekman T, Wallgren A. A qualitative study to explore the experience of fatigue in cancer patients. Eur J Cancer Care 1999;8:224-232.

16. Flesner G, Ek AC, Söderhamn, O. Lived experience of MS-related fatigue – a phenomenological interview study. Int J Nurs Stud 2003;40:707-717.
17. Dittner AJ, Wessely SC, Brown RG. The assessment of fatigue. A practical guide for clinicians and researchers. J Psychosom Res 2004;56:157–170.
18. Repping-Wuts H, van Riel P, van Achterberg T. Fatigue in patients with rheumatoid arthritis: what is known and what is needed. Rheumatology 2009;48:207-209.
19. Heller JE, Shadick NA. Outcomes in rheumatoid arthritis: incorporating the patient perspective. Curr Opin Rheumatol 2007;19:101-105.
20. Dupond J.-L. Fatigue in patients with rheumatic diseases. Joint Bone Spine 2011;78:156-160.
21. Repping-Wuts H, Repping T, van Riel P, van Achterberg T. Fatigue communication at the out-patient clinic of rheumatology. Patient Educ Couns 2008;76:57-62.
22. De Wit, M. From orphan outcome to darling pet – patient participation at OMERACT and the emergence of fatigue as an important outcome for research. Ann Rheum Dis 2010;69(Suppl3):23.
23. Kirwan J, Minnock P, Adebajo A, Bresnihan B, Choy E, de Wit M, et al. Patient perspective: fatigue as a recommended patient-centred outcome measure in rheumatoid arthritis. J Rheumatol 2007;34:1174-7.
24. DeLuca J ed. Fatigue as a window to the brain. Issues in clinical and cognitive neuropsychology. Jordam Grafman, series editor; 2005.
25. Lewis G, Wessely S. The epidemiology of fatigue: more questions than answers. J of Epidemiol Community Health 1992;46:92-97.
26. Stebbings S, Treharne GJ. Fatigue in rheumatic disease: an overview. Int J Clin Rheumatol 2010;5:487-502.
27. Belza BL. Comparison of self-reported fatigue in rheumatoid arthritis and controls. J Rheumatol 1995;22:639-643.
28. Stone AA, Broderick JE, Porter LS, Kaell AT. The experience of rheumatoid arthritis pain and fatigue: examining momentary reports and correlates over one week. Arthritis Care Res 1997;10:185-193.
29. Mancuso CA, Rincon M, Sayles W, Paget SA. Psychosocial variables and fatigue: A longitudinal study comparing individuals with rheumatoid arthritis and healthy controls. J Rheumatol 2006;33:1496-1502.
30. Treharne GJ, Lyons AC, Hale ED, Goodchild CE, Booth DA, Kitas GD. Predictors of fatigue over 1 year among people with rheumatoid arthritis. Psychol Health Med 2008;13:494-504.

31. Chauffier K, Salliot C, Berenbaum F, Sellam J. Effect of biotherapies on fatigue in rheumatoid arthritis: a systematic review of the literature and meta-analysis. Rheumatology (Oxford) 2012;51:60-68.
32. Nikolaus S, van de Laar MAFJ. Measuring fatigue in rheumatoid arthritis. Nat Rev Rheumatol 2011;7:562-4.
33. Hewlett S, Chalder T, Choy E, Cramp F, Davis B, Dures E, et al. Fatigue in rheumatoid arthritis: time for a conceptual model. Rheumatology 2011;50(6):1004-1006.
34. Nikolaus S, Bode C, Taal E, vd Laar MAFJ (in preparation for submission). Fatigue and factors related to fatigue in rheumatoid arhritis: a systematic review.
35. De Croon EM, Sluiter JK, Nijssen TF, Kammeijer M, Dijkmans BA, Lankhorst GJ et al. Work ability of Dutch employees with rheumatoid arthritis. Scand J Rheumatol 2005;34(4): 277-283.
36. Taal E, Hagens P, Braakman-Jansen L, van de Laar MAFJ. Perceived restrictions in participation in life situations among patients with rheumatoid arthritis. Arthritis Rheum 2009;60(Suppl10):2080.
37. Hewlett S, Dures E, Almeida C. Measures of Fatigue. Arthritis Care Res 2011; 63(S11):S263-S286.
38. Minnock P, Kirwan J, Bresnihan B. Fatigue is a reliable, sensitive and unique outcome measure in rheumatoid arthritis. Rheumatology 2009;48:1533-1536.
39. Belza Tack B. Dimensions and correlates of fatigue in older adults with rheumatoid arthritis. Dissertation, University of California; 1991.
40. Cella D, Lai JS, Stone A. Self-reported fatigue: one dimension or more? Lessons from the Functional Assessment of Chronic Illness Therapy Fatigue (FACIT-F) questionnaire. Support Cancer Care 2011;19:1441-50.
41. Ware JE Jr, Sherbourne CD. The MOS 36-Item Short-Form Health Survey (SF-36). I. Conceptual framework and item selection. Med Care 1992;30:473-483.
42. McNair D, Lorr M, Droppelman L. Profile of Mood States manual. New York: Multi-health Systems Inc; 1992.
43. Nicklin J, Cramp F, Kirwan J, Urban M, Hewlett S. Collaboration with patients in the design of patient reported outcome measures: Capturing the experience of fatigue in rheumatoid arthritis. Arthritis Care Res 2010;62:1552-1558.
44. Nikolaus S, Bode C, Taal E, van de Laar MAFJ (in press). Experts' evaluations of fatigue questionnaires used in rheumatoid arthritis – A Delphi study among patients, nurses and rheumatologists in the Netherlands. Clin Exp Rheumatol.
45. Nicklin J, Cramp F, Kirwan J, Greenwood R, Urban M, Hewlett S. Measuring fatigue in rheumatoid arthritis: A cross-sectional study to evaluate the Bristol Rheumatoid Arthritis

Fatigue Multi-Dimensional questionnaire, visual analogue scales, and numerical rating scales. Arthritis Care Res 2010;62:1559-1568.
46. Rose M, Bezjak A. Logistics of collecting patient-reported outcomes (PROs) in clinical practice: an overview and practical examples. Qual Life Res 2009;18:125–136.
47. Hambleton RK, Swaminathan H, Rogers HJ. Fundamentals of Item response theory. Thousand Oaks: Sage; 1991.
48. Walter OB. Adaptive tests for measuring anxiety and depression. In: WJ van der Linden, CAW Glas eds. Elements of adaptive testing. New York: Springer;2010:123–136.
49. Nikolaus S, Bode C, Taal E, van de Laar MAFJ. Four different patterns of fatigue in rheumatoid arthritis patients: results of a Q-sort study. Rheumatology 2010;49:2191-2199.
50. Streiner DL, Norman GR. Health measurement scales—a practical guide to their development and use. New York: Oxford University Press; 2003.
51. Fayers PM, Machin D. Quality of life—assessment, analysis and interpretation. Chichester: Wiley; 2000.
52. Yorkston KM, Johnson K, Boesflug E, Skala J, Amtmann D. Communication about the experience of pain and fatigue in disability. Qual Life Res 2010;19:243–251.
53. Nikolaus S, Bode C, Taal E, van de Laar MAFJ. Selection of items for a computer-adaptive test to measure fatigue in patients with rheumatoid arthritis – A Delphi approach. Qual Life Res. Published online 31 July 2011. doi: 10.1007/s11136-011-9982-8
54. Nikolaus S, Bode C, Taal E, van de Laar MAFJ. Which dimensions of fatigue should be measured in patients with rheumatoid arthritis? – A Delphi study. Musculoskeletal Care. Published online 11 November 2011. doi: 10.1002/msc.222
55. Nikolaus S, Bode C, Taal E, van de Laar MAFJ (in preparation for submission). Calibration of a multidimensional item bank to measure fatigue in patients with rheumatoid arthritis.

Chapter 2

Fatigue and factors related to fatigue in rheumatoid arthritis:
a systematic review

S. Nikolaus
C. Bode
E. Taal
M.A.F.J. van de Laar

In preparation for submission

ABSTRACT

Objective Patients with rheumatoid arthritis (RA) complain about fatigue. However, little is still known about causes and consequences of fatigue. A fully developed theoretical model explaining the experience of fatigue in RA is lacking. Goal of this study was to systematically review studies in RA that examined factors related to fatigue and differences in fatigue between RA patients and other patient groups, to gain more insight in possible causes and consequences of fatigue in RA.

Methods Four databases were searched for relevant studies; MEDLINE, Web of Science, SCOPUS and PsychInfo. All studies with RA samples about the relation between fatigue and other variables, that defined dependent and independent variables and used multivariate statistical methods, were included in the review. One hundred twenty nine studies were preliminary included. After reviewing the full-texts, we identified twenty-four studies on possible causes of fatigue, fifteen studies on possible consequences of fatigue and ten studies comparing levels of fatigue between groups.

Results Studies found possible causes of fatigue in illness-related aspects (e.g. pain), physical functioning (e.g. disability), cognitive/emotional functioning (e.g. depression) and social aspects (e.g. negative interpersonal events). Additionally, female gender was related to more fatigue. Inflammatory activity (e.g. ESR, DAS28) showed an unclear relationship with fatigue in RA. Possible consequences for fatigue were found among illness-related aspects (e.g. morning stiffness), physical functioning (e.g. physical quality of life), cognitive/emotional aspects (e.g. psychological distress) and social aspects (e.g. work ability). Patients reported higher severity of fatigue than healthy subjects and fibromyalgia patients reported worse levels of fatigue than other patient groups.

The most evidence for a relation between fatigue and other variables was found regarding pain, physical functioning and depression. Many cross-sectional and also longitudinal studies reported that these variables were associated with fatigue.

Conclusion This study gave an overview about variables that are related to fatigue in RA and information about fatigue levels in RA compared to other patient groups. However, most of the included studies were cross-sectional and not all longitudinal studies controlled for baseline fatigue so that hardly conclusions about causal relationships could be drawn.

Fatigue is commonly reported by patients with rheumatoid arthritis (RA).[1,2] Qualitative research has shown that patients experience fatigue as a multidimensional, annoying symptom with far-reaching consequences.[3-6] Primarily, fatigue is a subjective experience that can be described as 'extreme and persistent tiredness, weakness or exhaustion—mental, physical or both'.[7] A generally accepted definition of fatigue in RA does not exist,[8] and little is known about its aetiology.[9] The number of studies including fatigue as outcome measure has rapidly increased over the last years. However, a theoretical framework explaining the experience of fatigue in RA is lacking[10] and the phenomenon of fatigue is not yet described in detail.

Hewlett et al.[11] proposed a hypothetical model for fatigue in RA, suggesting interactions between different factors. The first factor, "RA", includes disease processes. The second factor, "cognitive/ behavioural", contains thoughts, feelings and behaviours. The third factor, "personal", is about personal issues in the life of a patient. The model vividly reflects the dynamic relations between fatigue and physical, psychological and environmental factors, but the authors did not provide evidence for the hypothesised relations so that the model remains on a heuristic level.

A recent overview paper about fatigue in rheumatic diseases[12] also underlines the multifactorial nature of fatigue by showing example evidence for predictors of fatigue in rheumatic diseases in longitudinal observation of routine care or following intervention; e.g. disease activity and severity, disability, pain, sleep disturbance, mood, self-efficacy, illness perceptions and coping. Our study has an added value to this overview paper as it is a systematic review. Moreover, not only studies about factors associated with fatigue indicating potential causes of fatigue but also studies indicating that fatigue has a potential impact on other variables will be included and results about comparing fatigue levels in different patient groups will be reported.

Although Hewlett et al.[11] proposed a bidirectional relationship between fatigue and many other concepts and variables, regarding this review we find ourselves in a 'causation dilemma'. The results of most studies on fatigue did not answer questions related to bidirectional causation. They did not pose research questions focused on bidirectional relationships and moreover did not use adequate designs to investigate the mutual influence of fatigue with other variables. A majority of studies merely report bivariate correlations. For this review we included studies which examined the relationships at least multivariately and which assumed a directional association between fatigue and the other constructs under consideration. In general, we report the directionality according the authors intentions of the original studies.

For the investigation of causality, studies have to use an adequate design. Due to the fact that most studies are cross-sectional, no causality can be examined. These studies either

statistically predict fatigue with other factors in a regression model or they predict other outcomes with fatigue. This gives insight into associations and "possible" causal relations, but no evidence for that. Studies with a longitudinal design do give some insight into causality, provided that adequate controls are conducted as control for baseline levels of the predicted outcome. Therefore we registered in this review whether a study is cross-sectional or longitudinal, and whether analyses are controlled for baseline levels of the predicted outcome.

Aim of this study was to systematically review existing scientific literature about the relations between fatigue and disease- and patient characteristics and environmental variables in RA to gain more insight into possible causes and consequences of fatigue. The following three research questions will be addressed based on the reviewed empirical studies: 1) What is reported about possible causes of fatigue in RA? 2) What is reported about possible consequences of fatigue in RA? 3) What is reported about differences in the level of fatigue between different groups (e.g. different diagnoses, patients and healthy controls)?

MATERIALS AND METHODS

Search strategy and study selection

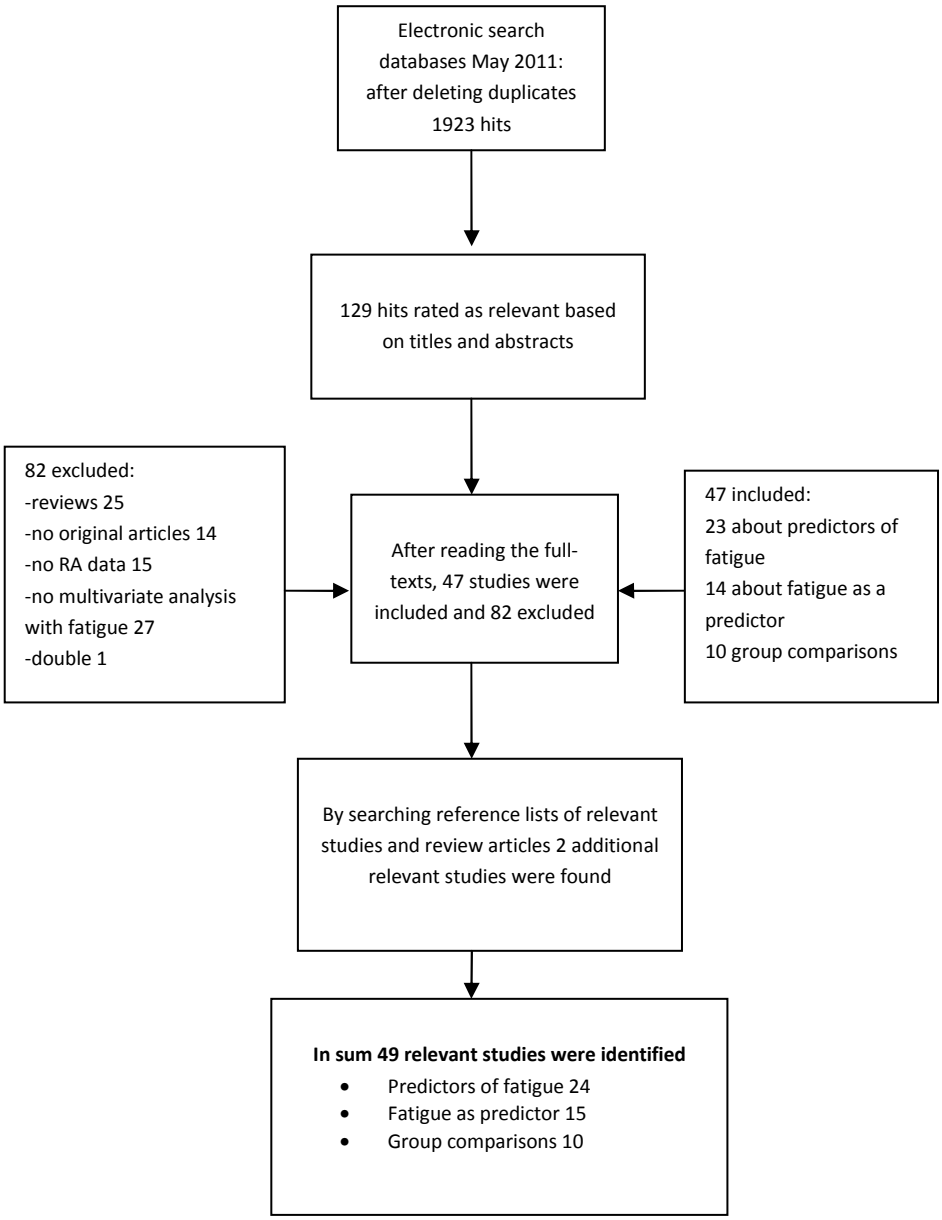
A systematic search of the literature was executed in MEDLINE, Web of Science, SCOPUS and PsychInfo. Main search terms were "fatigue / tiredness" in combination with "rheum* / arthritis / musculoskeletal / joint disease" and "model* / theor* / framework / predict* / etiology / pathophysiology / factor*". Whenever possible, proximity searches were used to make sure that the search terms, for example "fatigue" and "model", were mentioned in one sentence of a certain abstract. The detailed search strategies are included in the appendix.

The search was conducted in May 2011. All hits were saved in EndNote and duplicates were removed. After that, 1923 articles were present in the database. First, a researcher (SN) read all titles and abstracts and retrieved the potentially relevant articles. For our aim to summarize information about possible causes (statistical predictors) and consequences (fatigue as statistical predictor of another variable) of fatigue, we only included studies which defined dependent and independent variables and used multivariate statistical methods. So we did not include studies that merely provided correlations between variables. Case studies and qualitative studies were excluded. It was chosen to also exclude studies on effects of (medical) interventions, conference papers, letters, and papers in languages other than English. In cases where an abstract was not available or did

not give enough information to make a decision, the study was preliminary included so that the full-text could be screened. This procedure resulted in a preliminary set of 129 full-text articles.

These articles were read and summarized by three researchers (SN, CB, ET) and discussed in the team. Thereby consensus about the essential information of each article was obtained and agreement about the categorization of an article into one of the three research questions was reached. Question 1 was related to statistical predictors of fatigue (What is reported about possible causes of fatigue in RA?). In question 2, fatigue was referred to as statistical predictor (What is reported about possible consequences of fatigue in RA?). Finally question 3 asked for differences in fatigue severity between patient groups and / or healthy controls (What is reported about differences in the level of fatigue between different groups e.g. different diagnoses, patients and healthy controls?). On closer examination of the 129 abstracts/full-text manuscripts, it turned out that some studies did not fulfil our inclusion criteria and had to be excluded. Main reason for exclusion was that the sample did not include patients with RA or no data were provided for this group separately. Moreover some studies did not conduct relevant analyses with fatigue as variable and for some abstracts no original research article existed. Complementary to the electronic search, the reference lists of all 129 full-texts were searched for additional potentially relevant studies. The procedure of selecting relevant articles is shown in a flowchart (figure 1).

Figure 1 Flowchart of the selection of relevant studies from MEDLINE, Web of Science, SCOPUS, PsychInfo and searching reference lists



RESULTS

Description of studies

The selected studies were summarized in one of the three tables, according to our research questions. We identified 24 studies about statistical predictors of fatigue (What is reported about possible causes of fatigue in RA?), 15 studies wherein fatigue was the statistical predictor of another variable (What is reported about possible consequences of fatigue in RA?) and 10 studies with results about comparisons of fatigue between different groups (What is reported about differences in the level of fatigue between different groups e.g. different diagnoses, patients and healthy controls?)

Information extracted from the studies is summarized in tables 1 to 3. The tables consist of five columns. The first column includes the name of the first author and the year of publication. The second column contains information about the study design and the applied analyses. In the third column, the used measurement instrument for fatigue is shown. In the fourth column, the main results regarding our research questions are summarized briefly. This column differs per table. In the result column of table 1, the significant statistical predictors of fatigue are shown together with the most relevant statistical values and also a sub column is present for variables that turned out to be no significant statistical predictor of fatigue (insofar provided in the original studies). In table 2, the result column displays the variables that are statistically predicted by fatigue together with the most relevant statistical values and in a sub column variables are inserted that are not statistically significant predicted by fatigue. In table 3, the mean fatigue scores per group are provided together with the most relevant statistical values. In a sub column it is reported whether it was controlled for possible confounders in the comparison of fatigue between the groups. Some studies provided information for more than one research question. For a clear overview each study appears in only one of the three tables and the additional information is provided in the text.

In the next step, the findings of the studies were summarized briefly according to five categories:

1) illness-related characteristics: e.g. disease activity, pain, tender joint count, swollen joint count, radiographic damage
2) physical functioning: e.g. measures of disability, physical functioning, health related quality of life, quality of sleep
3) cognitive / emotional functioning: e.g. depression, anxiety, neuroticism
4) social / environmental aspects: e.g. work, roles, family, social support, life events
5) demographic aspects: e.g. gender, age

Table 1: Statistical predictors of fatigue (N=24)

First author, year publication	Sample (country, population, gender)	Study design		Measure of fatigue	Results	
					Statistical significant predictors of fatigue	**Statistically not significant predictors**
Belza, 1993[1)]	USA 133 RA patients 75 % female	Cross-sectional	Hierarchical multiple regression	Multi-dimensional Assessment of Fatigue (MAF)	Fatigue predicted by (beta weights) being female (-.36***), pain (.16*), quality of sleep (.31***), comorbid conditions (.15*), activity level (-.16*), functional status (.29**) and disease duration (-.14*) R^2 for demographic variables was .14, after adding disease-related variables .56 and after adding psychosocial variables .61 (adjusted R^2 for the model .57)	Education, age, depression, social support and helplessness
Bergman, 2009[2)]	USA 1) RA 14607 (22.7% male), OA 3173 (20.5% male), FM 2487 (4.7% male) 2) 1577 RA	Cross-sectional	Least-squares regression analysis	VAS fatigue (0-10)	1) No differences in fatigue between RA and OA, but higher in FM (M=4.5, 4.4, 6.3) 2) Regression analysis: Fatigue mainly explained by patient global (R^2=0.4)	Inflammatory activity, joint counts and ESR

Contreras-Yanez, 2010[3)]	Mexico 112 recent-onset RA patients 36 therapy-persistent 76 non-persistent 86.6% female	Longitudinal	Medical evaluations every 2, 4, or 6 months Group comparisons, correlational analyses, multivariate cox proportional hazard models	Presence / absence of substantial fatigue (not further defined)	-	Persistance on DMARDs, no other predictors either (socio-demographic variables, disease characteristics, PROs, physician-reported outcomes, laboratory outcomes, baseline treatment)
Davis, 2008[4)]	USA, 58 patients with RA	Cross-sectional	Multilevel modeling Hierarchical regression for association between fatigue and inflammatory markers over and above demographic behavioural and health-related factors	101-point NRS (0 no fatigue, to 100 fatigue as bad as it can be)	When pain included in initial step of the regression models no other covariate (age, BMI, etc) was significantly related to fatigue [R^2=0.4; b=0.7***] LPS-stimulated IL-6 level predicted fatigue over and above the contribution of pain [change in R^2 = 0.4; b = .21*]	Plasma levels of both CRP (r=-0.12*) and IL-6 (r=0.11*), age, sex, ethnicity, BMI, sleep disturbance alcohol use, steroid medication use, pain

Davis, 2010[5)]	USA 228 RA patients, 94 % women	Longitudinal 30 daily diaries 30 deviation scores for each variable	Descriptive statistics, comparisons between groups, multilevel and hierarchical linear regression models	101-point NRS (0 no fatigue, to 100 fatigue as bad as it can be)	Women more daily fatigue than men (men: M=29.6 SD=20.5, women: M=34.1 SD=23.8)** Same day-fatigue predicted by one or more comorbid pain conditions (beta=11.8***), days with more than average pain level (beta=.5***) and more than average number of negative interpersonal events (t=.4*), association between changes in positive interpersonal events and same-day fatigue varied by sex (-.36*), inverse relation in women but not in men Next day fatigue predicted by comorbid pain condition (beta=8.35***), same-day fatigue (.3***), pain (.1***), negative events (.6*) and interaction positive events and being female (.4*) Women positive events →less same day fatigue →more next day fatigue Relation negative events → fatigue is mediated by negative affect	For men no relations positive events with same day and next day fatigue No significant statistical predictors were age, education, income, employment status, marital / romantic relationship status, general health, daily pain
Dhir, 2009[6)]	North India 200 RA patients (30 FM, 170 without FM), mainly women 200 gender matched controls	Cross-sectional	Comparisons between groups, correlational analyses, linear regression	10 cm VAS fatigue about preceding week	FM was present in 15% of RA patients compared to 2.5% of controls RA patients with FM had more fatigue (M=3.9, SD=2.7 vs M=1.4, SD=2.3**) For RA, pain and fatigue were independently predicted by number of tender points and DAS28-3 (combined r=0.5*** and combined r=0.5***)	Not reported

Fifield, 2001[7]	USA 415 patients with RA, 83% female	10 year longitudinal study, fatigue assesed once per year from year 2 through 8 (so 7 years)	Descriptive analyses, hierarchical linear models, growth-curve analysis (to examine variations in initial fatigue levels and changes in fatigue during 7 years for patients with and without a history of affective disorder (AD)) history of affective disorder: fulfilling DSM-IV criteria for major depression (MD) or generalized anxiety disorder (GAD) at one time in the past as measured with Semi-Structured Assessment for the Genetics of Alcoholism, one interview	Question about fatigue: On a scale of 0 to 100, with 0 being no fatigue at all and 100 being the most fatigue possible, how much arthritis fatigue did you feel in the past week?	Initial fatigue scores were in middle range (mean initial fatigue =50, 95% confidence interval=47.1-52.3) and scores increased 5% per year (β01=mean growth rate = 0.5, p=0.3) individuals with AD history had higher levels of fatigue in first year and remained higher entire study compared to those without AD 3% of between-subject variance in initial fatigue due to history of AD (56% of total variance) effect of AD history on fatigue mediated through current distress History of AD → higher levels of distress aggregated over a full 7 years Higher levels of distress → higher initial fatigue in AD but less increase over time Fatigue is linked to history of AD (Fatigue 10% higher in subjects with AD, stable over 7 years and does not increase at faster time than in patients without AD) Changes in fatigue and changes in distress positively related within individuals	Co-morbidity, disease duration (RA)

Finan, 2010[8)]	USA 231 RA patients 70 male, 161 female	Longitudinal, cross-sectional analyses	30 daily diaries, multilevel modeling	Fatigue VAS (What number between 0 and 100 best describes your average level of fatigue today? 0 no fatigue, 100 fatigue as bad as it can be)	Daily increase in PIE (positive interpersonal events) results in decreased fatigue (β=-.6, SE=.1), F(1,6)=30.2***), Daily increase in NIE (negative interpersonal events) results in increased fatigue (β=.6, SE=.2), F(1, 6=13.1, **) Interaction PIE/NIE significant (R^2=.03) effect on fatigue Fatigue higher on days with few PIE or many NIE, reduction of fatigue only with combination of many PIE with few NIE (absence of NIE not enough to decrease fatigue) → Blunting hypothesis supported: reduction of fatigue related to days with many PIE diminished when patients also had many NIE Both PIE and NIE lead to more fatigue next day (β =.4, SE=.1), F(1,6)=19.5***) and (β =.6, SE=.2), F(1,6)=11**)	Neuroticism, extraversion were no moderator
Van Hoogmoed, 2010[9)]	NL 228 RA patients 63 % female	Cross-sectional	Multiple stepwise regression analysis	Checklist Individual Strength (CIS)	Predictors of fatigue (betas): age (-0.1***), RF (reumafactor, -0.2***), pain severity (0.1*), bodily pain (-0.2**) physical functioning (-0.1*), role functioning (-0.1*), Depressive mood (0.2**), self-efficacy on fatigue (-0.2***), coping (worrying, 0.2*), catastrophising (magnification of fatigue -0.1*), sleep disturbances (0.1**)	Gender, joint count, optimisme, self-esteem, social functioning, social support, physical activity

Huyser, 1998[10)]	73 RA patients 45% female	Cross-sectional	Principal component factor analysis, correlational analyses, Kruskal-Wallis tests, regression models with all possible combinations of predictor variables	Piper Fatigue Self-Report Scale (PFS)	Predictors of fatigue were pain (β=1.9***), depressive symptoms (β=2.6***), and female sex (β=9.9**) (adjusted R2 total model =0.5) entire model predictive of fatigue*** and explained 53% of the variance in PFS All other variables subsequently, one at a time, added to the model; the following contributed significantly (p<0.1) to increase in R2: longer symptom duration (beta=0.03 increase of 0.03), less perceived adequacy of social support (beta=-0.32, increase of 0.027), and less disease activity (beta=-1.6, increase of 0.02)	Many potential predictors included in analysis, e.g. anxiety, self-efficacy, coping, sleep, life stress, functional ability
Jump, 2004[11)]	USA 48 RA patients with prior affective disorder, 74 without prior affective disorder, more than 90% female	Cross-sectional	Differences between groups, hierarchical regression	Multi-dimensional Assessment of Fatigue (MAF) with GFI (global fatigue index)	27% without ADH (history of affective disorder) described fatigue as continuous, 45% of those with ADH did (X^2=6.2*) >50% reported that fatigue was not related to RA flares, 28% that fatigue occurred during flares (no difference between those with and without ADH) Patients with ADH rated fatigue, stiffness and pain as more serious than those without ADH. GFI higher for ADH group (34.4 vs 28.8) Hierarchical regression model to predict fatigue (R^2 whole model 0.27): ADH unique and independent contribution to the variance (8%, F(4,1)=3.1, *) when controlling for self-efficacy and neuroticism, self-efficacy and neuroticism also predictors (combined 17% F(6,1)=6.8, ***), but only SE uniquely predicted fatigue SE partly mediated relation between ADH and fatigue, Neuroticism does not sign. influence the relation between ADH and fatigue Fatigue predicted by a history of affective disorder, this relationship is mediated by self-efficacy	Control for age, education, illness duration (no betas reported)

Mancuso, 2006[12]	122 RA and 122 controls Follow-up about 14 months later 91 RA and 89 controls	Longitudinal	Multivariate regression analysis	Fatigue Severity Scale (FSS) at enrolment and one year later	In RA more fatigue than controls (M=4.2, SD=1.2 vs M=3.4, SD=1.1)***, also after 1 year (M=4.1, SD=1.3 vs M=3.2, SD=1.0) Cross-sectional MRA: For RA group: more fatigue associated with more anxiety (contributions to cumulative variance 0.3), more disability (0.1), less social support (0.02), and more social stress (0.04) ($p \leq 0.03$ for each variable, $R^2=0.5$). Enrolment variables associated with worse fatigue at follow-up, based on longitudinal multivariate regression analysis, were less help at home (contributions to cumulative variance 0.1), more anxiety (0.1), and more disability (0.2) ($p \leq 0.007$ for each variable, $R^2=0.3$).	Depressive symptoms, role satisfaction, physical activity, sleep quality at enrolment
Odegard, 2008[13]	Norway, RA patients 74.4% female N=550 at start N=216 later	10 year follow-up	ANOVA, ANCOVA, logistic regression	100 mm VAS fatigue	Gender difference in fatigue as measured 10 years after disease onset; women worse than men (women mean of 47.9 and men 33.3; 61% women > 40 mm and 35.8% of men)	-
Parrish, 2008[14]	US 89 RA, 76 OA, 90 FM Only female	Longitudinal 30 daily diaries	Multilevel analysis	VAS fatigue 0-100, 4 item fatigue PANAS-X	Fatigue higher in FM (M=55.8, SD=12.9) than in RA (M=33.5,SD=17.3) and OA (M=37.4,SD=19.4) Negative events predicted more fatigue in all patient groups Different reactions of patient groups on positive events; OA next day fatigue decreases, RA and FM next day fatigue increases	Always controlled for prior day fatigue

Pollard, 2006[15]	UK Sample 1: 228 RA Sex ratio 4 : 1 (female : male) Sample 2: 274 RA (3:1)	Cross-sectional	Descriptive statistics, group comparisons, simple linear regression followed by multiple linear regression fitted to all variables in stepwise manner	100 mm VAS SF-36 subscale vitality	Regression coefficients (beta): Sample 1: pain (0.5***), HAQ (7.2***), depression (11.0*), MTX (-8.1***), erosions (-7.5**) predicted fatigue (VAS) Sample 2: pain (0.4***), mental health (-0.4***), patient global (0.2*) predicted fatigue (VAS) and HAQ (2.0***), pain (1.0***) and mental health (1.0***) predicted energy and vitality (SF-36) Each 53% of variance in VAS fatigue explained	Excluded variables: DAS 28, TJC, SJC, physician global, patient global, ESR, CRP, age, sex, disease duration and other variables (especially regarding medication)
Repping-Wuts, 2007[16]	150 patients with RA, NL 137 at baseline and 123 follow-up	Longitudinal Follow-up design, with one-year duration	t-test, chi-squared test, univariate analysis, logistic regression analysis	Fatigue (CIS, baseline, one year) Two groups of patients (with and without severe fatigue based on CIS cut-off score)	Persistent severe fatigue predicted by disability (HAQ-DI, B=1.0)** at baseline and average general health (VAS-GH, B=0.7)*** Level of baseline fatigue related to level of fatigue at follow-up *** and remained constant for most patients	Starting model: ESR, average ESR, SW28, average SW28, TE28, average TE28, VAS-GH, average VAS-GH, HAQ-DI baseline, HAQ-DI follow-up

Riemsma, 1998[17)]	NL 229 RA patients, 61% female	Cross-sectional	Correlational analyses, regression analyses	Fatigue VAS, How tired were you on average during the past week due to your arthritis? (0 not tired at all, 100 very tired)	Fatigue is predicted by pain (beta=0.4***), self-efficacy toward coping with RA symptoms as pain, disability, depression (beta=-0.2**), self- efficacy expectations towards the mobilization of help (beta=-0.2**) and problematic social support (beta=0.1*) R^2 for model 0.37	Age, sex, education, disability, duration, laboratory, pain, social support
Scharloo, 1999[18)]	NL 71 RA patients without serious comorbidity 53 women	Longitudinal 2 time points of measure (baseline and follow-up about 1 year later)	Multiple regression	Tiredness over last week measured with VAS (0-100)	Baseline scores (tiredness at time 1, beta=0.4***) explained 40% of variance of in tiredness 1 year later. Perceived consequences of RA 8% (beta=0.2***), identity perceptions (belief that experienced symptoms as part of RA) 4% (beta=0.4***) and avoidant coping 3% (beta=0.2***) predictor of fatigue 1 year later.	Illness duration, disease severity, other illness perceptions and coping strategies
Sokka, 2009[19)]	6004 RA patients from multi-national cohort (QUEST-RA) 25 landen 79% female	Cross-sectional	Descriptive statistics, group comparisons, effect sizes	Fatigue measured with a VAS	Fatigue predicted by gender: Women had higher scores (poorer status) than men in all Core Data Set measures Mean fatigue was 4.6 for women versus 3.7 for men (***) in entire group, Cohen´s D: 0.3 (effect size of gender) Moreover examined gender differences in fatigue per number of swollen joints, gender differences were most pronounced in patients with low swollen joint counts	-

Stebbings, 2010[20]	New Zealand 103 RA and 103 OA patients RA: 29.1% male OA: 47.1% male	Cross-sectional study	Group comparisons, correlation analyses, multivariate linear regression analyses	MAF (GFI)	Fatigue OA>RA, CRP in RA>OA, in OA more persons with "moderate" fatigue (OA group had more severe disease than RA group) Difference in fatigue NOT significant between RA and OA after control for HAQ scores Regression analysis RA: predictors for fatigue: anxiety, depression (beta 0.8 and 1.1***)	Age, sex, disability, disease activity (DAS), erosive damage and pain, CRP, sleep
Stone, 1997[21]	N= 35 75% female	7 days (EMA measure 7 times per day)	Ecological momentary assessments (EMA) for better recalling of current states Ordinary least square (OLS) regression Repeated measures analysis of variance (RMANOVA)	7 point scale (0 not at all, 3 moderately)	OLS: higher fatigue variability significantly associated with more pain, fatigue, joint pain, muscle pain, stiffness on awakening, pain on awakening, swelling on awakening and poorer sleep quality No sign. differences in pain or fatigue by day of the week but strong effects of time of day for both (fatigue: F(1,1)=77.6,****), Fatigue moderate in morning, lowest between 10 am and noon, steep rise throughout rest of the day Higher pain levels on stressful days, but no sign. difference found for fatigue Sleep measured on daily basis: Fatigue not predicted by number of hours slept but by sleep quality (*)	Neuroticism, anxiety, depression, average number of hours sleep each night Sleep measured on daily basis: Fatigue not predicted by number of hours slept

Thyberg, 2009[22)]	Sweden n=320 (RA<1year) N=276 (191 women, 86 men)	Cross-sectional analyses Longitudinal data: measure at 3,6,12 months, then once per year to 8 years follow-up Here: follow-up 12, 24, 36 months (M12, M24, M36)	Correlations, differences between groups, principal component analysis (PCA), multiple linear regression analysis	VAS fatigue (0-100 mm)	At M12 and M36: Women more fatigue than men (M=40 vs M=30)*, age and fatigue only small correlations When comparing fatigue at M12 with M24 and M24 with M36, no differences, either in men or women - Underlying components (turned out same for men and women): I) physical disability: disease activity, Activity limit HAQ, Pain II) mental aspects: mental health, Sleep disturbance -MLRA: Women: fatigue explained more by I than by II at M 12 (beta 0.38 vs -0.31)but at M24 (0.33 vs -0.48) and 36 (0.34 vs -0.45), more by II than by I, age at inclusion was only significant predictor of fatigue at M36 (-0.17), R^2=0.46 Men: fatigue explained more by II than by I at M12 (beta 0.28 vs -0.44), at M24 I more than II (0.41 vs 0.28), at M36 still I more than II (0.62 vs ?)and age at inclusion was significant predictor (-0.30), R^2=0.47	-

Treharne, 2008[23]	UK 154 RA patients 74% female	Longitudinal (2 points of measuring: baseline and follow-up one year later)	Multiple hierarchical linear regression analysis Hierarchical regression of fatigue after 1 year:	Fatigue 100 mm VAS (anchors: no fatigue, unbearable fatigue) All variables measured at baseline and fatigue also at follow-up	Step1: fatigue baseline, Step2: demographics, Step3: disease impact, Step4: illness perceptions (consequences and self-efficacy), Step5: praying / hoping Baseline fatigue explained 13% of the variance in fatigue 1 year later*** In final model inflammation (ESR) (beta=-0.2*), perception of consequences of RA (beta=0.3*) and self-efficacy (beta=-0.3, p=0.06) were significant predictors, R^2=0.39 whole model	Age, sex, employment status, DMARD use, pain, impact of disability, sleep disruption, depressed mood, praying/ hoping coping
Wolfe, 1996[24]	N= 1488 (RA= 628; OA= 535; FM= 325) 81% female	Cross-sectional	Multivariate regression analysis, group comparisons	VAS fatigue (0-3)	Substantial, clinically important fatigue (>=2 on VAS) present in 42% RA 41% OA and 76% of FM RA predictors (standardized beta coefficient): Pain (0.2***), sleep disturbance (0.2***), depression (0.1***), tender points (0.1***), disability (0.2***) strongest independent predictors of fatigue in multivariate analysis, R^2 total = 0.49 In all 3 groups fatigue associated with pain, sleep disturbance, depression	Age, ESR, morning stiffness and sex

* = p ≤ 0.5, ** = p ≤ 0.01, *** = p ≤ 0.001

The studies which examined statistical predictors of fatigue revealed that concepts from all five categories (illness-related, physical functioning, cognitive/emotional and social/ environmental, demographics) showed significant independent relationships with the severity of fatigue as reported by patients with RA.

Within the category of illness-related variables, elevated pain most often came to the fore as complaint associated with increased levels of fatigue. This relation was supported by cross-sectional[1,4,9,10,15,17,22,24] and longitudinal studies[5,21] in that was controlled for previous fatigue levels. Only two studies reported that pain was not significantly related to fatigue; the cross-sectional study of Stebbings[20] and the longitudinal study of Treharne[23] in that was controlled for baseline fatigue.

Characteristics of inflammatory activity (e.g. ESR, DAS28, flares), however, showed an unclear relationship with fatigue in RA. In some studies these markers were significantly related to fatigue,[4,6,10,22] in other studies they did not contribute to the severity of fatigue at all.[2,20,24] All studies that found a significant relation between inflammation and fatigue were cross-sectional.[4,6,10,22] Davis[4] reported that LPS-stimulated IL-6 level predicting fatigue over and above the contribution of pain but fatigue was not related to plasma levels of both CRP and IL-6. In the study of Dhir[6] fatigue was significantly associated with DAS-28 while controlling for pain. In the analyses of Huyser[10] less disease activity was related to increased fatigue. The study of Thyberg[22] showed a significant association with fatigue by a cluster of disease activity, activity limitations and pain, labelled as physical disability.

In a longitudinal study that controlled for previous levels of fatigue, Stone found that higher fatigue variability was significantly associated with more stiffness and swelling on awakening[21] while Wolfe found in a cross-sectional analysis that the number of tender points were significantly related to fatigue in RA but morning stiffness had no important role in the regression model.[24]

Another illness-related factor that was significantly associated with fatigue was comorbidity as reported in a cross-sectional study[1] and a longitudinal study that controlled for previous fatigue levels.[5] Moreover, disease / symptom duration turned out to be significantly related to fatigue in two cross-sectional studies.[1,10] Contreras-Yanez[3] found no significant relation between the persistence on DMARDs and fatigue in a longitudinal study.

Aspects of physical functioning also contributed to the explanation of elevated levels of fatigue. Regarding quality of sleep or sleep disturbances significant associations with level of fatigue were found in four cross-sectional studies.[1,9,22,24] A longitudinal study[21] that

controlled for previous fatigue levels reported that fatigue was significantly related to sleep quality but not to the number of hours slept. Some studies found no support for significant relation between sleep and fatigue; two cross-sectional[10,20] and one longitudinal.[12] In some investigations also physical functioning, global health ratings and indications of disabilities were included. These characteristics were significantly related to fatigue in nearly all cross-sectional studies that included them in their analyses.[1,2,9,15,17,22,24] Only one cross-sectional study found no association between fatigue and disability.[20] Two longitudinal studies that controlled for baseline fatigue, also reported a significant relation between fatigue and disability.[12,16]

Within the cluster cognitive and emotional functioning the most often investigated construct in relation to fatigue was depression, operationalized as major depression or depressive mood. In several cross-sectional studies[9,10,11,15,17,20,22,24] and one longitudinal study,[7] depression was significantly associated with fatigue. Exceptions were two of the cross-sectional[1,12] and one of the longitudinal studies that controlled for baseline fatigue.[23] They did not find support for a significant relation between depression and fatigue. Regarding anxiety, contrasting results were reported. A cross-sectional study[20] found a significant relation between fatigue and anxiety, even so a longitudinal study controlling for baseline fatigue.[12] However the results of another cross-sectional[10] and longitudinal study controlling for previous levels of fatigue[21] did not support this relation. Perceptions such as self-efficacy also turned out to be related to fatigue. In cross-sectional studies[9,11,17] and a longitudinal study that controlled for baseline fatigue[18] significant associations were reported. In contrast, a cross-sectional study did not find support for a relation between self-efficacy perceptions and fatigue.[10]
For the cluster social and environmental aspects we found somewhat fewer studies than for the other characteristics. The reviewed cross-sectional[1,8,9,10,12,17] and longitudinal studies[5,12,14] that controlled for baseline or previous levels of fatigue, however, point to the importance of these characteristics for the explanation of fatigue. Negative interpersonal events for example were associated with higher levels of fatigue.[5,8,14] Adequate social support was significantly associated with less fatigue[12] and too less support or inadequate support (in the eye of the patient)[10,12,17] was significantly related to worse fatigue. Only two cross-sectional studies found that social support was not related to fatigue.[1,9]

In addition, age and gender were regarded as potential predictors of fatigue in most of the reviewed studies. A relation between fatigue and age of patients was not demonstrated, except in one cross-sectional study finding that younger patients reported more severe

fatigue.[9] Regarding gender of patients, studies showed consistent results. Female gender was significantly related to worse fatigue[1,10] and women reported higher levels of fatigue than men.[5,13,19,22] Also a study reported in table 2 found this pattern.[26] However, unfortunately these differences were not controlled for confounders in many studies. This seems necessary since the strength of gender differences in fatigue were less pronounced in patients with higher levels of disease activity.[19] The only study that did not find a gender difference was the longitudinal study of Belza;[40] women with RA did not report more fatigue than men with RA but healthy women did report higher levels of fatigue than healthy men.

Finally, longitudinal studies that included fatigue at baseline in the prediction model of fatigue at follow-up, reported that fatigue at baseline was significantly related to fatigue at follow-up one year later.[16,18,23]

Table 2: Fatigue as predictor (N=15)

First author, year publication	Sample (country, population, gender)	Study design		Measure of fatigue	Results	
					Fatigue predicts…	**Fatigue does not predict…**
Breedveld, 2005[25]	NL, 428 RA patients	Longitudinal 54 weeks measures at four week intervals from baseline to week 54	Stepwise multivariate linear regression and logistic regression analyses with disability as dependent variable To predict: physical function	One VAS (0-10)	Results baseline data (cross-sectional analyse at baseline): Linear regression model: Fatigue significantly related to disability (β = 0.1***) Logistic regression model: fatigue related to physical disability (HAQ ≥ 2): β= 0.2** OR = 1.2, OR 95% CI = 1.1 to 1.3 Longitudinal analysis: Multiple linear regression: baseline fatigue associated with physical disability at week 54 (β = 0.04*)	-
Chorus, 2003[26]	NL, 1056 RA patients (72.3% female) and 658 AS patients	Cross-sectional	Multiple stepwise regression, To predict: quality of life	MFI (general fatigue, physical fatigue, reduced motivation, reduced activity, mental fatigue), range 4-20	RA and AS: General fatigue worse in women than in men (scores of women and men in RA: 13.2 vs. 12.3, in AS: 13.4 vs 12.4) F=7.7***, no differences between the four groups (female RA and AS, male RA and AS) regarding other fatigue MFI General fatigue predicts physical (beta=-0.29***, total R^2=0.62) and mental (beta=-0.21***, total R^2=0.25) health related quality of life (physical and mental component summary scores of SF-36)	-

Covic, 2006[27)]	Australia 134 RA patients, 85% female 54 depressed 66 non-depressed	Cross-sectional	Pearson correlation coefficients, discriminant analysis with depression categories as dependent variable and 12 predictors as independent variables To predict: depression	VAS level of fatigue (10 cm line, 1 means no fatigue and 10 means extreme fatigue)	Fatigue predictor of depression, discriminant loading of 0.6 ($p<.01$) after tension, self-esteem and perceived RA impact	-
De Croon, 2005[28)]	NL, 78 working early RA patients, more than two-thirds were female	Cross-sectional	Logistic regression analysis To predict: work ability	Checklist Individual Strength (CIS) total score from 20-140, dichotomized in fatigued and non-fatigued patients by using cut-off point of >76	Fatigued employees with RA (N=17) report lower levels of work ability compared to non-fatigued employees (N=61) with RA [OR=24.0; 95% confidence interval (CI) 3.0-193.1] After adjustment for age, pain, DAS relation still significant	-

Davis, 2006[29]	USA 184 RA patients 52 men and 132 women	Longitudinal 30 daily diaries	Multilevel modelling, multivariate, multilevel, random effects regression analysis To predict: daily interpersonal events in pain patients	Fatigue NRS (0-100) "What number between 0 and 100 describes your average level arthritis fatigue today? A zero (0) would mean no fatigue and a one hundred (100) would mean fatigue as bad as it can be"	Fatigue is significant predictor (mean daily fatigue / current day´s fatigue) of negative daily events (beta .01* / .01**), psitive daily events (-.004, / -.01***), Daily relationship stress with friends (.002, 1.3 / .003, 12.3***), D. r. stress with family (.004, 4.9* / .001, .8), D. r. enjoyment with family (-.001, .3 / -.002, 3.9*), Enjoyment spouse (.00, .03 / -.002, 7.7**) at the same day Experience of intractable pain, fatigue, and disability explained only about 3% of the between-person variance in reports of both negative and positive events The amount of day-to-day variation in events accounted for by variations in pain and fatigue ranged from 1% for negative and 3.4% for positive events	Fatigue not predicts: Daily relationship enjoyment with friends, D. r. stress co-workers, D. r. enjoyment co-workers, D. r. stress with spouse
Katz, 1998[30]	USA 446 RA patients, 80% female	Cross-sectional	Multivariate regression analysis To predict: stressors of rheumatoid arthritis (e.g. impact of fatigue)	Fatigue 2 weeks prior, 1 item question; very mild, mild, moderate, severe, very severe (for analysis severe and very severe grouped together – dichoto-misation)	Fatigue severity predicts fatigue impact (beta=2.27***)	-

Parker White, 2009[31]	USA Mothers with MS, healthy mothers and 68 mothers with RA	Cross-sectional	Hierarchical regression analysis To predict: caregiving environment (the mother´s experience of daily hassles of parenting, discipline style she employed, how she monitored her child´s whereabouts)	21-item Modified Fatigue Impact Scale (FISK); physical fatigue, cognitive fatigue, psycho-social fatigue	Mothers with MS or RA had higher levels of fatigue than healthy mothers Higher levels of fatigue predicted both a greater frequency (standard β= 0.5**) and intensity (standard β =0.5**) of parenting daily hassles for mothers with RA mothers with RA who reported more fatigue also reported more difficulties monitoring their child (standard β=0.63***)	In RA fatigue did not explain significant variance in predictting laxness and over-reactivity
Rupp, 2004[32]	NL, 490 RA patients 72.7% female	Cross-sectional	Correlational analyses, 8 linear regression models with 8 dimensions of RAND-36 as dependent variables, and fatigue (physical fatigue, reduced activity, mental fatigue, and reduced motivation), RA-related pain and depressive symptoms as independent variables (general fatigue left out to reduce risk of (multi)collinearity) To predict: health-related quality of life	100-mm VAS ranging from 0 (no fatigue) to 100 (fatigue as bad as it could be) MFI-20 (general fatigue, physical fatigue, reduced activity, mental fatigue, reduced motivation)	HRQOL is predicted by fatigue; different aspects of fatigue selectively explain different dimensions of HRQOL while taking into account pain and depressive symptoms (beta p<0.05): physical fatigue: physical functioning (-2.2), social functioning (-1.2), role limitations physical (-2.7), vitality (-1.2), pain (-1.0), general health perceptions (-1.9) reduced activity: social functioning (-0.9), role limitations emotional (-1.2), vitality (-0.5), pain (-0.7) reduced motivation: vitality (-0.7) mental fatigue: role limitations physical (-1.1) and emotional (-1.0), mental health (-0.4) when entered last to the model, fatigue added 1 (mental health)-14% (vitality) to the explained variance of HRQOL	-

Smedstad, 1996[33)]	Norway, 238 RA patients, 73,5% female, 116 controls	Cross-sectional	t-tests and multiple linear regression analyses with GHQ (General Health Questionnaire, subscales for anxiety, depression, social dysfunction and somatization) as dependent variable and patient status (patient or control) and pain, disability and fatigue as independent variables To predict: psychological distress	Nottingham Health Profile (NHP): disability (8 items), pain (8 items), fatigue (3 items, score range 3-6, 3 being max. fatigue) Fatigue measure with NHP validated with Ritchie articular index (r=-.32)**	Fatigue higher in patients (4.8) than in controls (5.4)*** Differences between patients and controls regarding sum GHQ (mental distress) and all subscales no longer significant when controlling for fatigue	-
Waltz, 2000[34)]	136 Dutch RA patients, 98 German RA patients 70 % women	Longitudinal 3 points of measures (baseline, after 12 and after 24 months)	Examination of univariate associations Multivariate analysis with hierarchical regression model To predict: consumption patterns	Composite Index of Fatigue Impairment (CIFI) Symptoms: 11 point NRS and the 3 items of the NHP Sample split up into four groups according to level of fatigue (I low, IV high)	Fatigue contributes to physician consultations (change in R^2=0.04) and referral to a physical therapist in two years following baseline measures as assessed at 12 and 24 months	-

Wolfe, 2004[35])	USA 1) 825 RA patients 79.7% female 2) 12217 RA patients	Longitudinal At least 2 consecutive semi-annual measures	Ordered logistic regression, general estimating equations To predict: symptom importance in RA	Fatigue measured VAS with a range from 0-10, question: How much of a problem has fatigue or tiredness been for you in the past week? (anchors: Fatigue is no problem and fatigue is a major problem)	Patients ranking of symptom importance: 8% of the patients rated fatigue as the most important symptom above sleep problems but after functional status, pain, cognition and gastrointestinal symptoms In multivariate-ordered logistic regression –analysis and VAS fatigue, HAQ, pain, depression, age and sex as independent variables, only VAS fatigue predicted the ranked position for fatigue importance (OR 12., 95% CI 1.1 to 1.2,***), ranking of fatigue not related to use of anti-TNF agents or prednisone	-
Wolfe, 2008[36)]	USA 9921 RA patients (75.5% female), 2851 patients with noninflammatory rheumatic disorder (not FM, 79.6% female), 2867 FM patients	Cross-sectional	Logistic regression in univariate and multivariate analysis, Kendeall´s tau to examine association between combined dryness and RA symptom scales To predict: oral and ocular dryness	Fatigue VAS (0-10) alone Symptom intensity scale (SI): consisting of 1 VAS fatigue and regional pain scale (RPS)	Fatigue one of four predictors of dryness in RA (beside SI, pain and sleep scale) In RA (N=9921) SI strongest correlate of dryness, then RPS, then fatigue alone (mean 4.1, SD=2.9, Z-score 19.9***, Tau-a (95%CI: 0.08 (0.08 to 0.1)	-

Wolfe, 2009[37]	USA 20268 RA patients 1075 RA patients from RAES and ARCK datasets	Cross-sectional	Generalized linear model (GLM), Cox time-varying regression analyses Recursive partitioning to determine the best cutpoints for each clinical variable (very satisfied with health yes or no) using the RPART and Random forest programs To predict: RA outcome and treatment success and failure (satisfaction with health as surrogate for treatment success)	VAS fatigue (21 points from 0 to 10 at 0.5-unit intervals)	Patients without comorbidity who were very satisfied with their health (N=1774) and who were not very satisfied with their health (N=3958) differed on fatigue; not very satisfied: mean 4.1 and very satisfied: 1.5 (median 1.0) For fatigue the point of separating very satisfied and not satisfied states is <=1.7 (all patients) or <=1.3 (patients without comorbidity)	-
Wolfe, 2009[38]	USA 22131 RA patients, 3717 non inflamma-tory disease, 1002 SLE, 2674 FM	Cross-sectional	Several analyses to examine differences between groups and associations between variables, fractional polynominal logistic regression To predict: depression	Fatigue VAS Symptom intensity scale (SI): consisting of 1 VAS fatigue and regional pain scale	In RA, fatigue significantly different for depressed (Mean 6.5) and non-depressed (Mean 4.3) RA patients Fatigue and pain are important predictors of self-reported depression (especially correlated with SI score - SI scale best predictor of depression) in RA SI scale values for the quartiles for the SI scale are Q1 0-1.5, Q2 1.75-3, Q3 3.25-5 and Q4 5.25-9.75 (increasing quartiles associated with striking increase in cumulative risk of self-reported depression)	-

Yazici, 2004[39]	USA 337 patients with early RA 75% female	Cross-sectional and longitudinal design (patients with minimally two visits during 3,5 years).	Cross-sectional data analysis (ordered logit regression models) with: 5 confounders, disability, pain, patient global, fatigue, number of symptoms, swollen joint count, tender joint count, ESR Longitudinal analysis (random-effects logistic regression): age, disease duration, disability, pain, patient global, fatigue, number of symptoms To predict: morning stiffness	10 cm VAS fatigue	Analysis with cross-sectional data: regression model to estimate associations between 4-categories of morning stiffness and other variables: fatigue OR and 95% CI 1.3 (1.2,1.4) in univariate model and 0.9 (0.9,1.1) in multivariate model Longitudinal data: for fatigue OR 1.1 (1.0, 1.2)	-

In these 15 studies, fatigue turned out to be significantly related to several different variables.

Regarding illness-related aspects, it turned out that fatigue was significantly related to symptoms of ocular and oral dryness[36] and morning stiffness[39] in cross-sectional studies. Also functional aspects were associated with fatigue. A longitudinal study found that fatigue statistically predicted physical functioning[25] while controlling for baseline physical functioning. Two cross-sectional studies reported that fatigue was significantly related to physical quality of life[26,32] and that different aspects of health related quality of life were associated with different aspects of fatigue.[32]
Regarding cognitive/emotional aspects it turned out that fatigue was significantly related to mental health related quality of life,[26,32] and satisfaction with health[37] as reported by cross-sectional studies. A cross-sectional and a longitudinal study found that the level of fatigue statistically predicted the importance of fatigue as perceived by patients.[30,35] Two cross-sectional studies reported that fatigue was significantly associated with depression[27,38] and another that fatigue was associated with global psychological distress.[33]
The following social aspects were reported to be related to fatigue. A cross-sectional study found that fatigue was a significantly related to work ability.[28] In a longitudinal study[29] fatigue turned out to be related to same-day negative daily events, positive daily events, stress with friends, stress with family, enjoyment with family, enjoyment with the spouse but not to enjoyment with friends, stress or enjoyment with coworkers and stress with the spouse.[29] A cross-sectional study about the impact of fatigue on parenting found that fatigue was related to greater frequency and intensity of parenting daily hassles, and having enough energy to monitor a child's' whereabouts.[31] Fatigue was not significantly associated with laxness and over reactivity in mothers with RA.[31] Additionally, a longitudinal study found that fatigue was significantly associated with physician consultations, and referral to a physical therapist during two years.[34]

In conclusion, fatigue in RA is related to as well illness-related aspects as functional, cognitive/emotional and social aspects.

Table 3: Comparisons of levels of fatigue between patient groups and healthy subjects (N=10)

First author, year publication	Sample (country, population, gender)	Study design		Measure of fatigue	Results	
					Comparison between groups	Confounders
Belza, 1995[40]	USA N= 97: with RA n= 51, 85% female; without RA n= 46, 91% female	Longitudinal	Question-naires completed 3 times in 6-8 week-intervals, repeated measures of variance (ANCOVA)	Multidimensional Assessment of Fatigue (MAF) GFI (global fatigue index)	- Women with RA did not report more fatigue than men (healthy women reported higher fatigue than healthy men) -Fatigue scores calculated for each time point: Patients have higher mean scores on each dimension of fatigue than controls*** (severity 5.4 vs 3.2, degree 6.3 vs 4.2, distress 4.9 vs 3.0, impact 4.2 vs 2.6, timing 2.9 vs 2.1)	-
Crosby, 1991[41]	11 females, 4 males with RA (5 in disease flare, 10 either in remission or disease was mildly active) 12 controls	Cross-sectional	Descriptive statistics, analysis of variance	10 cm VAS fatigue (anchors 0% = no fatigue and 100% = extremely fatigued)	No significant differences in fatigue between controls (M=6.0, SD=2.5), RA flare (M=7.4, SD=2,4) and RA non-flare (M=5.1, SD=2.8) (ANOVA (F(2,2)=1.4,p=0.3)	-

Currey, 2003[42]	USA, N=619 RA 334 FM 152 OA 133 RA 71% vrouw	Survey (part of longitu-dinal study)	Univariate analyses and multiple linear regression models	10 cm VAS fatigue to assess degree to which unusual fatigue or tiredness has been a problem over the past week on a scale of 0 = no problem and 10 = a major problem	Fatigue in FM 6.4, RA 4.9, OA 4.9, FM significant more fatigue than RA and OA, no difference in fatigue between RA and OA -parameter estimates for differences between means by diagnosis: FM vs RA (beta 1.5***), FM vs OA (beta 1.6***), RA vs OA (beta 0.1 p>0.5)	Covariates: sex, age, education, disease duration and number of chronic conditions
Goodchild, 2008[43]	UK 34 SS 48 RA	Cross-sectional	One-way MANOVA in multiple fatigue scores between disease groups	Profile of Fatigue (ProF) Multidimensional Fatigue Inventory (MFI)	Patient groups did not differ in levels of fatigue on ProF or MFI	-
Goodchild, 2010[44]	UK, women, 25 with RA 14 with pSS	Longitudinal Diary study, 35 consecutive days	Repeated measures analyses of variance (RMANOVA), multilevel analysis	Profile of Fatigue (somatic and mental fatigue), 4 times a day	Patient groups did not differ in somatic or mental fatigue In both was an increase as the day progressed in the severity of both kinds of fatigue (somatic: F(1.7, 64.4)=25.4*** and mental: F(2.1, 78.9)=12.1***), time course of fatigue across day did not differ per group	-

Van Oers, 2010[45]	NL, 29 Sjögren, 23 SLE, 19 RA, 52 healthy controls	Measures at 8 time points /day, 2 consecutive days	Analyses with Kruskal-Wallis one-way analysis of variance by ranks, repeated measures analyses of variance	MFI: 5 most characteristic items (general fatigue, physical fatigue, reduced activity, reduced motivation, mental fatigue)	3 patient groups significantly higher levels of fatigue in comparison with healthy controls (no differences between patient groups) General fatigue: pSS M=2.3 IQR (1.9-3.4), SLE M=1.9 IQR (1.4-2.7), RA M=2.3 (0.8-2.9), healthy M=0.2 IQR (0.0-0.6) same pattern found for the other four fatigue items	-
Reading, 2009[46]	USA 164 RA patients 328 persons without RA (matched for age, sex) 72% female in both groups	Cross-sectional	Chi-squared tests to compare the two groups, logistic regression	Berlin Sleep Questionnaire to measure risk for sleep apnea (snoring, fatigue, co-morbidities)	More patients had fatigue as risk factor (38%) than subjects without RA (13% ***) More RA patients (50%) were at risk of sleep apnea than controls (31% ***)	Difference in fatigue remained after adjustment for smoking status, BMI and high blood pressure
Salaffi, 2009[47]	Italy 380 FM patients (92% female) 693 RA patients (91% female) 1579 healthy controls (50% female)	Cross-sectional	Between-group comparisons (ANOVA), multivariate linear regression models	VAS fatigue; what number between 0 and 10 best describes the average level of fatigue you experienced last week (0=no fatigue; 0=fatigue as bad as it can be),Vitality scale SF-36	FM mean 40.1 (SD=12.3) RA mean 43.4 (SD=20.6) healthy controls: mean 71.6 (SD=20.1) Difference between patients and healthy controls were significant ***	Gender differences only in the FM group (women reported worse health)

Wolfe, 2004[48]	11866 RA patients 1731 also fulfilled criteria for fibro-myalgia	Cross-sectional	Regression analyses	VAS fatigue	RA patients with FM had substantially more abnormal scores on all disease activity and status variables Regarding fatigue: Mean (SD) 3.8 (2.6) in those without FM and 8.1 (1.1) in those with FM Fatigue is higher in RA patients who also satisfy the criteria for FM	-
Zautra, 2007[49]	USA 89 RA, 76 OA, 90 FMS patients Only women	32 daily diaries	Multilevel analysis	Fatigue with one single item (number between 0 and 100)	FMS fatigue levels significantly higher than in RA and OA Fatigue levels OA 37.4, RA 33.5 and FMS 55.8, FMS patients more day-to-day variation in fatigue than RA and OA	-

The 10 studies shown in table 3 did not provide information about relations of fatigue with other variables, but about differences in fatigue between groups (patients with different diagnoses, patients and healthy controls).

Rheumatoid arthritis (RA) and healthy controls:

Belza[40] found that patients with RA reported higher levels of fatigue than controls. The studies of Mancuso[12] (table 1), van Oers,[45] Reading,[46] and Salaffi[47] also showed higher fatigue levels in patients with RA than in healthy controls and mothers with MS or RA reported higher levels of fatigue when compared to healthy mothers[31] (table 2). In contrast, Crosby[41] reported no differences between patients with RA in flare, those not in flare and healthy controls. However these results are based on very small samples so that it is doubtable whether conclusions about differences between groups can be drawn from this study.

Rheumatoid arthritis (RA), fibromyalgia (FM) and osteoarthritis (OA):

Fibromyalgia (FM) patients reported significantly more fatigue than those with rheumatoid arthritis (RA) and osteoarthritis (OA), while no difference in fatigue was found between patients with RA and OA.[42,49] Studies that are summarized in table 1 found comparable results.[2,14,24] In the study of Salaffi,[47] patients with FM reported less vitality than patients with RA.

Wolfe[48] found that fatigue reports were higher in RA patients who also satisfy the criteria for FM than in those with RA only. That is in line with the findings of Dhir,[6] a study displayed in table 1.

Stebbings[20] found that fatigue was reported as more severe in OA than in RA. However it has to be noted that the difference in fatigue did not remain significant after controlling for HAQ scores and the OA group had more severe disease parameters than the RA group.

Rheumatoid arthritis (RA), Sjögrens´ syndrome (SS) and systemic lupus erythematodes (SLE):

Patients with RA and SS did not differ on levels of fatigue.[43] In another study of Goodchild[44] it turned out that RA and SS patients did not differ in somatic or mental fatigue either. In both groups the severity of both kinds of fatigue increased as the day progressed, and the time course of fatigue across the day did not differ per group.[44]

Van Oers[45] reported about comparable levels of fatigue between patients with RA, SS and SLE. All those patient groups scored higher on fatigue than healthy controls.

In conclusion, most studies found that patients report higher levels of fatigue than healthy controls and that FM is the patient group that has the worst fatigue outcomes compared to other patient groups. RA patients had comparable levels of fatigue as patients with OA, SS and SLE.

DISCUSSION

This systematic review gave an overview of the existing literature on factors that are related to fatigue in RA. We systematically summarized the included studies according to our research questions: 1) What is reported about possible causes of fatigue in RA? 2) What is reported about possible consequences of fatigue in RA? 3) What is reported about differences in the level of fatigue between different groups (e.g. different diagnoses, patients and healthy controls)?

In conclusion, regarding question 1, we found the following factors: illness-related aspects (pain), physical functioning (sleep quality, sleep disturbances, global health ratings, indicators of disability), cognitive / emotional functioning (depression or depressive mood, anxiety, perceptions such as self-efficacy) and social / environmental aspects (interpersonal events, too less or inadequate social support) and female gender.

Not all illness-related aspects showed a clear relationship with RA. Regarding characteristics of inflammatory activity (e.g. ESR, DAS28, flares), inconsistent results were reported.

Regarding question 2, the following factors were reported by the reviewed studies: illness-related aspects (ocular and oral dryness, morning stiffness), physical functioning (physical quality of life), cognitive / emotional functioning (depression, global psychological distress, mental health related quality of life and satisfaction with health) and social / environmental aspects (work ability, negative and positive daily events, parenting). Fatigue also statistically predicted physician consultations and referral to a physical therapist.

Regarding the third research question, it came out that most studies found that patients reported higher levels of fatigue than healthy subjects and that FM is the rheumatic disease with the worst fatigue outcomes compared to other forms of arthritis.

Some of the factors came forward in answering the first as well as the second research question. That means that some studies reported that they statistically predicted fatigue and other studies found that they were predicted by fatigue. This applied for factors related to the aspects physical functioning, cognitive / emotional functioning (ratings of

health, depression, mental health / psychological distress) and social / environmental aspects (interpersonal events). This brings us back to the previously mentioned causation problem. Based on the reviewed studies, it is difficult to draw conclusions about causal relationships between fatigue and other variables. Most studies were cross-sectional and some of the longitudinal studies had shortcomings as not controlling for baseline fatigue or using small or selected samples.

The most convincing support was found for the prediction of fatigue by pain, disability / physical functioning and depression / depressive mood. For each of the three, more than ten of the reviewed studies reported significant relationships with fatigue. Among them were cross-sectional but also longitudinal studies. Regarding pain, two longitudinal studies that controlled for previous levels of fatigue, found a significant association with fatigue while one longitudinal study, that controlled for baseline fatigue, did not support this finding. Physical functioning was significantly associated with fatigue in two longitudinal studies that controlled for baseline fatigue. One longitudinal study found support for a significant relation between fatigue and depression while another did not find this. In conclusion, these three variables could be regarded as potential causes for fatigue in RA while for pain and physical functioning more convincing support was found than for depression.

However, for the closer examination of causality, it would be necessary to conduct prospective, longitudinal studies in representative samples. These studies should use an adequate measurement instrument of fatigue and should apply multivariate analysis techniques with check for confounders. Not all of the reviewed studies clearly described whether they checked for confounders and if so for which confounders. Moreover, in more than half of the reviewed studies fatigue was only measured with a single visual analogue scale (VAS) or numerical rating scale (NRS). That does not correspond to the multidimensional character of fatigue as reported by patients.[14,15] Only few studies used a fatigue questionnaire containing more than one dimension; those reported in more than one study were the Multidimensional Assessment of Fatigue (MAF) (e.g. *1, 20*), Checklist Individual Strength (e.g. *16,9*) and Multidimensional Fatigue Inventory (e.g. *26, 43*). It has to be noted that only one of these instruments, the MAF, is an RA specific questionnaire and some items of the other two scales could be related to disability or disease activity in RA.[13]

Furthermore, it would be important to measure other variables appropriately. For example regarding sleep disturbances, it turned out that patients are not aware of their actual number of awakenings during the night.[16]

We decided not to include a numerical quality assessment of the studies that are included in this review as it turned out to be very difficult due to the diversity of studies, different

measurements of fatigue and a lack of guidelines or standards for quality assessments of studies other than intervention studies as the Cochrane guideline.[17]

A limitation of this review is that it did not account for positive publication bias. Possibly, studies that did not find associations between fatigue and other variables were not published and therefore not identified by our search strategy. Moreover we cannot control whether some studies only reported the significant results and not the non-significant ones. Furthermore, we had to use proximity searches, meaning that the search in the databases was constrained to studies including a certain combination of search terms within a relatively small area of words in the abstract. Otherwise it would have been impossible to conduct this review with manageable efforts because the titles and abstracts to read would have been far more than the 1923 hits received now. However, we thoroughly checked the reference lists of the included studies for additional relevant studies.

The strength of this study is that it provides an overview of the state of the art regarding research about fatigue and factors related to fatigue in RA. Unfortunately we cannot draw solid conclusions about causal relationships because most of the included studies were cross-sectional and the longitudinal studies had some shortcomings either. When comparing our findings with the hypothetical model proposed by Hewlett et al.[11] it is obvious that we found support for relations between fatigue and all dimensions in the model. On a more detailed level, it turns out that the strongest support was found for relations between fatigue and variables on the RA dimension; disability, pain and sleep. However, the proposed relations between fatigue and the illness related aspects cortisol response, inflammation, joint damage, muscle effort and de-condition, drugs and anemia did not match well with the reviewed literature. On the dimensions “cognitive, behavioural” referring to behaviours, emotions and cognitions, the association between fatigue and depression was most frequently supported in our review. Regarding the other elements of this dimension (e.g. illness beliefs and stress, activity) and also the “personal” dimension referring to work/caring responsibilities, environment, health and social support, indications for the existence of relations with fatigue were also found in the literature, but by clearly less studies than those reporting about associations between fatigue and pain, disability and depression. Nevertheless, this review is no test for the causality of the considered bivariate associations in the model of Hewlett et al. It is a challenge for the future to conduct research that is able to give more insight into causes and consequences of fatigue in RA.

ACKNOWLEDGEMENT

We want to thank Dr. Marjolein Drent for supporting the development of the search strategy.

REFERENCES

1. Wolfe F, Hawley DJ, Wilson K. The prevalence and meaning of fatigue in rheumatic disease. J Rheumatol 1996;23:1407-1417.
2. Hewlett S, Carr M, Ryan S, Kirwan J, Richards P, Carr A, et al. Outcomes generated by patients with rheumatoid arthritis: how important are they? Musculoskel Care 2005;3:131-142.
3. Belza Tack B. Fatigue in rheumatoid arthritis. Conditions, strategies, and consequences. Arthritis Care Res 1990;3:65-70.
4. Hewlett S, Cockshott Z, Byron M, Kitchen K, Tipler S, Pope D, et al. Patients' perceptions of fatigue in rheumatoid arthritis: overwhelming, uncontrollable, ignored. Arthritis Care Res 2005;53:697-702.
5. Repping-Wuts H, Uitterhoeve R, van Riel P, van Achterberg T. Fatigue as experienced by patients with rheumatoid arthritis (RA): a qualitative study. Int J Nurs Stud 2008;45:995-1002.
6. Nikolaus S, Bode C, Taal E, van de Laar MAFJ. New insights into the experience of fatigue among patients with rheumatoid arthritis: a qualitative study. Ann Rheum Dis 2010;69:895-897.
7. Dittner AJ, Wessely SC, Brown RG. The assessment of fatigue. A practical guide for clinicians and researchers. J Psychosom Res 2004;56:157–70.
8. Hewlett S, Hehir M, Kirwan JR. Measuring fatigue in rheumatoid arthritis: a systematic review of scales in use. Arthritis Rheum 2007;57:429–39.
9. Repping-Wuts H, van Riel P, van Achterberg T. Fatigue in patients with rheumatoid arthritis: what is known and what is needed. Rheumatology 2009;48:207-209.
10. Hewlett S, Nicklin J, Treharne GJ. Fatigue in musculoskeletal conditions. Topical Reviews Series 6: No 1. Arthritis Campaign: Autumn 2008 www.arc.org.uk/ arthinfo/medpubs/6641/6641.asp (1 December 2009, date last accessed).
11. Hewlett S, Chalder T, Choy E, Cramp F, Davis B, Dures E, et al. Fatigue in rheumatoid arthritis: time for a conceptual model. Rheumatology 2011;50:1004-1006.
12. Stebbings S, Treharne GJ. Fatigue in rheumatic disease: an overview. Int J Clin Rheumatol 2010;5(4):487-502.
13. Hewlett S, Dures E, Almeida C. Measures of fatigue. Arthritis Care Res 2011;63:S263-S286.
14. Nikolaus S, Bode C, Taal E, van de Laar MAFJ. Which dimensions of fatigue should be measured in patients with rheumatoid arthritis? – A Delphi study. Musculoskeletal Care. Published online 11 November 2011. doi: 10.1002/msc.222

15. Hewlett S, Dures E, Almeida C. Measures of Fatigue. Arthritis Care Res, 2011;63,S263-S286. Doi 10.1002/acr.20579.
16. Mahowald MW, Mahowald ML, Bundlie SR, Ytterberg SR. Sleep fragmentation in rheumatoid arthritis. Arthritis Rheum 1989;32:974-983.
17. Higgins JPT, Green S (ed.). Cochrane handbook for systematic reviews of interventions. Version 5.0.2. John Wiley & Sons Ltd; 2009.

List of included studies

1) Belza BL, Henke CJ, Yelin EH, Epstein WV, Gilliss CL. Correlates of fatigue in older adults with rheumatoid arthritis. Nursing Research 1993;42(2): 93-99.

2) Bergman MJ, Shahouri SS, Shaver TS, Anderson JD, Weidensaul DN, Busch RE, Weng S, Wolfe F. Is Fatigue an Inflammatory Variable in Rheumatoid Arthritis (RA)? Analyses of Fatigue in RA, Osteoarthritis, and Fibromyalgia. J Rheumatol 2009;36:2788-2794.

3) Contreras-Yanez I, Cabiedes J, Villa AR, Rull-Gabayet M, Pascual-Ramos V. Persistence on therapy is a major determinant of patient-, physician- and laboratory-reported outcomes in recent-onset rheumatoid arthritis patients. Clin Exp Rheumatol 2010;28:748-751.

4) Davis MC, Zautra AJ, Younger J, Motivala SJ, Attrep J, Irwin MR. Chronic stress and regulation of cellular markers of inflammation in rheumatoid arthritis: Implications for fatigue. Brain Behav Immun 2008;22:24-32.

5) Davis MC, Okun MA, Kruszewski D, Zautra AJ, Tennen H. Sex Differences in the Relations of Positive and Negative Daily Events and Fatigue in Adults With Rheumatoid Arthritis. Journal of Pain 2010;11:1338-1347.

6) Dhir V, Lawrence A, Aggarwal A, Misra R. Fibromyalgia is common and adversely affects pain and fatigue perception in North Indian patients with rheumatoid arthritis. J Rheumatol 2009;36:2443-2448.

7) Fifield J, McQuillan J, Tennen H, Sheehan J, Reisine S, Hesselbrock V, Rothfield N. History of affective disorder and the temporal trajectory of fatigue in rheumatoid arthritis. Ann Behav Med 2001;23:34-41.

8) Finan PH, Okun MA, Kruszewski D, Davis MC, Zautra AJ, Tennen H. Interplay of Concurrent Positive and Negative Interpersonal Events in the Prediction of Daily Negative Affect and Fatigue for Rheumatoid Arthritis Patients. Health Psychology 2010;29:429-437.

9) Van Hoogmoed D, Fransen J, Bleijenberg G, van Riel P. Physical and psychosocial correlates of severe fatigue in rheumatoid arthritis. Rheumatology 2010;49:1294-1302.

10) Huyser BA, Parker JC, Thoreson R, Smarr KL, Johnson JC, Hoffman R. Predictors of subjective fatigue among individuals with rheumatoid arthritis. Arthritis Rheum 1998;41:2230-2237.

11) Jump RL, Fifield J, Tennen H, Reisine S, Giuliano AJ. History of Affective Disorder and the Experience of Fatigue in Rheumatoid Arthritis. Arthritis Rheum: Arthritis Care Res 2004;51:239-245.

12) Mancuso CA, Rincon M, Sayles W, Paget SA. Psychosocial variables and fatigue: A longitudinal study comparing individuals with rheumatoid arthritis and healthy controls. J Rheumatol 2006;33:1496-1502.

13) Odegard S, Kvien TK, Uhlig T. Incidence of clinically important 10-year health status and disease activity levels in population-based cohorts with rheumatoid arthritis. J Rheumatol 2008;35:54-60.

14) Parrish BP, Zautra AJ, Davis MC. The Role of Positive and Negative Interpersonal Events on Daily Fatigue in Women With Fibromyalgia, Rheumatoid Arthritis, and Osteoarthritis. Health Psychology 2008;27: 694-702.

15) Pollard LC, Choy EH, Gonzalez J, Khoshaba B, Scott DL. Fatigue in rheumatoid arthritis reflects pain, not disease activity. Rheumatology 2006;45:885-889.

16) Repping-Wuts H, Fransen J, van Achterberg T, Bleijenberg G, van Riel P. Persistent severe fatigue in patients with rheumatoid arthritis. J Clin Nurs 2007;16:377-383.

17) Riemsma RP, Rasker JJ, Taal E, Griep EN, Wouters JMGW, Wiegman O. Fatigue in rheumatoid arthritis: the role of self-efficacy and problematic social support. Br J Rheumatol 1998;37:1042-1046.

18) Scharloo M, Kaptein AA, Weinman JA, Hazes JMW, Breedveld FC, Rooijmans HGM. Predicting functional status in patients with rheumatoid arthritis. J Rheumatol 1999;26:1686-1693.

19) Sokka T, Toloza S, Cutolo M, Kautiainen H, Makinen H, Gogus F, et al. Women, men, and rheumatoid arthritis: Analyses of disease activity, disease characteristics, and treatments in the QUEST-RA Study. Arthritis Res Ther 2009;11:R7.

20) Stebbings S, Herbison P, Doyle TCH, Treharne GJ, Highton J. A comparison of fatigue correlates in rheumatoid arthritis and osteoarthritis: disparity in associations with disability, anxiety and sleep disturbance. Rheumatology (Oxford) 2010;49:361-367.

21) Stone AA, Broderick JE, Porter LS, Kaell AT. The experience of rheumatoid arthritis pain and fatigue: examining momentary reports and correlates over one week. Arthritis Care Res 1997;10:185-193.

22) Thyberg I, Dahlström Ö, Thyberg M. Factors related to fatigue in women and men with early rheumatoid arthritis: The swedish TIRA study. J Rehabil Med 2009;41:904-912.

23) Treharne GJ, Lyons AC, Hale ED, Goodchild CE, Booth DA, Kitas GD. Predictors of fatigue over 1 year among people with rheumatoid arthritis. Psychology Health Med 2008;13:494-504.

24) Wolfe F, Hawley DJ, Wilson K. The prevalence and meaning of fatigue in rheumatic disease. J Rheumatol 1996;23:1407-1417.

25) Breedveld FC, Han C, Bala M, van der Heijde D, Baker D, Kavanaugh AF, Maini RN, Lipsky PE. Association between baseline radiographic damage and improvement in physical function after treatment of patients with rheumatoid arthritis. Ann Rheum Dis 2005;64:52-55.

26) Chorus AMJ, Miedema HS, Boonen A, van der Linden SJ. Quality of life and work in patients with rheumatoid arthritis and ankylosing spondylitis of working age. Ann Rheum Dis 2003;62:1178-1184.

27) Covic T, Tyson G, Spencer D, Howe G. Depression in rheumatoid arthritis patients: demographic, clinical, and psychological predictors. J Psychosom Res 2006;60:469-476.

28) De Croon EM, Sluiter JK, Nijssen TF, Kammeijer M, Dijkmans BAC, Lankhorst GJ, Frings-Dresen MHW. Work ability of Dutch employees with rheumatoid arthritis. Scand J Rheumatol 2005;34:277-283.

29) Davis MC, Affleck G, Zautra AJ, Tennen H. Daily interpersonal events in pain patients: Applying action theory to chronic illness. J Clin Psychol 2006;62:1097-1113.

30) Katz PP. The stresses of rheumatoid arthritis: Appraisals of perceived impact and coping efficacy. Arthritis Care Res 1998;11:9-22.

31) Parker White C, White MB, Fox MA. Maternal fatigue and its relationship to the caregiving environment. Fam Syst Health 2009;27:325-345.

32) Rupp I, Boshuizen HC, Jacobi CE, Dinant HJ, van den Bos GAM. Impact of fatigue on health-related quality of life in rheumatoid arthritis. Arthritis Care Res 2004;51:578-85.

33) Smedstad LM, Moum T, Vaglum P, Kvien TK. The Impact of Early Rheumatoid Arthritis on Psychological Distress: A comparison between 238 patients with RA and 116 matched controls. Scandinavian J Rheumatol 1996;25:377-382.

34) Waltz M. The disease process and utilization of health services in rheumatoid arthritis: The relative contributions of various markers of disease severity in explaining consumption patterns. Arthritis Care Res 2000;13:74-88.

35) Wolfe F, Michaud K, Pincus T. Fatigue, rheumatoid arthritis, and anti-tumor necrosis factor therapy: an investigation in 24,831 patients. J Rheumatol 2004;31:2115-20.

36) Wolfe F, Michaud K. Prevalence, risk, and risk factors for oral and ocular dryness with particular emphasis on rheumatoid arthritis. J Rheumatol 2008;35:1023-1030.

37) Wolfe F, Michaud K. Proposed metrics for the determination of rheumatoid arthritis outcome and treatment success and failure. J Rheumatol 2009;36:27-33.

38) Wolfe F, Michaud K. Predicting depression in rheumatoid arthritis: The signal importance of pain extent and fatigue, and comorbidity. Arthritis Care Res 2009;61:667-673.

39) Yazici Y, Pincus T, Kautiainen H, Sokka T. Morning stiffness in patients with early rheumatoid arthritis is associated more strongly with functional disability than with joint swelling and erythrocyte sedimentation rate. J Rheumatol 2004;31:1723-1726.

40) Belza BL. Comparison of self-reported fatigue in rheumatoid arthritis and controls. J Rheumatol 1995;22:639-643.

41) Crosby LJ. Factors which contribute to fatigue associated with rheumatoid arthritis. J Adv Nurs 1991;16:974-981.

42) Currey SS, Rao JK, Winfield JB, Callahan LF. Performance of a Generic Health-Related Quality of Life Measure in a Clinic Population With Rheumatic Disease. Arthritis Care Res2003;49:658-664.

43) Goodchild CE, Treharne GJ, Booth DA, Kitas GD, Bowman SJ. Measuring fatigue among women with Sjogren's syndrome or rheumatoid arthritis: a comparison of the Profile of Fatigue (ProF) and the Multidimensional Fatigue Inventory (MFI). Musculoskeletal Care 2008;6:31-48.

44) Goodchild CE, Treharne GJ, Booth DA, Bowman SJ. Daytime patterning of fatigue and its associations with the previous night's discomfort and poor sleep among women with primary Sjögren's syndrome or rheumatoid arthritis. Musculoskeletal Care 2010;8:107-117.

45) van Oers ML, Bossema ER, Thoolen BJ, Hartkamp A, Dekkers JC, Godaert GLR, Kruize AA, Derksen RHWM, Bijlsma JWJ, Geenen R. Variability of fatigue during the day in patients with primary Sjögren's syndrome, systemic lupus erythematosus, and rheumatoid arthritis. Clin Exp Rheumatol 2010;28:715-721.

46) Reading SR, Crowson CS, Rodeheffer RJ, Fitz-Gibbon PD, Maradit-Kremers H, Gabriel SE. Do rheumatoid arthritis patients have a higher risk for sleep apnea? J Rheumatol 2009;36:1869-1872.

47) Salaffi F, Sarzi Puttini P, Girolimetti R, Atzeni F, Gasparini S, Grassi W. Helath-related quality of life in fibromyalgia patients: a comparison with rheumatoid arthritis patients and the general population using the SF-36 health survey. Experimental Rheumatology 2009;27:S67-S74.

48) Wolfe F, Michaud K. Severe rheumatoid arthritis (RA), worse outcomes, comorbid illness, and sociodemographic disadvantage characterize ra patients with fibromyalgia. J Rheumatol 2004;31:695-700.

49) Zautra AJ, Fasman R, Parish BP, Davis MC. Daily fatigue in women with osteoarthritis, rheumatoid arthritis, and fibromyalgia. Pain 2007;128:128-135.

APPENDIX Search strategy in the databases

MEDLINE:

"Fatigue"[Mesh] AND ("Models, Theoretical"[Mesh] OR "Causality"[Mesh] OR "etiology"[Subheading]) AND "Musculoskeletal Diseases"[Mesh]

Web of Science:
Topic=((fatigue OR tiredness) SAME (model* OR theor* OR framework OR predict* OR etiology OR pathophysiology OR factor*)) AND Topic=(rheum* OR arthritis OR musculoskeletal OR joint disease)

SCOPUS:
Fatigue w/15 model* AND rheum* OR arthritis OR musculoskeletal OR joint disease, Fatigue w/15 theor* AND rheum* OR arthritis OR musculoskeletal OR joint disease, Fatigue w/15 framework AND rheum* OR arthritis OR musculoskeletal OR joint disease, Fatigue w/15 predict* AND rheum* OR arthritis OR musculoskeletal OR joint disease, Fatigue w/15 etiology AND rheum* OR arthritis OR musculoskeletal OR joint disease, Fatigue w/15 pathophysiology AND rheum* OR arthritis OR musculoskeletal OR joint disease, Fatigue w/15 factor* AND rheum* OR arthritis OR musculoskeletal OR joint disease
The same search was done with the word “tiredness” instead of “fatigue”.

PsychInfo:
((fatigue N15 model*) OR (tiredness N15 model*) OR (fatigue N15 theor*) OR (tiredness N15 theor*) OR (fatigue N15 framework) OR (tiredness N15 framework) OR (fatigue N15 predict*) OR (tiredness N15 predict*) OR (fatigue N15 etiology) OR (tiredness N15 etiology) OR (fatigue N15 pathophysiology) OR (tiredness N15 pathophysiology) OR (fatigue N15 factor*) OR (tiredness N15 factor*)) and (rheum* OR arthritis OR musculoskeletal OR joint disease)

Chapter 3

Measuring fatigue in rheumatoid arthritis

S. Nikolaus
M.A.F.J. van de Laar

Nat Rev Rheumatol 2011 Aug 30;7(10):562-4. Advance online publication 30 August 2011;
doi:10.1038/ nrrheum.2011.124

Irksome to debilitating, fatigue is a frequent symptom of rheumatoid arthritis (RA). Patients regard it as an important, neglected facet of the disease. An analysis of the effect of biologic therapies on fatigue has yielded disappointing results—do current treatments lack efficacy, or are we failing to measure adequate parameters?

When TNF-blockers were introduced into clinical practice, many patients anecdotally reported a meaningful reduction in fatigue. As fatigue is experienced as one of the most important symptoms of RA, such unsystematic observations were noted with interest. Today, fatigue is regarded as an important patient-reported outcome measure in rheumatology,[1] and is increasingly measured in clinical trials. This trend is reflected in the publication, by Chauffier et al.[2] in *Rheumatology (Oxford),* of a meta-analysis of the effects of biologic agents on fatigue in patients with RA. Forming one of the first concrete answers to the expectations and hopes that rest on biologic therapies to tackle this symptom, the results are disappointing.[2]

Patients with RA frequently report feeling unable to control fatigue, and that the symptom is mostly neglected by clinicians when evaluating the severity and treatment of the disease. Moreover, patients consider fatigue to be an annoying symptom with far-reaching consequences.[3] The experience of fatigue is multidimensional: it manifests not only as a physical sensation, but also in emotional and cognitive ways. Accordingly, fatigue has a serious impact on many different areas of daily life; for example, it is an important predictor of work productivity.[4]

Thus, the analysis by Chauffier and colleagues,[2] which included 10 randomized controlled trials, is an anticlimax to the hopes invested in biologic therapies, reporting that the impact of such treatments on fatigue is small, and comparable to their effects on physical functioning.

Two main scenarios might explain the disappointing effect sizes calculated. Firstly, biotherapies might be inefficacious for treating fatigue in RA. If so, the cause could be either that they do not have an effect on the mechanisms that underlie fatigue, or that their effects are limited to only some of the relevant pathways, which are eclipsed by those mechanisms that remain unaffected. Alternatively, it could be that biotherapies do have effects on fatigue in RA, but that we are unable to measure them.

The first of these hypotheses is not yet provable in its simplest form, as we lack sufficient knowledge of the causes and pathways of fatigue and RA (and of fatigue in RA). As

patients with RA mention fatigue as a specific feature of their disease, it is plausible that RA and fatigue have some mechanisms in common. Nevertheless, fatigue in RA might be driven by processes other than the immunological pathways that drive other symptoms of the disease. As a consequence, medication directed at the immune system alone would not lead to measurable reduction in fatigue. The relationship between fatigue and inflammation in RA is not yet clear, as underlined by a recent review[5] that discussed studies with conflicting results: some investigators have demonstrated a correlation between fatigue and inflammatory markers in RA, whereas others have found no such associations, reflecting the complexity of the links between disease activity and fatigue in RA. Plausibly, therefore, we can assume multi-causality of fatigue; physical components might be involved, but psychological processes might also play a part in the etiology.[6]

The physical component of fatigue is likely to be subject to modification by biologic agents, via a reduction in inflammatory processes. Improvements in pain and sleep would lead to a reduction in fatigue. However, the psychological component will be more difficult to influence, because reactions to the experience of a disease symptom—such as avoidance behaviors, feelings of helplessness and/or depressive mood—can become habitual. Such unconscious behaviors and attributions can be, or can become, independent of current levels of inflammation and pain, and no longer controllable by medication. Thus, medication alone is probably not enough to control fatigue; additional patient education, behavioral change and cognitive restructuring components are needed; for example, in the form of group cognitive–behavioral therapy.[7]

Another explanation for the disappointing conclusion of the meta-anlaysis[2] concerns measurement. When we are unable to measure fatigue appropriately we cannot evaluate it adequately—the measurement of fatigue is an issue in itself. Several multi-item questionnaires are used in clinical practice and research, but they all have shortcomings. Indeed, a systematic review of measurements of fatigue in RA showed that reasonable evidence of their validity in the disease exists for only a few of them.[8] Furthermore, none met all criteria for validity.[8] The main problem identified was content validity: the perspective of patients with RA was not included during development of the measures.
All studies that were included in the analysis by Chauffier *et al.*[2] used the Short Form 36 (SF‑36) vitality subscale and/or the Functional Assessment of Chronic Illness Therapy Fatigue Scale (FACIT‑F). Whether these two scales are the most appropriate for the measurement of fatigue in RA is disputable. Both measures have shown some validity,[8] but might be open to distortion when used in populations of patients with RA. The vitality subscale of the SF‑36, a generic scale, might not be able to differentiate well between

depression and fatigue.[8] Moreover, one must question whether vitality and fatigue reflect a one-dimensional construct. The FACIT-F was originally developed to measure fatigue in patients with cancer, and some items might reflect physically disabling symptoms of RA, rather than fatigue. To ensure content validity of a measurement instrument, canvassing the opinion of patients and professionals in the field is essential during development.[9] Only patients can report on the experience of fatigue.

The Bristol RA Fatigue Multi-Dimensional Questionnaire (BRAF-MDQ) is a recently devised fatigue questionnaire that was not used in any of the studies analyzed by Chauffier and colleagues.[2] Its approach is promising, as it embraces the patient's perspective by developing items based on interviews, focus groups and cognitive testing. Nevertheless, the BRAF-MDQ is a traditional instrument, meaning that a fixed number of items have to be filled in by each patient.

By contrast, computer-adaptive testing (CAT) is a new technology that allows comprehensive measurement of patient-reported outcomes (PROs) with relatively few items.[10] Items are respectively selected from an item bank, based on a patient's previous answer, so that precise measurement at individual level with few items becomes possible. By using Item Response Theory, items can ideally be matched to the patient's individual level, and inter-individual comparisons of the measured construct can be made, even if patients filled in different items. Primarily, CAT was developed and has been used for ability and achievement testing. To date, few CAT applications have been validated and implemented in medical settings, but interest in CAT for health-related measures is growing.[10] CATs constructed recently to measure depression and anxiety seem to be reliable, valid and efficient instruments that are more precise than traditional questionnaires.[10] Moreover, as they are shorter, the burden for patients is reduced. None of the studies included in the *Rheumatology (Oxford)* paper[2] used this type of assessment; a CAT for fatigue in RA is not yet available, but will be developed in the near future. Potentially, this promising technology might contribute to a better measurement of fatigue in RA and mean that the effects of treatment can be demonstrated more clearly in the future.

The meta-analysis of the effect of biotherapies on fatigue in patients with RA by Chauffier *et al.*[2] clearly shows that measuring and treating fatigue in RA are major problems. Either our current treatments are insufficient to influence one of the major outcomes of the RA disease process, or we are unable to measure RA-associated fatigue in itself.

ACKNOWLEDGEMENT

The authors wish to thank Dr Christina Bode of the Arthritis Center Twente at the University of Twente, the Netherlands, for substantial contributions to discussions during the preparation of this article.

REFERENCES

1. Kirwan JR, Minnock P, Adebajo A, Bresnihan B, Choy, de Wit M, et al. Patient perspective: fatigue as a recommended patient centered outcome measure in rheumatoid arthritis. J Rheumatol 2007;34:1174–1177.
2. Chauffier K, Salliot C, Berenbaum F, Sellam J. Effect of biotherapies on fatigue in rheumatoid arthritis: a systematic review of the literature and meta-analysis. Rheumatology (Oxford) 2012;51:60-8. doi:10.1093/rheumatology/ker162.
3. Nikolaus S, Bode C, Taal E, van de Laar MAFJ. New insights into the experience of fatigue among patients with rheumatoid arthritis: a qualitative study. Ann Rheum Dis 2010;69:895–897.
4. De Croon EM, Sluiter JK, Nijssen TF, Kammeijer M, Dijkmans BA, Lankhorst GJ, et al. Work ability of Dutch employees with rheumatoid arthritis. Scand J Rheumatol 2005;34:277–283.
5. Stebbings S, Treharne GJ. Fatigue in rheumatic diseases: an overview. Int J Clin Rheumatol 2010;5:487–502.
6. Norheim KB, Jonsson G, Omdal R. Biological mechanisms of chronic fatigue. Rheumatology (Oxford) 2011;50:1009–1018.
7. Hewlett S, Ambler N, Almeida C, Cliss A, Hammond A, Kitchen K, et al. Self-management of fatigue in rheumatoid arthritis: a randomised controlled trial of group cognitive-behavioural therapy. Ann Rheum Dis 2011;70:1060–1067.
8. Hewlett S, Hehir M, Kirwan JR. Measuring fatigue in rheumatoid arthritis: a systematic review of scales in use. Arthritis Rheum 2007;57:429–439.
9. Streiner DL, Norman GR. Devising the items. In: Streiner DL, Norman GR eds. Health measurement scales—a practical guide to their development and use. New York: Oxford University Press, 2003:14-28.
10. Walter OB. Adaptive tests for measuring anxiety and depression. In: WJ van der Linden, CAW Glas eds. Elements of Adaptive Testing. New York, Springer, 2010:123-126.

Chapter 4

New insights into the experience of fatigue among patients with rheumatoid arthritis: a qualitative study

S. Nikolaus
C. Bode
E. Taal
M.A.F.J. van de Laar

Ann Rheum Dis 2010;69:895–897. doi:10.1136/ard.2009.118067 895

ABSTRACT

Objectives Patients with rheumatoid arthritis (RA) commonly experience fatigue. The aim of this study was to gain further insight into the experience of fatigue in RA.

Methods Participants were 31 outpatients with RA of Medical Spectrum Twente, Enschede, The Netherlands, with all levels of fatigue. In-depth structured interviews on the patients' experience of fatigue were conducted and analysed using a bottom-up coding scheme, meaning that answers of patients were summarised and categorised.

Results Patients' mean fatigue severity score was 50 (visual analogue scale (VAS); theoretical range 0–100). Interviews showed interindividual differences in the experience and impact of fatigue. Different patterns in emotions, consequences and management of fatigue were found. Especially younger women with multiple daily roles seemed to be vulnerable to the negative impact of RA fatigue. Patients also reported positive aspects of fatigue. Moreover, varying forms of fatigue were described.

Conclusions Results point to the existence of differences in fatigue experience according to gender, age and daily roles. This finding indicates a need for targeting advice and interventions to the individual situation of a patient. Furthermore, the positive aspects patients reported about their fatigue could facilitate an approach that is focused on remaining opportunities and not just on restrictions.

INTRODUCTION

Patients with rheumatoid arthritis (RA) mention fatigue as one of their most bothersome symptoms.[1] An internationally accepted definition of fatigue in RA does not exist,[2] and little is known about its aetiology and the possibilities for support by health professionals.[3,4] Three studies on the experience of fatigue in RA[5–7] showed that fatigue is a multidimensional, bothersome symptom with far-reaching consequences. However, these results were found in small or highly selective samples[5,7] or with predetermined coding schemes.[6] These studies give a first explorative insight into the experience of RA fatigue, but did not address differences between patients in their descriptions of fatigue. As the subjective experience of a symptom is always influenced by an interaction of several individual patient characteristics,[8] we were interested in intraindividual differences in the experience of RA fatigue. Patients with osteoarthritis have reported varying forms of fatigue, making a distinction between physical and mental fatigue.[9] Hence we were interested whether fatigue experience differs between patients with RA and whether one patient can have different experiences of fatigue.

PATIENTS AND METHODS

In February and March 2008, consecutive outpatients visiting the rheumatology clinic of Medical Spectrum Twente, Enschede, The Netherlands, were invited to participate. Selection criterion was a diagnosis of RA minimally 2 years ago, so that patients were able to talk about a certain amount of experience with RA fatigue. Interested patients received an information letter about the study and made an appointment with the interviewer (SN). Patients completed a questionnaire via a computer with touch screen at the hospital. More information about the methods can be found in appendix 1.

Interviews

The interviewer (SN) used a hierarchical interview scheme comprising eight main topics (table 1) with several subquestions. The interview scheme was based on the interview schedules used in the previous interview studies[5,7] and a literature search on fatigue models.

Table 1 Interview topics

Topic	Example question
(1) Personal situation and disease-related information	- What are your daily activities?
(2) Manifestation of fatigue	- How often are you fatigued?
(3) Experience of fatigue	- How would you describe your fatigue?
(4) Determinants of fatigue	- According to your opinion, what causes your fatigue?
(5) Consequences of fatigue	- What are the consequences of fatigue for your daily life?
(6) Coping with fatigue	- How do you deal with your fatigue?
(7) Support by others (personal and	- What kind of things do other people do to professional) lessen your fatigue?
(8) Additional comments	-Is there anything I did not ask you about your fatigue that you would like to add?

Analyses

We used an approach that combines deductive and inductive elements of analysis. This approach is similar to the framework approach,[10] but the deductive element is slightly more dominant. All interviews were recorded on audiotape and transcribed verbatim. After careful reading of the interviews by three researchers (SN, CB, ET), a coding scheme was developed bottom-up, meaning that the answers of patients were summarised and categorised, without using predetermined categories. With some new codes emerging during the coding process, previously coded interviews were checked for them. Three researchers (SN, CB, ET) coded 10% of the interviews to check the inter-rater reliability. Kappa values were calculated for each interview question. The mean value of 0.63 was sufficient.[11] Statements with different codes were discussed until full consensus was reached.

RESULTS

Sample

The sample consisted of 31 outpatients; 8 men and 23 women with a mean age of 58 years (SD=13.78), ranging from 32 to 83. Further sample characteristics are shown in appendix 2. In the following paragraph, we present our qualitative results with focus on meaning and description of fatigue. Appendix 3 also contains further information about striking differences between patients.

Interindividual differences

Negative emotions related to fatigue

Most negative emotions were reported by women patients. Only two men mentioned unpleasant emotions at all: slightly worsened mood and feeling useless. Women reported "being oversensitive, becoming irritated more quickly, not being able to deal with things or being catty", being angry, being in a bad or depressive mood, feeling miserable, rebellious, unpleasant, sad, down, despondent, frustrated, aggressive, anxious, self-pitying or misunderstood. "Well, if you're really tired, you're just angry with yourself. (…) That you're tired again. You don't want to be, but you are. (…)" (respondent (R) 13, woman, 39 years old, mother, spouse, housewife, employee). It was remarkable that especially younger women with multiple daily roles experienced negative emotions related to fatigue.

Consequences of fatigue

Most patients reported fatigue having consequences for daily life; they felt impeded in their mobility and activities. Only a few older patients (>67 years) reported no consequences at all. Many patients felt restricted by fatigue in housekeeping and gardening. "Yes, that you can't get on with things like you used to. That you just have to rest more (…) that you have to become more easygoing about your work. That everything here doesn't sparkle or shine like a new pin, and yes, you just have to learn to accept it" (R3, woman, 52, mother, spouse, housewife, employee). No man, but several women reported consequences for social contacts, for example, having to cancel appointments because of fatigue. For some women helping and supporting others became more difficult, and some women experienced stress from fatigue. Several women reported work-related consequences, for example, being unable to work (in a certain job), or being tired during work.

Coping with fatigue

Only women reported that their ability to successfully cope with fatigue was variable. "Yes, it can vary; it depends on how I feel. Sometimes I accept it, but other times I can't; I can get angry, rebellious" (R32, woman, 55, spouse, housewife, employee). No man reported having to dose activities and find a balance, but several women did. "Yes, I used to work full time, on top of the housekeeping and the children, you just did it. And now you have to plan what you're going to do, which means there are some things you can't. Ah, your entire day revolves around it and spent thinking how you're going to get through the day" (R1, woman, 44, mother, spouse, housewife, employee).

Different forms of fatigue

Some patients reported varying fatigue experiences. They distinguished between mental and physical fatigue, fatigue with or without a prior reason (such as poor sleep or physical activity), fatigue in combination with or without pain, with or without dizziness and with or without the desire to go to bed and sleep. “Around quarter to 11 I start to yawn, and want to go to bed. Right, but that’s not the kind of tiredness I feel at 3 in the afternoon. (…) I don’t want to go to bed. Then it’s just sitting and being so tired you can’t even put your cup down on the table” (R7, woman, 65). Furthermore, patients described differences between forms of fatigue based on severity, frequency or duration.

Most patients had difficulties to specify at least one of the following points: how severe the fatigue is, how often the fatigue occurs and how long it lasts. Patients thought the reason was that fatigue is not always the same. “I have weeks in which I’m really tired for 3 or 4 days. And I have weeks that are fine. (…) But it depends of course on what you’re doing, what the circumstances are; even the weather can have an effect” (R6, woman, 59).

Benefit finding

Some patients saw positive aspects of fatigue, as making more conscious decisions in life “You just go about things a bit differently and you make more conscious decisions about what is really important. (…) That’s the positive side of it. Normally I would probably have just gone on racing about doing all sorts of things, but now you just think about it more” (R14, woman, 42), learning to let things go “I’m sure you also learn from it that not everything is so important that you have to do everything you want to. That you can also just give in and leave things as they are. (…) Of course I’m lucky that I don’t have a busy job, so it’s easy for me to give in. (…)” (R6, woman, 59), or learning to appreciate the beneficial aspects of rest “Undoubtedly, because you come to rest. (...) rest is a natural phenomenon I think, which is good for humans. So it [the fatigue] will undoubtedly have positive aspects” (R5, man, 68). Patients also reported that fatigue can make daily life easier; being able to fall asleep easier than before, that the boss can see when one is really tired and being able to use fatigue as an excuse.

DISCUSSION

This is the first study showing interindividual and intraindividual differences in the patients’ experience of RA fatigue, and that fatigue can also have benefits for patients as well. Compared to men, women reported gender-specific coping patterns, more negative emotions and more consequences of fatigue. This applied especially to younger women

with multiple daily roles (eg, spouse, employee, housekeeping, childcare). In contrast, only older patients reported no consequences of fatigue at all. In this study, fewer men than women participated, which may reduce the generalisability of the results. However, the spreading of the sample almost corresponds to the epidemiological distribution in RA.[12] We obtained information about daily roles from demographic data and asking patients about their daily activities. For future fatigue research we recommend to comprehensively measure daily roles, because it seems likely that not gender and age as such are associated with the experience of fatigue, but the amount of daily roles. Compared to men, women more often fulfil multiple roles in daily life.[13,14] After retirement, people have fewer roles,[15] and expectations from others are less demanding. Therefore women might feel more restricted by fatigue than men, and older people might be able to cope with fatigue more easily. Women reported managing fatigue by dosing and planning their activities to find a balance between activity and rest. Negative emotions related to fatigue might occur when personal possibilities do not meet with one's expectations and wishes, and might be worsened by external demands related to multiple roles. Consequently, to reduce or prevent negative emotions, adjustment of personal goals and standards as well as negotiating external demands might be a good strategy.[16,17] Moreover, developing or strengthening an awareness of positive fatigue aspects, could be valuable in reducing negative emotions. Previous interview studies did not report beneficial aspects of fatigue, probably because the patients were not asked about it. However, their existence is of special importance since the ability to find something beneficial in a negative experience can lead to improvements in adaptation and health.[18] In daily clinical practice professionals should have a supportive attitude to allow patients to discuss their fatigue, because its description can be difficult due to various forms and variations in severity, frequency and duration. A theoretical, comprehensive model explaining the experience of RA fatigue is still lacking.[19] In future research, the dimensions of RA fatigue and possibly related patient characteristics should be examined in more detail. This study was qualitative, and did not test hypotheses about differences between patients. However, our data show new indications for the existence of such differences in the experience and description of RA fatigue. These new results could inform future quantitative research. The results of this study point to the relevance for an individualised professional approach and a precise measurement of fatigue in patients with rheumatoid arthritis. The positive aspects patients reported about their fatigue could facilitate an approach that is focused on remaining opportunities and not just on restrictions.

ACKNOWLEDGEMENT

We would like to thank the patients with RA who participated in this study.

REFERENCES

1. Hewlett S, Carr M, Ryan S, et al. Outcomes generated by patients with rheumatoid arthritis: how important are they? Musculoskeletal Care 2005;3:131–42.
2. Hewlett S, Hehir M, Kirwan JR. Measuring fatigue in rheumatoid arthritis: a systematic review of scales in use. Arthritis Rheum 2007;57:429–39.
3. Swain MG. Fatigue in chronic disease. Clin Sci 2000;99:1–8.
4. Repping-Wuts H, van Riel P, van Achterberg T. Fatigue in patients with rheumatoid arthritis: what is known and what is needed. Rheumatology 2009;48:207–9.
5. Hewlett S, Cockshott Z, Byron M, et al. Patients' perceptions of fatigue in rheumatoid arthritis: overwhelming, uncontrollable, ignored. Arthritis Care Res 2005;53:697–702.
6. Repping-Wuts H, Uitterhoeve R, van Riel P, et al. Fatigue as experienced by patients with rheumatoid arthritis (RA): a qualitative study. Int J Nurs Stud 2008;45:995–1002.
7 Belza Tack B. Fatigue in rheumatoid arthritis. Conditions, strategies, and consequences. Arthritis Care Res 1990;3:65–70.
8. Gijsbers van Wijk CM, Kolk AM. Sex differences in physical symptoms: the contribution of symptom perception theory. Soc Sci Med 1997;45:231–46.
9. Power JD, Badley EM, French MR, et al. Fatigue in osteoarthritis: a qualitative study. BMC Musculoskelet Disord 2008;9:63.
10. Pope C, Ziebland S, Mays N. Qualitative research in health care. Analysing qualitative data. BMJ 2000;320:114–16.
11. Altman DG, ed. Some common problems in medical research. In: Practical Statistics for Medical Research. London: Chapman & Hall/CRC, 1991:396–439.
12. Alamanos Y, Drosos AA. Epidemiology of adult rheumatoid arthritis. Autoimmun Rev 2005;4:130–6.
13. Cloin M, Hermans B. Onbetaalde arbeid en de combinatie van arbeid en zorg. [Unpaid work and the combination of work and care]. In: Portegijs W, Hermans B, Lalta V, eds. Emancipatiemonitor 2006 – Veranderingen in de leefsituatie en levensloop [Emancipationmonitor 2006 – Changes in living situation and life course]. Den Haag: Sociaal en Cultureel Planbureau, Centraal Bureau voor de Statistiek, 2006:100–140.
14. Eurostat. The Life of Women and Men in Europe – A Statistical Portrait. Luxembourg: Offi ce for Offi cial Publications of the European Communities, 2008.
15. Riley M, Kahn RL, Foner N, eds. Age and Structural Lag. New York: Wiley, 1994.
16. Bode C, de Ridder DTD, Kuijer RG, et al. Effects of an intervention promoting proactive coping competencies in middle and late adulthood. Gerontologist 2007;47:42–51.

17. Bode C, de Ridder DTD, Bensing JM. Preparing for aging: Development, feasibility and preliminary results of an educational program for midlife and older based on proactive coping theory. Patient Educ Couns 2006;61:272–278.
18. Danoff-Burg S, Revenson TA. Benefi t-fi nding among patients with rheumatoid arthritis: positive effects on interpersonal relationships. J Behav Med 2005;28:91–103.
19. Hewlett S, Nicklin J, Treharne GJ. Fatigue in musculoskeletal conditions. Topical Reviews Series 6: No 1. Arthritis Campaign: Autumn 2008. www.arc.org.uk/arthinfo/ medpubs/6641/6641.asp (accessed 30 September 2009).

APPENDIX 1

The questionnaire contained questions about age, gender, marital status, education, work status and disease duration. Furthermore, the participants completed visual analogue scales (VAS) for pain and general health, the Health Assessment Questionnaire (HAQ), and three VAS fatigue scales[1] for severity ("Please place a mark on the line to show your level of fatigue over the past 7 days" (anchors: No fatigue – Totally exhausted)), impact ("Please place a mark on the line to show the effect that fatigue has had on your life over the past 7 days" (anchors: No effect – A great deal of effect)), and coping ("Please place a mark on the line to show how well you have coped with fatigue over the past 7 days" (anchors: Not at all well - Very well)). As these three VAS have been developed in the UK, they were translated into Dutch, according to the recommended translation procedure for cross-cultural adaptation.[2] We collected further information during the interviews, such as medication use, co-morbidities and the patients' daily activities.

According to local regulations in the Netherlands (WMO) the study did not need approval of the ethical review board. We used patient informed consent to make sure that patients decided consciously to participate in this study. Most patients preferred an interview at home, two patients chose to come to the university.

1. Nicklin JK, Kirwan JR, Cramp F, Hewlett S. Standardizing visual analogue scales to measure fatigue in rheumatoid arthritis. Arthritis Rheum 2007;56,S778.

2. Guillemin F, Bombardier C, Beaton D. Cross-cultural adaptation of health-related quality of life measures: literature review and proposed guidelines. J Clin Epidemiol 1993;46:1417-1432.

APPENDIX 2

Table A1 Description of the sample RA patients (N=31)

	N
Sex	
Women	23
Men	8
Age *	
32-57 years (younger group)	15
58-83 years (older group)	16
Marital status	
Single	3
Living with partner/married	20
Widow/widower	8
Level of education	
Low (≤ 12 years of education)	15
Moderate (13-14 years of education)	10
High (≥ 14 years of education)	6
Work status	
Working full-time	4
Working part-time	9
Housewife	4
Disabled/Retired	14
Roles **	
Few roles	17
Multiple roles	14
Co-morbidities	
Other rheumatic disease/s	4
Other disease/s	13
Medication	
DMARDs (Disease-modifying antirheumatic drugs)	26
NSAIDs (Non-steroidal anti-inflammatory drugs)	10
Prednisone	5
Biologicals	9
Medication for other diseases	13

* Two age groups were formed by splitting the group at the median age.

** The variable "roles" was constructed based on demographic data from the interviews and answers to a question about the patients' daily activities. Patients were allocated to the group "multiple roles" when they fulfilled minimal three of the following criteria: responsibility for children living at home, being a spouse/partner, working at least part-time, being responsible for household/garden. Patients fulfilling fewer criteria were allocated to the group "few roles".

Table A2 Health-related characteristics of the sample RA patients

	Mean (S.D.)	Range
VAS Pain	4.00 (2.31)	1.00 – 9.00
VAS General health	4.00 (2.28)	1.00 – 9.00
VAS Fatigue severity	50.00 (18.15)	14.00 – 90.00
VAS Impact of fatigue	49.00 (23.65)	1.00 – 90.00
VAS Coping with fatigue	67.00 (21.72)	16.00 – 99.00
HAQ-DI	0.83 (0.63)	0.00 - 2.88
Disease duration	11.00 years (7.95)	2.00 – 35.00 years

VAS= Visual Analogue Scale, HAQ-DI=Health Assessment Questionnaire Disability Index

The mean fatigue severity was 50, the mean impact of fatigue was 49 and the mean coping with fatigue score was 67. VAS fatigue scores did not differ significantly between men and women or between younger and older patients ($p>0.05$).

APPENDIX 3

In Table A3, the inter-individual differences in fatigue experiences between men and women, two age groups, and people with few or multiple daily roles are summarized.

Table A3 Gender, age and role patterns in the experience of fatigue

Topic	Gender	Age	Roles
Negative emotions			
no negative emotions	> ♂	> older	> few roles
one or more emotions	> ♀	> younger	> multiple roles
Consequences			
no consequences	> ♂	only older	only few roles
mobility and activity	> ♀	-	-
housekeeping/gardening	> ♀	> younger	> multiple roles
social contacts	only ♀	> younger	-
helping others	only ♀	-	-
work	only ♀	only younger	> multiple roles
stress	only ♀	> younger	> multiple roles
Coping/Management			
varying ability to cope	only ♀	> younger	> multiple roles
dosing activities	only ♀	> younger	> multiple roles

♂ = men, ♀ = women, > = more, - = no difference; a difference is defined as at least 15% difference between the percentages of respondents in every group (men vs. women/older vs. younger/few vs. multiple).

Chapter 5

Four different patterns of fatigue in rheumatoid arthritis patients: results of a Q-sort study

S. Nikolaus
C. Bode
E. Taal
M.A.F.J. van de Laar

Rheumatology 2010;49:2191-2199. doi:10.1093/rheumatology/keq210

ABSTRACT

Objectives. Many patients with RA complain about fatigue. Whereas qualitative studies have covered the meaning of fatigue for RA patients, it was still unknown whether subgroups of patients could be distinguished. This study aimed to describe different perspectives on the experience of fatigue.

Methods. Participants were 30 outpatients with established controlled RA of the Medical Spectrum Twente, with a mean fatigue severity score of 4.67 (visual analogue scale 0–10). They evaluated 57 statements about fatigue according to Q-methodology. Data were analysed with PCQ for Windows (Portland, OR, USA), using centroid factor analysis with varimax rotation (i.e. the participants but not the items of a scale are the variables). Factor scores of statements on the different dimensions were calculated to investigate which items are relevant when describing and distinguishing fatigue experiences. Demographic and clinical patient characteristics were collected to describe each of the dimensions that resulted from centroid factor analysis.

Results. After statistical and theoretical considerations, a four-factor structure of the data was obtained. Each factor represented a perspective on the experience of fatigue, shared by a certain group of patients. Physical, psychological and social patient characteristics seemed to be associated with those experiences. The factors were labelled as: 'Little impact of fatigue'; 'Good coping and bad sleep'; 'Search for balance'; and 'High distress'.

Conclusions. Results indicated that fatigue experience is a complex phenomenon. Existing questionnaires to measure fatigue do not meet this complexity. Extensive research is warranted and new efforts to develop instruments that take into account all aspects of fatigue are indicated.

INTRODUCTION

Fatigue is a common problem for patients with RA.[1,2] Patients mention fatigue as one of their most annoying problems.[3,4] Primarily, fatigue is a subjective experience that can be described as 'extreme and persistent tiredness, weakness or exhaustion—mental, physical or both'.[5] However, an internationally accepted definition of fatigue in RA does not exist,[6] and little is known about its aetiology.[7] Nevertheless, the number of studies including fatigue as outcome measure has rapidly increased over the last 2 years. This is problematic as so little is still known about fatigue. A theoretical model explaining the experience of fatigue in RA is still lacking[8] and consensus about an adequate way to measure it has hence not yet been reached. First, more insight into the complexity of the phenomenon fatigue should be obtained.

In analogy with pain we should assume that fatigue is what a patient says it is, and it is as bad as the patient says it is. Consequently, the most appropriate starting point is the perspective of the patients.[4] To date, four qualitative studies examining patients' experience of fatigue in RA have been conducted in the USA, the UK and The Netherlands.[9–12] They all showed that RA patients experience fatigue as a multidimensional, annoying symptom with far-reaching consequences. Our previous study[12] revealed inter- and intra-individual differences in the patients' experience of fatigue in RA: emotions, consequences and management of fatigue differed with age and gender. Some of these differences were related to the number of daily roles that patients had to fulfil. Many patients reported that the severity, frequency and duration of their fatigue varied, with some even reporting they experienced different forms of fatigue. It is not yet fully understood which factors cause such inter- and intraindividual differences in fatigue in RA patients. Although we know about the experience of fatigue in general, it is still not known whether there are certain groups of RA patients who experience fatigue in a similar way, and which patient characteristics are related to differences in fatigue experiences between such groups.

Studies on predicting factors of fatigue in RA show that not only physical but also psychological and social aspects can be related to fatigue. Several studies found that pain,[1,2,13–15] physical disabilities,[1,2,16,17] impaired general health,[16] longer symptom duration,[13] less physical activities, comorbid conditions,[2] lower sleep quality[1,2] and female gender[2,13] were predictive of fatigue in RA. Moreover, psychosocial factors emerged, which are: depressive symptoms;[1,13,14] anxiety;[17] social stress and problematic social support;[13] lower self-efficacy;[15,18] a lower perceived sense of control;[19] and in female RA patients, negative interpersonal events predicted higher levels of fatigue.[20] Regarding the

relationship between disease activity and fatigue, the literature shows inconsistent results.[1,7,13–16,18,21,22]

The purpose of this study was to identify different perspectives on the experience of fatigue in RA. Furthermore, in our aim to investigate whether several patient characteristics are associated with the fatigue experiences, we collected a broad range of characteristics that covered not only disease-specific information but also topics relating to the patient's daily life. So besides disease-specific variables (e.g. disease activity, disability, pain), we included social (e.g. age, marital status, daily roles) and psychological variables (e.g. anxiety, depression).

PATIENTS AND METHOD

Q-methodology

We used Q-methodology, which is a valuable tool in research on highly subjective topics, such as fatigue. Especially in situations when a topic is not yet well understood, this method is recommended.[23] In comparison with the Likert attitude questionnaires, Q-methodology provides highly detailed information and identifies groups of people sharing similar views on a certain topic.[24] It has already been used to examine the acceptance of chronic pain[25] and to identify characteristics of adolescent cancer patients sharing a common experience of fatigue.[26]

In Q-methodology, a by-person factor analysis is applied, whereby the participants but not the items of a scale are the variables.[27] First, the researchers had to prepare 40–60 statements that cover all aspects of the topic under study.[28] We developed a comprehensive pool of statements by collecting items from multi-item fatigue scales and by selecting relevant statements from our interview material.[12] For this, we used the items of the Multidimensional Assessment of Fatigue (MAF) scale,[29,30] Functional Assessment of Chronic Illness Therapy Fatigue (FACIT-F) scale,[31,32] Profile of Mood States (POMS)[33] and Short Form 36 (SF-36) vitality subscale.[34] In a systematic review on measurement of fatigue in RA,[6] evidence of reasonable validation could be identified for these scales, although even they did not meet all the criteria for validation. Moreover, we used the items of the Checklist Individual Strength (CIS)[35,36] and the Bristol RA Fatigue Multi-Dimensional Questionnaire (BRAF-MDQ),[37] two scales which are undergoing validation in RA at the moment. This resulted in 200 items, which were systematically reduced. In cases of overlapping content of items, we chose the formulation from the interviews or the clearest formulation (when two questionnaire items overlapped). We discussed the

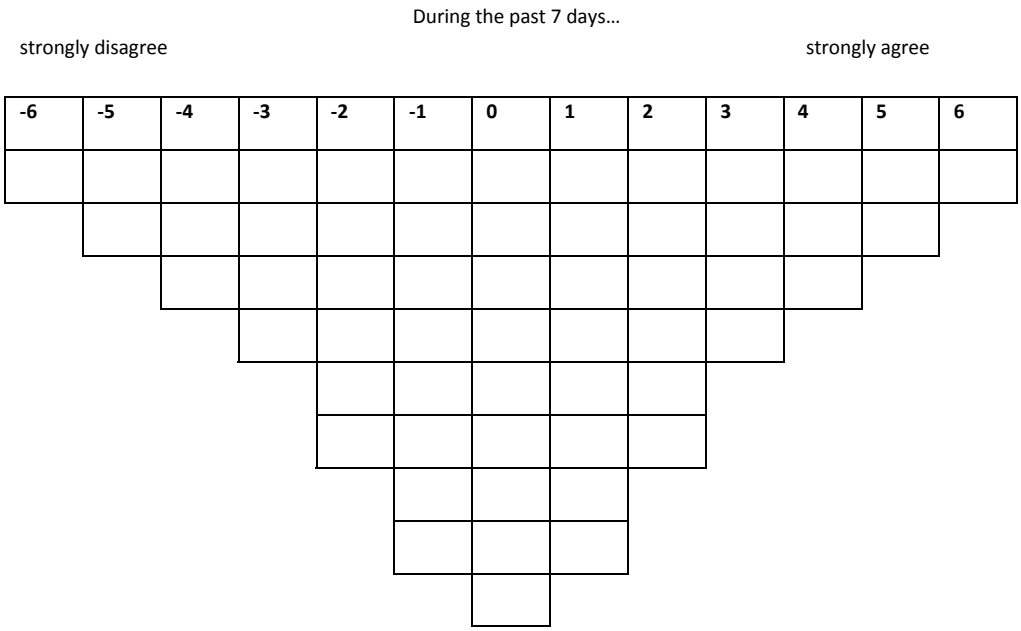
remaining items, whereby those were identified that represented the most relevant categories, based on the interview results and relevant scientific literature. Ultimately, 57 items were selected that covered 14 categories: negative emotions, consequences, impact/ meaning, benefit finding, different forms of fatigue, severity, duration, frequency, cognition/ concentration, energy, sleep, description, social support and coping/ management.

The items were then printed on cards, which had to be sorted by each participant on a worksheet with a Q-sort figure, to reflect his/her agreement with the statements[28] during the previous 7 days. We used a Q-sort figure (Fig. 1) with 11 piles, ranging from strongly disagree (-6) to strongly agree (+6). The piles consisted of different numbers of squares, whereby most of the cards could be sorted in the neutral regions and the fewest in the extremes. To check the feasibility of the selected items, we piloted the Q-sort with two patients. Both the patients judged the method as clear and the items as comprehensive and relevant.

Patients

This study was performed in collaboration with the patient research partners of our Arthritis Centre Twente. In March 2009, 30 RA outpatients of the rheumatology clinic of Medical Spectrum Twente participated in the study. When using Q-methodology, the adequate number of participants depends on the number of statements that have to be sorted.[38] Ideally, the number of statements is at least twice the number of participants.[39] Since we used 57 statements in this study, 30 participants is the absolute maximum number to be able to conduct a methodologically correct analysis. Our aim was to include a diverse sample of patients, giving an adequate reflection of the usual RA outpatient population with all levels of fatigue. Selection criterion was a diagnosis of RA at least 2 years previously to make sure that the patients had sufficient experience with RA to be able to sort the statements on fatigue. According to local regulations in The Netherlands (WMO), the study did not need approval of the ethical review board; only (non-intervention) studies with a high burden for patients have to be reviewed. Patients underwent no intervention or treatment and the burden of participation was relatively low. Before starting the Q-sort, patients signed an informed consent, to make sure that they decided consciously to participate in this study, and completed a questionnaire.

Figure 1 Q-sort worksheet



Method

For conducting the Q-sort, the researcher (S.N.) made individual appointments with the patients. Most of them chose to complete the Q-sort at home, and two at the university. They were asked to read the statements about fatigue and to make three initial global piles (agreed with; did not agree with; and neutral/doubtful). Then patients sorted the cards on the worksheet, starting with the statement they agreed with least, until all statements they disagreed with were on the worksheet. Then they did the same with statements they did agree with. Next, patients sorted the remaining statements in the middle of the worksheet. After finishing the task, they read the items again, so they could rearrange the cards if necessary. Finally, the researcher (S.N.) noted the numbers of the statements (which were written on the back of the cards) on a record sheet with the Q-sort figure.

For analysis the Q-sorts were correlated, whereby a high correlation indicated that participants sorted the cards in a similar way. By-person factor analysis was used to analyse the Q-sort correlation matrix, obtaining factors that represented groups of participants with a similar opinion.[28] Factor scores were calculated for each statement,

which resulted in one representative Q-sort per group. The diversity of participants was more important than the sample size, which was large enough when factor stability and a clear understanding of the factors' meaning were ensured.[38] For an appropriate analysis, considerably more statements were needed than participants (ideally, the number of items should be at least twice the number of participants).[39]

Measurements

Patient characteristics

Patients answered questions about marital status, children living at home, work situation, important life events, exercise, medication and recent medication change, comorbidity, health status, hours of sleep during the night and napping during daytime.

Daily roles

With eight items, patients were asked to indicate how many hours per week they normally spent on the following eight areas: work; study; housekeeping (cooking, shopping, washing, cleaning); caring for people at home (e.g. children, ill spouse, parents); caring for people outside home; voluntary work; gardening and pets. They could choose one of the following six answers: not applicable to me (scored as 0); on average >0 and <3 h/week; >3 and <10 h; >10 and <20 h; >20 and <30 h; and >30 h (scored as 5). By adding up all the scores, we calculated a personal role score (range: 0–40).

Visual analogue scales

Patients completed a visual analogue scale (VAS) for pain and impact of the disease, and three VAS fatigue scales[40] for severity ('Please place a mark on the line to show your level of fatigue over the past 7 days' (anchors: no fatigue - totally exhausted); impact ('Please place a mark on the line to show the effect that fatigue has had on your life over the past 7 days' (anchors: no effect—a great deal of effect); and coping ('Please place a mark on the line to show how well you have coped with fatigue over the past 7 days' (anchors: not at all well—very well). As these three VASs were developed in the UK, they were translated into Dutch according to the recommended translation procedure for cross-cultural adaptation.[41]

Anxiety, depression and disability

The Dutch versions of the anxiety subscale of the Hospital Anxiety and Depression Scale (HADS),[42] the 15-item Geriatric Depression Scale (GDS)[43] and the 10-item version of the HAQ-II[44] were completed.

Laboratory values

The 28-joint DAS, CRP, ESR and haemoglobin (Hb) values were collected.

Analysis

The Q-sort data were analysed with PCQ for Windows,[45] using centroid factor analysis with varimax rotation. The decision for four factors with a significance level of 0.40 was based on the following considerations: a maximum number of Q-sorts loading significantly on one factor; all factors having an eigenvalue >1; a maximal explained variance; no factors with only one significantly loading Q-sort; and as few as possible consensus items and no bipolar factors. For the theoretical interpretation of the factors, we used those statements with extreme factor loadings (+4, +5, +6 and -4, -5, -6). First, we looked at the four patterns of fatigue experience based on the content of these statements. Then, we also looked at which patient variables (see the appendix at the end of this article) characterized the four groups. To this end, we used percentages and median values with inter-quartile ranges. We described the most prominent patient characteristics in each group, without testing for statistically significant differences. Owing to the small sample size in the groups, the power would be too small in such an analysis. However, describing patient characteristics in each group gave interesting indications about which kind of fatigue experience is perceived by which patients. All four authors were involved in the decision process about the number and naming of the factors, whereby professional knowledge from psychology and clinical rheumatology was incorporated.

RESULTS

Sample

The sample consisted of 13 men and 17 women with a mean age of 60 years (S.D. 12.15 years), ranging from 38 to 82. The mean score of daily roles was 6.8 (S.D. 3), ranging from 1 to 12. Further sample characteristics and health-related outcomes are shown in the appendix. The mean disease duration was 9.9 years (S.D. 6 years; range: 2–26 years). The range of VAS fatigue scores for severity (range 1–9), impact (range 1–9) and coping (range 2–10) showed that patients with all levels of fatigue were included. The Cronbach's α values of the scales to measure anxiety, depression and disability were 0.76 (HADS), 0.60 (GDS) and 0.85 (HAQ-II).

Patterns of fatigue experience

Factor analysis showed that the optimal solution for this Q-sort data was a four-factor structure. Thereby, two-thirds of the Q-sorts loaded significantly on one factor. Of the remaining 10 Q-sorts, 8 did not load significant on any factor and 2 were confounders

(significant factor loadings on 2 factors). The explained variance was 40%. The four-factor solution was also the best alternative for theoretical interpretation.

The factor groups are described as follows: those items with extreme factor loadings (+4, +5, +6 and -4, -5, -6) were used for the description of the fatigue experience. Furthermore, per group, the most distinctive patient characteristics are described briefly. These results are summarized in Table 1. Extensive tables are provided in the appendix: table A1 quantitatively summarizes the patient characteristics per factor; table A2 shows the factor scores per statement and eigenvalues, percentage explained variance and median factor scores per factor; and table A3 shows the most characterizing statements per group (statements with the strongest agreement and also the strongest disagreement).

Group A: little impact of fatigue (n = 7)

Group A consisted of seven patients having Q-sorts with negative factor loadings, so the Q-sorts had to be interpreted in the reverse way.[46] It means that, for example, the patients strongly agreed with the statement in column 6. For better readability, the interpretative correct algebraic signs were used in tables A2 and A3 (see appendix). Group A patients reported days without any fatigue, only being tired after doing too much and that the fatigue would pass if they rested. Furthermore, they agreed with having a lot of energy and having good coping competencies. However, the ability to cope well with fatigue varied. These patients belonged to the oldest group, with the fewest daily roles (see table A1of the appendix). Group A had moderate disease activity in absolute terms, which reflects the highest disease activity of all four groups. They scored low on pain, fatigue severity and fatigue effect, and had high fatigue coping scores. More than half of the Group A patients reported to seldom rest during the day. Compared with Groups B and C, Group A had average disease duration and reported less exercising. They had higher disability scores than Group C, but the same as Group B. In this group, no biologicals were used.

Group B: Good coping and bad sleep (n = 5)

Group B patients reported disrupted and non-refreshing sleep. They also had days without any fatigue and could cope with it well. They did not let on to anyone when they were tired and felt physically impeded by their fatigue. Group B consisted of more women than men (see table A1 of the appendix). Compared with the other groups, Group B patients had an average number of daily roles and reported the least number of hours they usually slept during the night. Four of the five patients reported not resting during the day on a regular basis. None of them had fatigue-related comorbidities. They had the longest

disease duration, the lowest fatigue severity and impact scores of all four groups, and also their pain score was low. Four of the five patients in Group B reported exercising on a weekly basis. Moreover, Group B patients rated their general health as better and the impact of their disease on their life as lower than patients in the other groups. Three of the five patients used a biological.

Group C: search for balance (n = 6)

Group C consisted of six patients. They were not refreshed in the morning, took a nap during the day, lived in a constant struggle to find a balance between activity and rest and just carried on when they were fatigued. Moreover, they reported only being tired after doing too much and could cope well with their fatigue. Group C was the youngest group (see table A1 of the appendix). All patients were living with a partner, most of them had paid work and half of them had children living at home. The role score was the highest of all groups. In addition, most of the patients exercised on a weekly basis. This group scored higher on pain, fatigue severity and impact, lower on the ability to cope with fatigue and had shorter disease duration than Groups A and B. Disability scores were lower than those in the other groups. In this group, as many biologicals are used as in Group B and also DMARDs were frequently reported.

Group D: high distress (n = 2)

Group D was the smallest group with two patients. They found it difficult to accept the fatigue, were frustrated when they were too fatigued to do things they wanted to do, had difficulty concentrating when they were fatigued, were often awake during the night and had to plan everything they did because of their fatigue. Group D patients reported the highest levels of pain and fatigue (severity and effect), most comorbidities, high levels of anxiety and depression and the shortest disease duration (see table A1 of the appendix). Group D patients had the highest disability scores, did no regular exercise and their general health was worse. Of all the groups, they reported the most hours' sleep during the night, and regularly taking a nap during the day. They had slightly lower coping fatigue scores on the VAS than patients in the other groups.

TABLE 1 Summary of fatigue experience and the most striking patient characteristics per group

Group	Fatigue experience	Patient characteristics
Group A	Days without any fatigue	Oldest group
n=7	Only fatigued if too much undertaken	Fewest number of daily roles
	Good coping competencies	Moderate but highest disease activity
	Fatigue would pass if they rested	Low pain scores on VAS
	A lot of energy	Low fatigue severity and impact on VAS
	Varying ability to cope with fatigue	High fatigue coping scores on VAS
		Most of them seldom rest during the day
		Average disease duration
		Less exercising than Groups B and C
		Higher disability scores than Group C and comparable to Group B
		No biologicals are used
Group B	Good coping with fatigue	More women than men
n=5	Days without any fatigue	Average number of daily roles
	Not showing fatigue to anyone	Least hours of sleep during the night
	Often awake during the night	Most of them do not rest during the day
	Unrefreshing sleep	No comorbidities
	Fatigue impedes physical activity	Longest disease duration
		Lowest fatigue severity and impact (VAS)
		Low pain scores
		Most of them regularly exercise
		Most positive general health
		Lowest impact of the disease on daily life
		Three patients used a biological
Group C	Trying to control fatigue by seeking balance between activity and rest	Youngest group
n=6		All living with partner
	Taking a nap during the day	Most of them in paid work
	Unrefreshing sleep	Half of them with children at home
	Only fatigued if too much undertaken	Highest role score
	Just carrying on when fatigued	Most of them regularly exercise
	Good coping with fatigue	Higher pain scores on VAS than A and B
		Higher fatigue severity and impact on VAS than Groups A and B
		Lower fatigue coping scores on VAS than Groups A and B
		Shorter disease duration than A and B
		Lowest disability scores
		As many biologicals as in Group B
		Frequent use of DMARDs
Group D	Difficulty in thinking clearly when fatigued	Highest pain scores on VAS
n=2		Highest fatigue severity and impact on VAS
	Often awake during the night	Most with comorbidities
	Difficult to accept fatigue	High on anxiety and depression
	Frustration when too fatigued to do sth	Shortest disease duration
	Need to plan everything	Highest disability scores
	Difficulty in concentrating when fatigued	No regular exercise
		Worst general health
		Most hours of sleep during the night
		Regular nap during the day
		Slightly lower fatigue coping on VAS

The fatigue experiences in this table are based on the statements with extreme factor loadings (+4, +5, +6 and -4, -5, -6).

DISCUSSION

In this study, we identified four different perspectives on the experience of fatigue in RA. Although parts of the fatigue experiences overlapped in the groups, there were remarkable differences, also regarding the patients' living situation and other characteristics.

For Group A patients, fatigue had almost no impact on their lives. Their fatigue was comparable to normal fatigue. The higher age in Group A did not seem to be associated with longer disease duration. So the low impact of fatigue cannot be explained by a longer time of experience of coping with RA. It is more likely that the higher age in this group was related to the few daily roles that patients had to fulfil. After retirement, people have fewer roles,[47] and expectations from others are less demanding. In addition, older people's psychological resources to adapt to fatigue might be higher because they have learned to adjust to changing life circumstances and accommodate their goals and standards with greater flexibility than younger people.[48] Group B patients reported disrupted and non-refreshing sleep, and being physically impeded by their fatigue. However, four of the five patients said they exercised on a weekly basis and, in general, fatigue had little impact on their lives. Some explanations for their good coping ability could be the fact that the patients had the longest disease duration and the least fatigue severity of all groups, and that they had an average but not high number of daily roles. Group C patients experienced more impact of fatigue than Group A and B patients. They knew how to deal with their fatigue, but in line with the high amount of activity in Group C, they lived in a constant struggle to find a balance between activity and rest. That this effort was not always successful is reflected by their agreement to just carry on when they were fatigued (Item 54). This struggle might be related to the fact that Group C was the youngest group with the most daily roles. Those patients with multiple roles (e.g. housekeeping, working, being a spouse, parent), seemed to live in a constant tension between adequate fulfilment of their roles and the need for rest or relaxing activities. This finding is in line with the results of our interview study that the impact of fatigue was worst in young women with multiple roles.[12] Patients in Group D were very distressed and fatigue had a great impact on their lives. Compared with the other groups, they had the most physical and mental health problems. The high levels of physical and psychological distress in this group probably made it difficult for patients to accept their fatigue and cope with it in a constructive way.

This study is descriptive and cannot examine causal relationships. Our aim was not to find such relationships or statistically significant differences between groups. Rather, we made a first attempt to describe dominant perspectives on fatigue experience in RA and to

describe the patient groups which share a certain perspective. However, it is an outstanding finding that Group A patients had the highest but still moderate disease activity, although they experienced the least impact of fatigue. We assume that this points to the importance of psychological and social aspects, beside physical ones, in the experience of fatigue in RA. It seems likely that the fatigue experience in patients with established controlled RA is not so much influenced by disease-specific characteristics as by personal and situational characteristics. The differences in fatigue experience appeared to be related to tasks, external demands and coping resources. The complexity of the fatigue experience in RA is further underlined by the fact that one-third of the patients could not be allocated to one of the groups. Furthermore, there was no item with which all patients strongly agreed (sorted between +4 and +6 in all groups) or disagreed (sorted between -4 and -6 in all groups). This shows that measurement instruments should include items that reflect a broad spectrum of fatigue experiences in order to be able to adequately describe fatigue experiences in RA patients and to distinguish between patients with different patterns of fatigue experience.

When looking at the existing multi-item fatigue scales, MAF, FACIT-F, POMS, SF-36 vitality subscale, CIS and the BRAF-MDQ, it becomes obvious that none of them covers all categories. A comparable conclusion has already been drawn by Nicklin *et al.*,[49] who showed that neither the MAF, SF-36 vitality subscale, POMS or FACIT-F covers all of the following patient-reported categories of fatigue: frequency/severity, duration, energy, impact, emotion, coping, social life, planning, relationships, cognition, quality of life and sleep. All of these categories also turned out to be relevant in this study; the categories social life, planning, relationships and quality of life were included in the 'consequences' category in our study. The CIS consisting of the dimensions 'subjective feeling of fatigue', 'concentration', 'motivation' and 'physical activity' does not cover all dimensions either. The BRAF-MDQ was recently developed and validated in a British population and is based on interviews with patients scoring seven or higher on a fatigue severity VAS. So we assume that there is still no measurement instrument available which embraces the full diversity of the subjective experience of fatigue with its several dimensions and possible ranges of fatigue severity. To be able to measure fatigue in RA with all its facets, new ways of measurement are needed in the future. Not only should more aspects of fatigue then be included but also new developments in measurement applications should be used. Computer-adaptive testing seems to be a promising method to comprehensively measure fatigue with relatively few items per patient.[50]

New measurement instruments should also be used for screening purposes. In clinical practice, a patient's perspective on fatigue experience should be identified. Consequently, it would be possible to provide tailored support. Patients with a fatigue experience such as

patients in Group D might benefit from a psycho-social intervention that is focused on the high level of distress, whereas patients with an experience similar to patients in Group C might be aided by an approach emphasizing the balancing of roles and demands. Patients with a fatigue experience such as patients in Group B could probably benefit from support focused on sleep. However, for patients with an experience similar to patients in Group A, support seems unnecessary.

Our sample of 30 RA outpatients was sufficient for using Q-methodology.[28,38,39] Since this research raises new hypotheses, generalizations to the entire population of RA patients should be made with caution and need more research. However, evidence for the test–retest reliability and also the reliability across subjects in Q-sort results could be identified in research concerning different topics.[24] All patients in our study had a disease duration of at least 2 years. We acknowledge that we missed patients with recently developed RA and patients in flare. As those patients mostly have anaemia, high disease activity with fever and very high levels of fatigue, it is important to include them in future research on fatigue experiences. Moreover, DASs were relatively low in this sample, which raises the question of to what extent our results can be generalized to patients with high disease activity. However, nowadays, high disease activity is not very common among RA outpatients. In a recent study about the patients' perspective on fatigue in RA, relatively low disease activity was reported in a comparable sample also.[11]

This study revealed different types of fatigue experience, reflecting the patients' perspective and showing the match with patient characteristics. Results indicated that fatigue experience is a complex phenomenon related to physical, psychological and social resources and competencies.

ACKNOWLEDGEMENTS

We would like to thank the patients who participated in this study and also Alisa van Zonneveld for assistance with data collection.

REFERENCES

1. Wolfe F, Hawley DJ, Wilson K. The prevalence and meaning of fatigue in rheumatic disease. J Rheumatol 1996;23:1407–17.
2.Belza BL, Henke CJ, Yelin EH, Epstein WV, Gilliss CL. Correlates of fatigue in older adults with rheumatoid arthritis. Nurs Res 1993;42:93–9.
3. Hewlett S, Carr M, Ryan S et al. Outcomes generated by patients with rheumatoid arthritis: how important are they? Musculoskel Care 2005;3:131–42.
4. Kirwan JR, Newman S, Tugwell PS, Wells GA. Patient perspective on outcomes in rheumatology—a position paper for OMERACT 9. J Rheumatol 2009;36:2067–70.
5. Dittner AJ, Wessely SC, Brown RG. The assessment of fatigue. A practical guide for clinicians and researchers. J Psychosom Res 2004;56:157–70.
6. Hewlett S, Hehir M, Kirwan JR. Measuring fatigue in rheumatoid arthritis: a systematic review of scales in use. Arthritis Rheum 2007;57:429–39.
7. Swain MG. Fatigue in chronic disease. Clin Sci 2000;99: 1–8.
8. Hewlett S, Nicklin J, Treharne GJ. Fatigue in musculoskeletal conditions. Topical Reviews Series 6: No 1. Arthritis Campaign: Autumn 2008 www.arc.org.uk/arthinfo/medpubs/6641 /6641.asp (1 December 2009, date last accessed).
9. Belza Tack B. Fatigue in rheumatoid arthritis: conditions, strategies, and consequences. Arthritis Care Res 1990;3: 65–70.
10. Hewlett S, Cockshott Z, Byron M et al. Patients' perceptions of fatigue in rheumatoid arthritis: overwhelming, uncontrollable, ignored. Arthritis Care Res 2005;53:697–702.
11. Repping-Wuts H, Uitterhoeve R, van Riel P, van Achterberg T. Fatigue as experienced by patients with rheumatoid arthritis (RA): a qualitative study. Int J Nurs Stud 2008;45:995–1002.
12. Nikolaus S, Bode C, Taal E, van de Laar MAFJ. New insights into the experience of fatigue among patients with rheumatoid arthritis: a qualitative study. Ann Rheum Dis 2010;69:895–7.
13. Huyser BA, Parker JC, Thoreson R, Smarr KL, Johnson JC, Hoffman R. Predictors of subjective fatigue among individuals with rheumatoid arthritis. Arthritis Rheum 1998;41:2230–7.
14. Pollard LC, Choy EH, Gonzalez J, Khoshaba B, Scott DL. Fatigue in rheumatoid arthritis reflects pain, not disease activity. Rheumatology 2006;45:885–9.
15. Riemsma RP, Rasker JJ, Taal E, Griep EN, Wouter JMGW, Wiegman O. Fatigue in rheumatoid arthritis: the role of self-efficacy and problematic social support. Br J Rheumatol 1998;37:1042–6.

16. Repping-Wuts H, Fransen J, van Achterberg T, Bleijenberg G, van Riel P. Persistent severe fatigue in patients with rheumatoid arthritis. J Clinical Nurs 2007;16: 377–83.
17. Mancuso CA, Rincon M, Sayles W, Paget SA. Psychosocial variables and fatigue: a longitudinal study comparing individuals with rheumatoid arthritis and healthy controls. J Rheumatol 2006;33:1496–502.
18. Treharne GJ, Lyons AC, Hale ED, Goodchild CE, Booth DA, Kitas GD. Predictors of fatigue over 1 year among people with rheumatoid arthritis. Psychol Health Med 2008;13:494–504.
19. Younger J, Finan P, Zautra A, Davis M, Reich J. Personal mastery predicts pain, stress, fatigue, and blood pressure in adults with rheumatoid arthritis. Psychol Health 2008; 23:515–35.
20. Parrish BP, Zautra A, Davis MC. The role of positive and negative interpersonal events on daily fatigue in women with fibromyalgia, rheumatoid arthritis, and osteoarthritis. Health Psychol 2008;27:694–702.
21. Davis MC, Zautra AJ, Younger J, Motivala SJ, Attrep J, Irwin MR. Chronic stress and regulation of cellular markers of inflammation in rheumatoid arthritis: implications for fatigue. Brain Behav Immun 2008;22:24–32.
22. Repping-Wuts H, van Riel P, van Achterberg T. Fatigue in patients with rheumatoid arthritis: what is known and what is needed. Rheumatology 2009;48:207–9.
23. Redburn FS. Q factor analysis: application to educational testing and program evaluation. Educ Psychol Meas 1975; 35:767–78.
24. Ten Klooster PM, Visser M, de Jong MDT. Comparing two image research instruments: the Q-sort method versus the Likert Attitude Questionnaire. Food Qual Prefer 2008;19: 511–8.
25. Risdon A, Eccleston C, Crombez G, McCracken L. How can we learn to live with pain? A Q-methodological analysis of the diverse understandings of acceptance of chronic pain. Soc Sci Med 2003;56:375–86.
26. Lai JS, Kupst MJ, Cella D, Brown SR, Peterman A, Goldman S. Using Q-methodology to understand perceived fatigue reported by adolescents with cancer. Psychooncology 2007;16:437–47.
27. Barbosa JC, Willoughby P, Rosenberg CA, Mrtek RG. Statistical methodology: VII. Q-methodology, a structural analytic approach to medical subjectivity. Acad Emerg Med 1998;5:1032–40.
28. Brown SR. Q technique and method: principles and procedures. In: Berry WD, Lewis-Beck MS, eds. New tools for social scientists: advances and applications in research methods. Beverly Hills, CA: Sage, 1986:57–76.

29. Belza Tack B. Dimensions and correlates of fatigue in older adults with rheumatoid arthritis [dissertation]. San Francisco: University of California, 1991.
30. Belza B. Comparison of self-reported fatigue in rheumatoid arthritis and controls. J Rheumatol 1995;22:639–43.
31. Cella D, Webster K. Linking outcomes management to quality-of-life measurement. Oncology 1997;11:232–5.
32. Yellen SB, Cella DF, Webster K, Blendowski C, Kaplan E. Measuring fatigue and other anemia-related symptoms with the Functional Assessment of Cancer Therapy (FACT) measurement system. J Pain Symptom Manage 1997;13: 63–74.
33. McNair D, Lorr M, Droppelman L. Profile of Mood States Manual. New York: Multi-health Systems Inc., 1992.
34. Ware JE Jr, Sherbourne CD. The MOS 36-Item Short-Form Health Survey (SF-36). I. Conceptual framework and item selection. Med Care 1992;30:473–83.
35. Vercoulen JHHM, Swanink CMA, Fennis JFM, Galama JMD, van der Meer JWM, Bleijenberg G. Dimensional assessment of chronic fatigue syndrome. J Psychosom Res 1994;38:383–92.
36. Van Hoogmoed D, Fransen J, Bleijenberg G, van Riel PLCM. How to assess fatigue in rheumatoid arthritis: validity and reliability of the checklist individual strength. Arthritis Rheum 2008;58:868.
37. Nicklin JK, Kirwan J, Cramp F, Urban M, Hewlett S. Development and initial validation of the Bristol RA Multi-Dimensional Questionnaire (BRAF-MDQ). Rheumatology 2009;48:12.
38. McKeown BF, Thomas DB. Q-methodology. Newbury Park, CA: Sage Publications, 1988.
39. Thompson B. Q-technique factor analysis: one variation on the two-mode factor analysis of variables. In: Grimm LG, Yarnold PR, eds. Reading and understanding more multivariate statistics. Washington DC: American Psychological Association, 2002:207–26.
40. Nicklin JK, Kirwan JR, Cramp F, Hewlett S. Standardizing visual analogue scales to measure fatigue in rheumatoid arthritis. Arthritis Rheum 2007;56:S778.
41. Guillemin F, Bombardier C, Beaton D. Cross-cultural adaptation of health-related quality of life measures: literature review and proposed guidelines. J Clin Epidemiol 1993;46:1417–32.
42. Zigmond AS, Snaith RP. The Hospital Anxiety and Depression Scale. Acta Psychiatr Scand 1983;67:361–70.
43. Sheikh JI, Yesavage JA, Brooks JO, Friedman L, Gratzinger P. Proposed factor structure of the Geriatric Depression Scale. Int Psychogeriatr 1991;3:23–8.
44. Wolfe F, Michaud K, Pincus T. Development and validation of the Health Assessment Questionnaire II. Arthritis Rheum 2004;50:3296–305.

45. Stricklin M, Almeida R. PCQ for Windows (Academic Edition 1.4, 2004). http://pcqsoft.com/ (1 December 2009, date last accessed).
46. Ramlo SE, McConnell D, Duan Z-H, Moore FB. Evaluating an inquiry-based bioinformatics course using Q methodology. J Sci Educ Technol 2008;17:219–25.
47. Riley M, Kahn RL, Foner N, eds., Age and structural lag. New York: Wiley, 1994.
48. Brandtstädter J, Rothermund K. The life-course dynamics of goal pursuit and goal adjustment: a two-process framework. Dev Rev 2002;22: 117–50.
49. Nicklin JK, Kirwan J, Cramp F, Hewlett S. Do existing scales capture the multi-dimensional aspects of RA fatigue? Presentation ACR/ARHP Scientific Meeting, San Francisco, CA, October 2008.
50. Rose M, Bezjak A. Logistics of collecting patientreported outcomes (PROs) in clinical practice: an overview and practical examples. Qual Life Res 2009;18: 125–36.

APPENDIX

Table A1 Summary table patient characteristics

Characteristics	Group A (N=7)	Group B (N=5)	Group C (N=6)	Group D (N=2)	Other (N=10)	Total group (N=30)
Gender						
Women, n (%)	4 (57%)	4 (80%)	3 (50%)	1 (50%)	5 (50%)	17 (57%)
Men, n (%)	3 (43%)	1 (20%)	3 (50%)	1 (50%)	5 (50%)	13 (43%)
Age						
Median	78.00	54.00	54.50	58.50	58.00	59.50
Inter quartile range (IQR)	66.00-79.00	46.00-64.50	40.25-60.00	45.00-72.00	56.00-64.25	53.75-69.00
Marital status						
Married, n (%)	3 (43%)	5 (100%)	6 (100%)	1 (50%)	7 (70%)	22 (73%)
Widowed, n (%)	4 (57%)	0 (0%)	0 (0%)	0 (0%)	1 (10%)	5 (17%)
Single, n (%)	0 (0%)	0 (0%)	0 (0%)	1 (50%)	2 (20%)	3 (10%)
Children at home						
Yes, n (%)	0 (0%)	2 (40%)	3 (50%)	1 (50%)	2 (20%)	8 (27%)
No, n (%)	7 (100%)	3 (60%)	3 (50%)	1 (50%)	8 (80%)	22 (73%)
Paid work						
Yes, n (%)	1 (14%)	2 (40%)	5 (83%)	1 (50%)	3 (30%)	12 (40%)
No, n (%)	6 (86%)	3 (60%)	1 (17%)	1 (50%)	7 (70%)	18 (60%)
Weekly exercise						
Yes, n (%)	3 (43%)	4 (80%)	5 (83%)	0 (0%)	10 (100%)	22 (73%)
No, n (%)	4 (57%)	1 (20%)	1 (17%)	2 (100%)	0 (0%)	8 (27%)
Medication						
DMARDs, n (%)	7 (100%)	3 (50%)	6 (100%)	1 (50%)	8 (80%)	25 (83%)
NSAIDs, n (%)	0 (0%)	1 (20%)	2 (33%)	0 (0%)	4 (40%)	7 (23%)
Prednisolon, n (%)	3 (43%)	0 (0%)	0 (0%)	0 (0%)	1 (10%)	4 (13%)
Biologicals, n (%)	0 (0%)	3 (60%)	3 (50%)	1 (50%)	2 (20%)	9 (30%)
Co-morbidity (fatigue-related)						
Yes, n (%)	1 (14%)	0 (0%)	1 (17%)	2 (100%)	3 (30%)	7 (23%)
No, n (%)	6 (86%)	5 (100%)	5 (83%)	0 (0%)	7 (70%)	23 (77%)
Health status						
Very good, n (%)	0 (0%)	2 (40%)	0 (0%)	0 (0%)	0 (0%)	2 (7%)
Good, n (%)	5 (71%)	3 (60%)	4 (67%)	1 (50%)	8 (80%)	21 (70%)
Good–moderate, n (%)	0 (0%)	0 (0%)	1 (17%)	0 (0%)	0 (0%)	1 (3%)
Moderate, n (%)	2 (29%)	0 (0%)	1 (17%)	0 (0%)	2 (20%)	5 (17%)
Bad, n (%)	0 (0%)	0 (0%)	0 (0%)	1 (50%)	0 (0%)	1 (3%)

Characteristics	Group A (N=7)	Group B (N=5)	Group C (N=6)	Group D (N=2)	Other (N=10)	Total group (N=30)
Hours sleep						
Less 6, n (%)	0 (0%)	1 (20%)	0 (0%)	0 (0%)	1 (10%)	2 (7%)
6-7, n (%)	1 (14%)	3 (60%)	1 (17%)	0 (0%)	4 (40%)	9 (30%)
7-8, n (%)	4 (57%)	0 (0%)	3 (50%)	1 (50%)	5 (50%)	13 (43%)
More 8, n (%)	2 (29%)	1 (20%)	2 (33%)	1 (50%)	0 (0%)	6 (20%)
Regular nap during day						
Yes, n (%)	2 (29%)	1 (20%)	5 (83%)	2 (100%)	6 (60%)	16 (53%)
No, n (%)	5 (71%)	4 (80%)	1 (17%)	0 (0%)	4 (40%)	14 (47%)
Disease duration						
Median	10.00	15.00	5.50	5.00	10.00	10.00
IQR	5.25-11.75	9.00-17.5	3.75-12.25	3.00-7.00	5.25-17.00	5.00-14.00
Role-score						
Median	3.00	7.00	10.00	6.00	6.50	7.00
IQR	2.00-6.00	5.50-7.00	8.50-12.00	2.00-10.00	5.75-8.25	5.00-9.00
Pain						
Median	2.00	2.00	3.00	6.00	5.00	3.00
IQR	1.00-4.00	2.00-3.00	2.75-4.75	3.00-9.00	2.75-6.00	2.00-5.00
Impact of disease						
Median	3.00	2.00	4.00	6.50	3.50	3.00
IQR	2.00-5.00	2.00-2.50	3.00-6.25	5.00-8.00	2.75-5.00	2.00-5.00
Fatigue severity						
Median	3.00	3.00	4.50	7.00	6.50	5.00
IQR	2.00-5.00	2.00-3.50	2.75-7.00	5.00-9.00	5.00-7.20	3.00-7.00
Impact of fatigue						
Median	3.00	2.00	3.50	7.00	5.50	3.00
IQR	2.00-5.00	1.50-2.50	2.00-6.25	5.00-9.00	3.00-7.25	2.00-6.00
Coping with fatigue						
Median	8.00	8.00	6.50	6.00	7.00	7.00
IQR	6.00-8.00	4.50-8.50	3.50-7.75	5.00-7.00	5.75-8.00	6.00-8.00
HADS anxiety						
Median	3.00	4.00	3.00	12.00	5.00	4.00
IQR	1.00-6.00	2.50-8.00	2.50-3.25	11.00-13.00	3.75-5.00	3.00-5.00
GDS depression						
Median	2.00	1.00	1.00	6.00	2.00	2.00
IQR	2.00-2.00	0.50-4.00	0.00-2.50	6.00-6.00	1.50-4.75	1.00-4.00
HAQ-II disability						
Median	0.90	0.90	0.70	1.10	0.90	0.85
IQR	0.20-1.30	0.50-1.20	0.38-0.90	0.70-1.50	0.73-1.00	0.58-1.13
DAS 28						
Median	3.33	1.96	2.40	2.52	2.91	2.51
IQR	2.08-4.90	1.20-3.00	1.98-2.91	1.23-3.81	2.11-3.54	2.05-3.39

Characteristics	Group A (N=7)	Group B (N=5)	Group C (N=6)	Group D (N=2)	Other (N=10)	Total group (N=30)
Hb value						
Median	7.70	9.30	8.50	9.20	8.30	8.30
IQR	6.43-8.08	8.15-9.6	8.08-9.25	9.00-9.40	8.05-8.65	8.00-9.10

Table A2 Item scores (varimax)

Items	During the past 7 days… Groups	A	B	C	D
01	…I was angry about my fatigue	-2	-4	-3	0
02	…I was despondent about my fatigue	-3	-2	-3	0
03	…I was sad about my fatigue	-1	-2	-5	2
04	…I was miserable about my fatigue	-1	-5	-1	1
05	…I was frustrated if I was too tired to do things I wanted to do	1	-1	-1	4
06	…I felt useless because of my fatigue	0	-5	-3	0
07	…I was worried about my fatigue	-2	-3	-3	-4
08	…I was embarrassed about my fatigue	-2	-4	-4	-2
09	…I was upset about my fatigue	-4	-4	-4	-1
10	…I got irritated quickly because of my fatigue	-4	-2	1	2
11	…I wasn't able to do anything because of my fatigue	0	0	0	2
12	…my fatigue prevented me from any physical activity (walking, cycling, sports, etc.)	-1	4	0	2
13	…I was too tired to do any chores in and around the house (cleaning, cooking, gardening, etc.)	1	1	0	-1
14	…my fatigue meant it was difficult to do the shopping	-1	2	0	0
15	…my fatigue had a negative influence on the relationship with my partner and/ or others close to me	-4	0	-1	-5
16	…my fatigue restricted in my contacts with family, friends and acquaintances	-1	-2	-4	2
17	…I was too tired to help others	-2	1	1	1
18	…I was too tired to do my most important tasks properly	0	1	-1	-1
19	…my fatigue restricted me from doing pleasant or relaxing things	1	1	-1	3
20	…my fatigue made me unhappy about my life	-6	-6	-2	-1
21	…I experienced the fatigue as a handicap	-1	-3	-2	-2
22	…I was also able to see positive sides of my fatigue	3	-1	-1	-6
23	…I have experienced different forms of fatigue	1	0	2	0
24	…my fatigue was unpredictable	0	2	-2	-5
25	…I was sometimes too tired to eat	-3	-1	-5	-2
26	…I felt exhausted	-1	-1	1	-2
27	…my fatigue would pass if I rested	4	3	3	-2
28	…I was continually tired	-5	-2	-2	1
29	…I was tired almost every day	-3	-3	0	3
30	…the fatigue would take me by surprise	-3	3	-2	-3
31	…I was only tired if I had done too much	5	2	4	-3
32	…I tired quickly	2	-1	2	0
33	…I also had days when I wasn't tired at all	6	5	2	-3
34	…I had difficulty thinking clearly when I was tired	2	0	1	6

35	…I had difficulty concentrating when I was tired	3	0	2	4
36	…I had a lot of energy	4	3	0	-4
37	…I felt like doing lots of pleasant things	3	2	2	-4
38	…I was often awake during the night	-5	4	0	5
39	…I did not awake in the morning feeling refreshed	0	4	5	1
40	…I took a nap during the day	0	2	5	-1
41	…I couldn't help falling asleep	1	1	-6	1
42	…the fatigue made my body feel heavy	-1	1	2	0
43	…the fatigue made me feel listless	-2	0	1	0
44	…the fatigue made me feel mentally worn out	2	-1	0	-1
45	…my fatigue also resulted in other rheumatic symptoms	0	-1	1	3
46	…others helped me with my fatigue by doing things for me	-2	0	1	-2
47	…others were understanding about my fatigue	2	1	3	1
48	…I didn't let on to anyone that I was tired	0	5	3	1
49	…I was able to cope well with my fatigue	5	6	4	-1
50	…it varied how well I could cope with my fatigue	4	-2	3	3
51	…I found it difficult to accept the fatigue	1	-3	-1	5
52	…I tried everything to decrease my fatigue	1	0	-1	-1
53	…I sought distraction when I was tired	2	3	0	2
54	…I just carried on, however tired I was	3	0	4	1
55	…I rested more because of my fatigue	0	2	1	0
56	…my fatigue meant I had to plan everything I did	1	-1	-2	4
57	…I tried to control my fatigue by seeking a balance between activity and rest	2	1	6	-3
Median factor scores		-0.57	0.61	0.53	0.51
Eigenvalues		3.70	3.46	3.36	1.57
Percentages variance explained		12	12	11	5

Items with differences in factor scores between the groups ≥ 3 are marked in gray. These items are able to distinguish between the groups, meaning they are relevant items to describe and distinguish fatigue experience.

Table A3 Agreement and disagreement with distinguishing items in the four groups

Group	**Agreement with statement (factor score)**	**Disagreement with statement (factor score)**
A	33: having days without any fatigue (+6) 31: only being tired if too much undertaken (+5) 49: good coping competencies (+5) 36: having a lot of energy (+4) 27: fatigue would pass if they rested (+4) 50: varying ability to cope well with their fatigue (+4)	20: fatigue makes them unhappy about their life (-6) **28: being continually tired (-5)** **38: often being awake during the night (-5)** 9: being upset about fatigue (-4) 10: getting irritated quickly because of fatigue (-4) 15: negative influence of fatigue on relationship with partner and/or others close to them (-4)
B	49: good coping competencies (+6) 33: having days without any fatigue (+5) 48: not showing fatigue to anyone (+5) 38: often being awake during the night (+4) 39: unrefreshing sleep (+4) 12: fatigue prevents them from any physical activity (+4)	20: fatigue makes them unhappy about their life (-6) 6: feeling useless because of fatigue (-5) **4: being miserable about fatigue (-5)** 1: being angry about fatigue (-4) 9: being upset about fatigue (-4)
C	**57: trying to control fatigue by seeking balance between activity and rest (+6)** **40: taking a nap during the day (+5)** 39: unrefreshing sleep (+5) 31: only being tired if too much undertaken (+4) 54: just carrying on when they were fatigued (+4) 49: good coping competencies (+4)	**41: falling asleep unintendedly (-6)** **3: being sad about fatigue (-5)** 25: being too tired to eat (-5) 9: being upset about fatigue (-4) 16: being restricted in contacts with family, friends and acquaintances because of fatigue (-4)
D	**34: difficulties thinking clearly when they are tired (+6)** 38: often being awake during the night (+5) **51: difficulties to accept fatigue (+5)** **5: being frustrated when too tired to do things they want to do (+4)** **56: having to plan everything they do because of fatigue (+4)** 35: having difficulties concentrating when they are tired (+4)	**22: seeing positive aspects of fatigue (-6)** **24: fatigue being unpredictable (-5)** 15: negative influence of fatigue on relationship with partner and/or others close to them (-5) **36: having a lot of energy (-4)** **37: feeling like doing lots of pleasant things (-4)**

Statements are reported in bold letters if they were not only able to distinguish between groups in general (differences in factor scores between groups were ≥ 3), but also distinguished one group from all the others (difference of ≥ 3 between its factor score and the factor scores of the three other groups).

Chapter 6

Selection of items for a computer-adaptive test to measure fatigue in patients with rheumatoid arthritis: a Delphi approach

S. Nikolaus
C. Bode
E. Taal
M.A.F.J. van de Laar

Qual Life Res. Published online 31 July 2011. doi: 10.1007/s11136-011-9982-8

ABSTRACT

Purpose Computer-adaptive tests (CATs) can measure precisely at individual level with few items selected from an item bank. Our aim was to select fatigue items to develop a CAT for rheumatoid arthritis (RA) and include expert opinions that are important for content validity of measurement instruments.

Methods Items were included from existing fatigue questionnaires and generated from interview material. In a Delphi procedure, rheumatologists, nurses, and patients evaluated the initial pool of 294 items. Items were selected for the CAT development if rated as adequate by at least 80% of the participants (when 50% or less agreed, they were excluded). Remaining items were adjusted based on participants' comments and re-evaluated in the next round. The procedure stopped when all items were selected or rejected.

Results A total of 10 rheumatologists, 20 nurses, and 15 rheumatoid arthritis patients participated. After the first round, 96 of 294 items were directly selected. Nine items were directly excluded, and remaining items were adjusted. In the second round, 124 items were presented for reevaluation. Ultimately, 245 items were selected.

Conclusion This study revealed a qualitatively evaluated item pool to be used for the item bank/CAT development. The Delphi procedure is a beneficial approach to select adequate items for measuring fatigue in RA.

INTRODUCTION

Computerized adaptive testing allows comprehensive measurement of patient-reported outcomes (PROs) with relatively few items.[1] Items are, respectively, selected from an item bank, based on a patient's previous answer, so that precise measurement at individual level with few items becomes possible. For the computerized selection of the best matching items, a large item pool is needed that contains more items than are presented to a patient.[2] Before a computer-adaptive test (CAT) can be developed, such an item pool has to be scaled according to item response theory (IRT). With IRT, item parameters as the difficulty level can be assessed for each item independently.[2] This information is required to ideally match the items to the patient's individual level and for inter-individual comparisons on the measured construct even if patients filled in different items. Primarily, CATs were used for ability and achievement testing, but the interest in computerized adaptive testing for health-related measures is growing,[3] such as attempts to use it for cognitive testing in dementia[4] or assessment of cancer-related fatigue.[5,6] Furthermore, a joint initiative is currently developing a Patient-Reported Outcome Measurement Information System (PROMIS) that aims to construct a large item bank and computerized adaptive testing system for assessing patient-reported outcomes in chronic diseases.[7] Nevertheless, few computerized adaptive testing applications are validated and implemented in medical settings so far.[3] Recently constructed CATs for depression, anxiety, and stress perception appeared to be reliable, valid, and efficient instruments that measure more precisely than traditional questionnaires.[8–12] Moreover, as they are shorter, the burden for patients is reduced.

Also, measurement of fatigue in rheumatoid arthritis (RA) could benefit from this new measurement methodology and technology. RA is a chronic autoimmune disease and is characterized by inflammation of the joints.RA is a chronic autoimmune disease and is characterized by inflammation of the joints.[13] Many RA patients experience fatigue[14,15] and rate it as one of their most annoying symptoms.[16,17] The number of studies on RA incorporating fatigue as outcome has rapidly increased over the last years. Experts endorse the inclusion of fatigue in the core set of outcome variables in clinical trials.[18,19] Measuring fatigue provides additional information on disease outcomes that does not overlap other established measures of RA and is essential for the understanding of the patients' perspective.[20] Qualitative studies on the experience of fatigue in RA have shown that patients experience fatigue as a multidimensional, annoying symptom with far-reaching consequences.[21–24]

Several multi-item questionnaires are used in clinical practice and research. Hewlett et al. conducted a systematic review of measurements of fatigue in RA,[25] whereby a validation

list was applied covering validity (face validity, content validity, criterion, and construct validity), reliability (internal consistency and stability), and sensitivity to change and feasibility. This review has shown that reasonable evidence for validity in RA was found for only four of the used questionnaires: Short Form 36 subscale vitality (SF-36),[26] Functional Assessment of Chronic Illness Therapy Fatigue Scale (FACIT-F),[27] Profile of Mood States subscale fatigue/inertia (POMS),[28] and the Multidimensional Assessment of Fatigue Scale (MAF).[29] However, none of these met all criteria for validity, so further validation in RA patients is needed.[25] The main problem concerned content validity; a limitation of these traditional instruments is that the perspective of RA patients was not included during their development. For instance, instruments include items that reflect disability or inflammation in RA rather than fatigue.[25] Moreover, different time frames are used in the questionnaires: the past 4 weeks, the past 7 days, right now, and the past week. It is not clear which time frame is the most appropriate.

The Bristol RA Fatigue Multi-Dimensional Questionnaire (BRAF-MDQ)[30] is a new instrument, recently developed and evaluated in a British RA population. Although it has not yet been tested in other countries, the BRAF-MDQ is promising since it was developed from the patient's perspective with interviews, focus groups, and cognitive testing. The aim of this study was to select items for the development of a CAT for fatigue in RA.

Before IRT analyses can be conducted, an adequate item pool must be designed. In the development of the stress- CAT, for example, items from fixed-length questionnaires were collected and subsequently selected. Items were independently rated according to their ability to represent the construct under consideration by five members of the research team. In cases of non-agreement, items were openly discussed and rated again. Only items achieving full consensus for inclusion remained in the item bank.[12]

It is not yet clear how valid existing fatigue questionnaires are for RA, especially for the Dutch RA population. Moreover, we did not know how much face validity the items generated from the interviews would have. Therefore, we opted for an extensive Delphi approach and let patients, rheumatologists, and nurses evaluate fatigue items. With the Delphi technique, opinions are collected individually by a questionnaire in at least two rounds accompanied with systematic feedback on the results of the previous round, aiming to reach consensus about a certain topic among a group of experts.[31] The opinion of patients and professionals in the field is essential in the development of questionnaire items to ensure content validity.[32,33] Only patients can report on the subjective experience of fatigue[34,35] as clinicians have the most experience with the outward manifestation of a symptom or condition.[34]

In our Delphi study, we presented a question about the appropriate time frame to measure fatigue in RA and included all fatigue items from traditional questionnaires (SF-

36, POMS, FACIT-F,MAF), the BRAF-MDQ, and newly items generated from interview material and other questionnaires. The four traditional scales were included since they turned out to be the most adequate in RA;[25] moreover, we were interested in the experts' evaluation thereof because their perspective had not been included in the scale development. Our aim was to check whether our item pool contained all relevant aspects to assess fatigue in RA and whether items and response options were adequate and clear.
This paper focuses on the method of the Delphi study as able to select CAT items for clinical populations and the appropriate time frame to measure fatigue in RA. Related concise papers discuss which dimensions of fatigue should be measured[36] and report on the evaluation of the traditional questionnaires and the BRAF-MDQ.[37]

METHODS

Delphi process

With the Delphi process.[38] opinions about a certain topic can be collected by a questionnaire, which is sent by (electronic) mail to a panel of potential participants who then fill it out individually. Contrary to group discussion, it is avoided that especially more dominant participants express their opinion. So the phenomenon of ''group think'' cannot obstruct the viewing of different opinions. A Delphi study consists of at least two rounds and aims to reach consensus among the participants. After each Delphi round, systematic feedback on the results of the previous round is provided to the participants.[31] By doing so, the experts are informed about the opinions that are present in the group. The extent of agreement is determined by statistical measures as no in person meetings are conducted.

Preparation of the item pool

In our aim to develop a CAT for fatigue in RA, we needed a large and comprehensive item pool to cover all dimensions of the fatigue experience in RA. We included the items of the validated Dutch versions of those multi-item questionnaires for which reasonable evidence of validity in RA could be identified[25]: RAND SF-36 subscale vitality,[39] FACIT-F,[40] POMS subscale fatigue/inertia,[41] and MAF.[42] We also included all items of the BRAF-MDQ, which we translated from English into Dutch using the recommended cross-cultural translation procedure.[43] All items of these five scales were sorted according to their dimension (e.g., frequency, duration, severity, consequences). This was supplemented with additional items based on interview material on the experience of fatigue in patients with RA[24] and used in our previous Q-sort study.[44] To check for possibly missing aspects,

we also looked at fatigue questionnaires that are not validated in RA. To ensure that the most prominent fatigue questionnaires were included in our check, recent papers reviewing fatigue scales were used as a guide.[25,45–51] We systematically compared our item pool with the fatigue items included in the preliminary PROMIS fatigue item bank[52] and other fatigue questionnaires as the Checklist Individual Strength (CIS), Cancer Fatigue Scale (CFS), Chronic Fatigue Syndrome— Activities and Participation Questionnaire (CFS-APQ), Fatigue Assessment Instrument (FAI), Functional Impact of Fatigue (FIS), Fatigue Questionnaire (FQ), Lee fatigue scale, Multidimensional Fatigue Inventory (MFI), Multidimensional Fatigue Symptom Inventory (MFSI), Piper scale, Profile of Fatigue (ProF), and Wu fatigue scale. About 70 additional items were constructed as a result of these questionnaires and added to our item pool.

This procedure resulted in a large item pool of 294 items, spread over 12 dimensions of fatigue (severity, frequency, duration, changes in fatigue, perceived causes of fatigue, energy, sleep/rest, body feeling, cognition/concentration, coping, negative emotions/mood, consequences, see Table 1). These were based on dimensions of the existing fatigue questionnaires and were supplemented with additional dimensions that came out of our Q-sort study and interview studies on the experience of fatigue.[21–24] Items taken from an already existing measurement scale were included in their original formulation and with original response options so that the item pool ultimately included items with different kinds of formulations (e.g., statements, questions) and response options (e.g., level of agreement, frequency).

Participants and data collection

We contacted the experts (40 rheumatologists, 40 nurses, and 31 patients) by e-mail, informing them about the study and requesting their participation. E-mail addresses of patients were collected from the database of Patient Research Partners of the Arthritis Centre Twente and those of rheumatologists and nurses from the member lists of a Dutch professional association in Rheumatology (NVR) and the DREAM registry. When selecting potential participants, we aspired to reach a distributed sample across the Netherlands. E-mails included a link to the online survey. We used SurveyMonkey[53] to build the Delphi questionnaire, in which the fatigue aspects and items were presented. The questions posed are displayed in Table 2.

Table 1 Overview of fatigue dimensions and number of selected items

Dimension (original number of items)	Meaning of the dimension	Number of finally selected items
Severity (16)	Intensity of fatigue	5
Frequency (11)	Occurrence of fatigue	9
Duration (6)	Presence of fatigue over time	5
Changes in fatigue (7)	Variability and unpredictability of fatigue	9
Perceived causes of fatigue (17)	Circumstances that according to patients influence their fatigue	18
Energy (22)	Amount of energy, for example, to undertake an activity	18
Sleep/rest (19)	Patients' need for sleep or rest	14
Body feeling (26)	Feeling/manifestation of fatigue in the body	15
Cognition/concentration (18)	Impact of fatigue on the ability to concentrate or on tasks that require thinking	15
Coping (33)	Ways of dealing with fatigue, for example, used strategies or thoughts about fatigue	23
Negative emotions/mood (30)	Negative emotions or mood caused by fatigue	29
Consequences (89)	Impact of fatigue on daily life	85

Table 2 Structure of Delphi questionnaire

Participants were asked to…
1. Indicate whether they agree with a time frame of the last 7 days (we presented an example question and participants could choose between the answers ''agree with'' or ''not agree with'' and give reasons for their choice)
2. Judge the selected dimensions of fatigue for importance on a 4-point Likert scale: very important, rather important, less important, and not at all important
3. Indicate how appropriate they thought the selected items described each dimension on a 4-point Likert scale: very appropriate, rather appropriate, less appropriate, and not at all appropriate
4. Give comments and suggestions in the blank fields available under each question, e.g., give reasons for their evaluation of the items and their response format
5. Nominate other dimensions of fatigue and related items to be included in a comprehensive measure of fatigue

The questionnaire was piloted with one patient and one professional. Both rated the invitation e-mail, the questionnaire, and the instructions as clear and feasible. After each round, the data were analyzed and summarized using SPSS 18.

A main characteristic of the Delphi procedure is that after each round, participants are given feedback information about the results of the previous round. For the items that had to be re-evaluated in round 2, information about its evaluation in the previous round was provided by presenting the item from round 1, with a brief summary of the given comments, and then the adjusted item. Also, newly created items, based on the participants' comments and suggestions, were presented in round 2. Items had to be evaluated in the same way as in round 1.

The amount of time our respondents had to invest was rather large: completing the questionnaire of round 1 required about 1 h, whereas the questionnaire of round 2 still required about 45 min.

In the first round, 15 patients, 10 rheumatologists, and 20 nurses returned the completed questionnaire (N = 45); a response rate of 40.5%. Professionals worked at 21 hospitals spread over the Netherlands (see acknowledgment). In the second round, 80% of the participants of the first round (15 nurses, 13 patients, and 8 rheumatologists) participated (N = 36). In both rounds, our expert panel clearly consisted of more women than men, and the mean age was 51.4 (SD = 11.4) in the first round and 52.4 (SD = 10.5) in the second. For details of the samples, see Table 3.

Table 3 Gender and age of the participants in round 1 and round 2

Group	Round 1		Round 2	
	Gender (N) (female/male)	Mean age (in years)	Gender (N) (female/male)	Mean age (in years)
Patients	13/2	57.3 (SD = 11.4)	11/2	57.7 (SD = 10.3)
Rheumatologists	1/9	54.4 (SD = 10.8)	1/7	54.9 (SD = 12.1)
Nurses	19/1	45.4 (SD = 8.8)	14/1	46.5 (SD = 6.8)
Total	33/12	51.4 (SD = 11.4)	26/10	52.4 (SD = 10.5)

Analysis—criteria for item selection

In Delphi studies, different rules for defining sufficient consensus are applied,[31] ranging from 55 to 80%. To establish the rule for our study, we searched the literature for other Delphi studies in health research. Repeatedly, the criterion of 80% was used for selection of an item or topic, as for EULAR recommendations.[54] Based on these examples, we also applied the conservative criterion of 80% agreement. Further properties of our rule were developed in discussion among the authors to meet our special situation with different expert groups.

An item was directly selected for inclusion in the CAT item pool if 80% or more of the participants rated it as appropriate (response options ''very appropriate'' and ''rather appropriate''). An item rated as appropriate by 50% or less of the participants was directly excluded from the CAT item pool.

Besides the global percentage, we also calculated the percentages per expert group. If the mean percentage was high (80%) or low (50%) enough, but one group had a percentage of more than 10% different from the criterion (below 70% or above 60%), the item was not directly selected or removed. These and also those items rated as appropriate by 50–80% of the experts were discussed by the authors and adapted according to the participants' comments and suggestions and re-evaluated if necessary. If the same comment was raised by at least two participants, the item was adapted. The adapted items were judged again in round 2 by the panel. After round 2, comments were discussed among the researchers and the final decision about in- or exclusion of the items was made.

RESULTS

Time frame

Round 1

Precisely 80% of the experts agreed with a time frame of the last 7 days (patients 86.7%, rheumatologists 70%, and nurses 80%). Nevertheless, many comments were still given on the time frame. Some participants argued that a longer time frame would be more appropriate, because it would give a more reliable reflection of the experienced fatigue. They mentioned acute conditions like an infection or other physical reasons or complaints, weather, medication, stress, or mood that could affect fatigue. Using a longer time frame would be less sensitive to such circumstances. An argument for using a shorter time frame than 7 days was that it can be difficult to precisely remember the fatigue, so that people often give answers based on the last 2 or 3 days and how they feel today. Due to these comments, we decided to ask about the time frame again in the second round. We

showed a summary of the comments to the experts and asked whether they would prefer a time frame of the last 7 days or whether they would prefer a longer time frame of the last 14 days.

Round 2

The majority of participants (64%) preferred the shorter 7-day time frame. Although this percentage did not reach our 80% criterion, we chose to use this time frame in our instrument. Convincing arguments for this decision were that the changing character of fatigue and possible acute circumstances can also effect questions concerning the last 14 days and that the danger of recall bias increases with the length of the used time frame.

Items

Round 1

Step 1: An immediate decision could be made on 105 of the 294 items: 96 items were rated as appropriate by at least 80% of the participants so that they were immediately selected. Nine items were rejected because they were rated as appropriate by 50% or less. These rejected items are displayed in Table 4.

Table 4 Rejected items due to negative evaluation by experts in round 1

Item	Comments	Origin
Severity		
Over the past 7 days I felt bushed. 1 not at all / 2 a little / 3 moderately / 4 quite a bit 5 extremely	Too negative, dramatic, extreme or exaggerated. Item could be confusing because it could also be related to physical circumstances as loss of strength in the joints, or could have a psychological component such as depressive feelings.	POMS
How severe is the fatigue[1] which you have been experiencing during the past week? 1 – 10 / 1 mild, 10 severe	It is assumed that exhaustion is always present. It is not possible to measure exhaustion with the used scale since you can only be exhausted or not, and exhaustion is always serious.	MAF
During the past 7 days I was so tired that I could only sit and stare. -6 - +6 / -6 Not at all applicable, +6 Totally applicable	This item is difficult to grade, confusing, and too cursory. It contains two questions (about sit and about stare), and it could also measure depression and sounds very negative.	Interviews

Item	Comments	Origin
During the past 7 days I was so tired that even the tiniest movement seemed too strenuous. -6 - +6 / -6 Not at all applicable, +6 Totally applicable	The response options were rated as not appropriate, and the item as unclear, too dramatic and probably related to other aspects than fatigue.	Interviews
Energy		
During the past 7 days I felt efficient. 1 – 100 / 1 not at all efficient, 100 extremely efficient	This item was unclear for many participants. Furthermore it could also refer to depression.	Lee fatigue scale
Body feeling		
During the past 7 days I had a prickly feeling in my feet when I was tired. -6 - +6 / -6 Not at all applicable, +6 Totally applicable	This item has too many response options, could apply more to fibromyalgy patients, could be related to other diseases as polyneuropathy, lack of vitamine B12 etc.	Interviews
During the past 7 days I had a headache when I was tired. -6 - +6 / -6 Not at all applicable, +6 Totally applicable	This item has too many response options and does not refer to RA. A headache can have many causes, is more applicable to women than to men.	Interviews
During the last 7 days I kept yawning when I was tired. -6 - +6 / -6 Not at all applicable, +6 Totally applicable	This item has too many response options. It is doubtful what the relevance is as sleep disorders will not be assessed. Yawning is a symptom but all people do this.	Interviews
Negative emotions / mood		
Over the past 7 days, have you felt embarrassed because of fatigue? Not at all, a little, quite a bit, very much	Item is unclear and too difficult.	BRAF-MDQ

[1] in Dutch translated as „exhaustion“

Step 2: The remaining 189 items were discussed by the authors whereby the participants' comments and suggestions were taken into account. The authors decided to select 60 items (whereby one item was split into two) for inclusion in the item pool when the criticism did not concern the content of the item (e.g., a change in formulation such as replacement/deletion of a word/phrase) or when it merely referred to the response options or suggested placing an item under another dimension. For an example of such adaptations, see Table 5.

Table 5 Example for small adaptations in selection by authors

Item in round 1	Comments	Adapted and selected item
Response options		
During the past 7 days, I have fallen asleep without wanting to. - 6 to + 6 / -6 Not at all applicable + 6 Totally applicable	The response options are not clear, they are confusing.	During the past 7 days, I have fallen asleep without wanting to. 1 often, 2 sometimes, 3 rarely, 4 never
Formulation		
Over the past 7 days, have you lacked physical energy because of fatigue? Not at all, a little, quite a bit, very much	Unclear / difficult formulation, better to use "bodily" instead of "physical", with "bodily" the item would be clear.	Over the past 7 days, have you lacked bodily energy because of fatigue? Not at all, a little, quite a bit, very much

Having received comments with convincing arguments against their inclusion, the authors also decided to remove eight items from the item pool. One example was ''The fatigue I feel now is worse than the fatigue I felt before I had rheumatoid arthritis.'' Experts commented as follows: ''This is asking for something you already know. Patients experience the fatigue as different but that does not automatically mean that it is worse—it's more about quality than about severity. You ask patients to compare apples with oranges (fatigue vs. tiredness). Question is only applicable to patients with a short disease duration.''

Step 3: The remaining 121 items had to be re-evaluated in the second round (most of them in an adapted version). An example of such an adapted item is question 2 of the FACIT-F: ''During the past 7 days, I felt weak all over (0 not at all, 1 a little bit, 2 somewhat, 3 quite a bit, and 4 very much)'' (in the Dutch translation: helemaal niet, een beetje,

enigzins, in vrij hoge mate, in zeer hoge mate). Participants evaluated the response options as unclear and the item as vaguely formulated and unclear. Consequently, we changed the item to: ‘‘During the past 7 days, my fatigue made me feel weak (1 not at all, 2 a little, 3 rather, and 4 to a great extent)’’ (in Dutch: helemaal niet, een beetje, nogal, in sterke mate). Another example: ‘‘During the past 7 days, I wanted to do all kinds of nice things (-6 Not at all applicable and +6 Totally applicable).’’ Comments: The response options are unclear and ‘‘wanted to’’ does not say anything about its feasibility. Hence, we made the following adaptations: ‘‘During the past 7 days, I had enough energy to do nice things (1 always, 2 usually, 3 sometimes, 4 rarely, and 5 never).’’

Based on the comments regarding missing aspects, two new items were constructed and included in round 2. One item was split into two, bringing the number of items that had to be evaluated in round 2 to 124. Figure 1 gives an overview of the item selections in the two Delphi rounds (see the appendix at the end of this article).

Round 2

Step 1: As 58 of the 124 items were judged as appropriate by at least 80% of the participants in the second round, they were selected for the development of the CAT. Only one item had to be rejected due to 50% or less of the participants rating it as appropriate. This item was constructed based on our interview material: ‘‘During the past 7 days, I had an unpleasant feeling in my body when I was tired (-6 Not at all applicable and +6 Totally applicable).’’ Based on the comments given, we adjusted the item to: ‘‘During the past 7 days, my fatigue made me feel drowsy (1 not at all, 2 a little, 3 rather, and 4 to a great extent).’’ However, participants still rated this item as unclear.

Step 2: The remaining 65 items were discussed by the authors whereby the comments and suggestions of the participants were taken into account. The researchers decided to select 30 items, which could be adjusted according to the participant’s comments. So some items that were rated as appropriate by more than 50% of the participants but did not reach the 80% criterion were included in the item pool. These items were rated in a normative way by our expert panel. Some items reflecting strongly undesired or awesome aspects of fatigue were evaluated as not appropriate. Professionals indicated that they could be confronting for patients, and patients reported that fatigue should not have so much impact on life. Simply excluding those items would be problematic for the construction of a CAT as items measuring all levels of fatigue, also extreme ones (in methodological terms: difficult items), are needed for precise measurement. Therefore, we made use of a slightly modified Delphi approach to ensure the quality of our item pool for the CAT. In contrast, 35 items appeared to be insufficiently appropriate, even after the adaptations made after the first round. In cases where comments still indicated the same

problems as in the first round or where new points of concern emerged which could not be solved by a simple adaptation of the item (e.g., another formulation or using other response options), items were removed. So after the second round, all items could be selected or rejected, and no third round was necessary.
The classification of items into the dimensions was also satisfactory: only 6 of 245 finally selected items had to be placed under another dimension than presented in round 1.

DISCUSSION

This Delphi study resulted in a large item pool consisting of 245 items that were rated as adequate by patients and professionals to measure fatigue in RA. Moreover, it provided insights into problems regarding the measurement thereof.
Our first question was about the appropriate time frame to measure fatigue. The questionnaire items of the scales included in this study had different time frames and response options. We chose to use the original response options. For a CAT, this is no problem because as only one item is simultaneously displayed on the computer screen, patients will not be confused by different response options. In the CAT for depression,[10] also items with different response formats were included and even different recall periods were used. Although most patients did not express any concerns, it is not yet clear if and how the different formats influence patients' answers. Different response formats might increase the patients' concentration while completing the CAT. In contrast, a continually changing time frame would not be feasible.[33] So our aim was the decision for one time frame, and after round 2, we chose the time frame of the last 7 days. Lai et al.[55] compared self-reported fatigue in cancer patients using a time frame of the last 7 days and the last 4 weeks. At item level, some items provided more information when asking about the last 7 days, while others provided more information when asking about the last 4 weeks. Regarding information at scale level, the 7-day time frame provided slightly more information than the 4-week time frame. The authors concluded that both time frames are appropriate but recommend using the 7-day time frame as it is more informative.[55] This recommendation is in accordance with the opinion of the majority of participants in the current study. Nevertheless, our decision remains disputable. In the first round, 80% of the participants agreed with a time frame of the last 7 days when asked to choose between agreement and non-agreement. When given the choice between 7 and 14 days in round 2, only 64% preferred the 7-day time frame. This is no huge majority, and it reflects the experts' perspective that no ideal solution exists for the question which time frame should be used.

Furthermore, we wanted to investigate which items are appropriate for measuring fatigue in RA. A notable number of the rejected items came from the interview material. This suggests that it is not sufficient to transform patients' stories directly into items when measuring fatigue. Items based on patients' stories have to undergo further expert checks to find the right formulation. Before starting to develop a scale and its related statistical analyses, further evaluation by experts is recommended. We also found that only 40% of the items from traditional questionnaires (SF-36, POMS, FACIT-F, MAF) were clear and adequate in their original style for at least 80% of the participants.[37] This finding further underlines the relevance of the perspective of professionals and patients in the development of measurement instruments for fatigue in RA. Generally, it appeared that items should be formulated as clearly and briefly as possible. The same applies to the response options: clearly distinguishable and not too many.

Some experts commented on the importance of careful use of different terms for fatigue, with the Dutch frequently using the words ''moeheid'' and ''vermoeidheid''. Although comparable, the difference between these words is not exactly the same as between ''tiredness'' and ''fatigue'' in the English language. Experts indicated that they associate the word ''vermoeidheid'' with being fatigued for a reason and with being more seriously fatigued than with the word ''moeheid''. Our item pool also contained several items referring to ''moe'' or ''moeheid'' instead of ''vermoeidheid'' in an endeavor to measure the entire range of fatigue and thereby also less severe fatigue.

Remarkable is that most of the experts in this study were female. This reflects the gender distribution in RA[13] and the fact that most nurses are still women.[56] Nevertheless, it is important not to neglect the perspective of men in future studies.

It is beyond question that the perspective of experts is essential for the development of questionnaire items. Nevertheless, it appeared that a Delphi approach has its limitations. The experts' comments showed that patients and professionals tended to evaluate the fatigue items in a normative way. Items reflecting a rather annoying impact of fatigue were sometimes judged as inadequate with the explanation that these could be confronting for patients (professionals' perspective) or that they do not want the fatigue to have so much impact (patients' perspective). The idea that undesired aspects should be omitted is especially problematic when constructing a CAT. For a precise measurement, the item pool also requires items referring to an extreme impact of fatigue, meaning that items are also needed that might have a moral implication for patients and professionals. Therefore, it was important to discuss items among the authors and to scrutinize the comments given when deciding on in- or exclusion of an item. Furthermore, comments from patients showed that some of them had difficulties distinguishing the value of an item for measuring fatigue in general from their own experience and with their potential

answer. Patients may have rated some items as inappropriate because they were not applicable to their own fatigue. Because of this, we chose to include some items that did not reach our 80% criterion, meaning a slight deviation from the conservative decision rule.

For participants, an important disadvantage of this Delphi approach was the length of time it took. Therefore, we think that our response rate of 40.5% (111 invitations and 45 participants in round 1) is quite satisfactory. As fewer comments were given in the last part of the questionnaire, we compensated for this by reversing the order of dimensions/items in the second round. Also, for the researchers, the study was time-consuming due to the large number of items that had to be prepared and presented in the online questionnaire, and the huge amount of data that had to be analyzed in two rounds.

Although the Delphi procedure has some bottlenecks, it is a valuable method to include the opinions of experts in a thorough development of dimensions and items for a measurement instrument. We assume that our item pool is comprehensive and adequate for the measurement of fatigue in RA. It will be used in our next study in which the items will be calibrated according to IRT.

ACKNOWLEDGEMENTS

We would like to thank the participants in this study; patients from the forum of Patient Research Partners of the Arthritis Centre Twente; and rheumatologists and nurses from Universitair Medisch Centrum Groningen, Medisch Centrum Leeuwarden, Ziekenhuisgroep Twente, Zorg Groep Twente, Medisch Spectrum Twente, St. Maartenskliniek Nijmegen, TweeSteden Ziekenhuis Locatie Tilburg, Universitair Medisch Centrum Utrecht, Rijnstate Ziekenhuis Arnhem, Meander Medisch Centrum Amersfoort, Ziekenhuis Gelderse Vallei Ede, Medisch Centrum Haaglanden s-Gravenhage, Tweesteden Ziekenhuis Tilburg, Evean Thuiszorg Purmerend, Streekziekenhuis Midden Twente Hengelo, Isala Klinieken Zwolle, Albert Schweitzer ziekenhuis Dordrecht, Spaarne Ziekenhuis Hoofddorp, Scheperziekenhuis Emmen, Thuiszorg West- Brabant Roosendal, Omring thuiszorg Hoorn.

REFERENCES

1. Rose M, Bezjak A. Logistics of collecting patient-reported outcomes (PROs) in clinical practice: an overview and practical examples. Qual Life Res 2009;18:125–136.
2. Hambleton RK, Swaminathan H, Rogers HJ. Fundamentals of Item response theory. Thousand Oaks: Sage, 1991.
3. Walter OB. Adaptive tests for measuring anxiety and depression. In: WJ van der Linden, CAW Glas eds. Elements of adaptive testing. New York: Springer, 2010:123-136.
4. Wouters H, Zwinderman AH, van Gool WA, Schmand B, Lindeboom R. Adaptive cognitive testing in dementia. Int J Methods Psychiatr Res 2009;18:118–127.
5. Lai JS, Cella D, Chang CH, Bode RK, Heinemann AW. Item banking to improve, shorten and computerize self-reported fatigue: An illustration of steps to create a core item bank from the FACIT-fatigue scale. Qual Life Res 2003;12:485–501.
6. Lai JS, Cella D, Dineen K, Bode R, von Roenn J, Gershon RC, et al. An item bank was created to improve the measurement of cancer-related fatigue. J Clin Epidemiol 2005;58:190–197.
7. Cella D, Yount S, Rothrock N, Gershon R, Cook K, Reeve B, et al. The patient-reported outcomes measurement information system (PROMIS): Progress of an NIH Roadmap Cooperative Group during its first two years. Med Care 2007;45:I3–I11.
8. Fliege H, Becker J, Walter OB, Bjorner JB, Klapp BF, Rose M. Development of a computer-adaptive test for depression (D-CAT). Qual Life Res 2005;14:2277–2291.
9. Walter OB, Becker J, Bjorner JB, Fliege F, Klapp BF, Rose M. Development and evaluation of a computer adaptive test for ‘‘anxiety’’ (Anxiety-CAT). Qual Life Res 2007;16:143–155.
10. Fliege H, Becker J, Walter OB, Rose M, Bjorner JB, Klapp BF. Evaluation of a computer-adaptive test for the assessment of depression (D-CAT) in clinical application. Int J Methods Psychiatr Res 2009;18:23–36.
11. Becker J, Fliege H, Kocalevent RD, Bjorner JB, Rose M, Walter OB, et al. Functioning and validity of a computerized adaptive test to measure anxiety (A-CAT). Depress Anxiety 2008;25:182–194.
12. Kocalevent RD, Rose M, Becker J, Walter OB, Fliege H, Bjorner JB, et al. An evaluation of patient-reported outcomes found computerized adaptive testing was efficient in assessing stress perception. J Clin Epidemiol 2009;62:278–287.
13. Kvien TK, Scherer HU, Burmester GH. Rheumatoid arthritis. In: JWJ Bijlsma ed. EULAR compendium on rheumatic diseases. London: BMJ Publishing Group and European League Against Rheumatism, 2009:61-80.

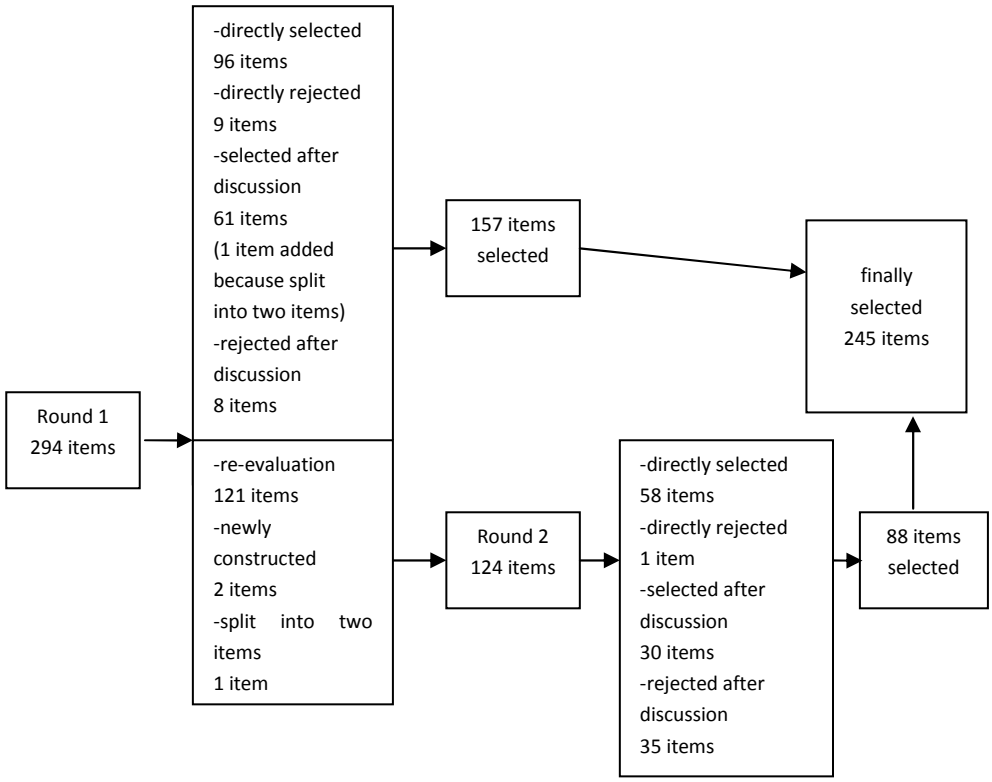
14. Wolfe F, Hawley DJ, Wilson K. The prevalence and meaning of fatigue in rheumatic disease. J Rheumatol 1996;23:1407–1417.
15. Belza BL, Henke CJ, Yelin EH, Epstein WV, Gilliss CL. Correlates of fatigue in older adults with rheumatoid arthritis. Nurs Res 1993;42:93–99.
16. Hewlett S, Carr M, Ryan S, Kirwan J, Richards P, Carr A, et al. Outcomes generated by patients with rheumatoid arthritis: how important are they? Musculoskeletal Care 2005;3:131–142.
17. Kirwan JR, Newman S, Tugwell PS, Wells GA. Patient perspective on outcomes in rheumatology—a position paper for OMERACT 9. J Rheumatol 2009;36:2067–2070.
18. Kirwan JR, Hewlett S. Patient perspective: reasons and methods for measuring fatigue in rheumatoid arthritis. J Rheumatol 2007;34:1171–1173.
19. Kirwan JR, Minnock P, Adebajo A, Bresnihan B, Choy E, de Wit M, et al. Patient perspective: fatigue as a recommended patient centered outcome measure in rheumatoid arthritis. J Rheumatol 2007;34:1174–1177.
20. Minnock P, Kirwan J, Bresnihan B. Fatigue is a reliable, sensitive and unique outcome measure in rheumatoid arthritis. Rheumatology 2009;48:1533–1536.
21. Belza Tack B. Fatigue in rheumatoid arthritis. Conditions, strategies, and consequences. Arthritis Care Res 1990;3:65–70.
22. Hewlett S, Cockshott Z, Byron M, Kitchen K, Tipler S, Pope D, et al. Patients' perceptions of fatigue in rheumatoid arthritis: Overwhelming, uncontrollable, ignored. Arthritis Care Res 2005;53:697–702.
23. Repping-Wuts H, Uitterhoeve R, van Riel P, van Achterberg T. Fatigue as experienced by patients with rheumatoid arthritis (RA): A qualitative study. Int J Nurs Stud 2008;45:995–1002.
24. Nikolaus S, Bode C, Taal E, van de Laar MAFJ. New insights into the experience of fatigue among patients with rheumatoid arthritis: A qualitative study. Ann Rheum Dis 2010;69:895–897.
25. Hewlett S, Hehir M, Kirwan JR. Measuring fatigue in rheumatoid arthritis: A systematic review of scales in use. Arthritis Rheum 2007;57:429–439.
26. Ware JE Jr., Sherbourne CD. The MOS 36-item short-form health survey (SF-36). I. Conceptual framework and item selection. Med Care 1992;30:473–483.
27. Cella D, Yount S, Sorensen M, Chartash E, Sengupta N, Grober J. Validation of the Functional Assessment of Chronic Illness Therapy Fatigue Scale relative to other instrumentation in patients with rheumatoid arthritis. J Rheumatol 2005;32:811–819.
28. McNair D, Lorr M, Droppelman L. Profile of mood states manual. New York: Multi-health Systems Inc; 1992.

29. Belza Tack B. Dimensions and correlates of fatigue in older adults with rheumatoid arthritis (dissertation). San Francisco: University of California; 1991.
30. Nicklin J, Cramp F, Kirwan J, Greenwood R, Urban M, Hewlett S. Measuring fatigue in rheumatoid arthritis: A cross-sectional study to evaluate the Bristol Rheumatoid Arthritis Fatigue Multi-Dimensional questionnaire, visual analogue scales, and numerical rating scales. Arthritis Care Res 2010;62:1559–1568.
31. Powell C. The Delphi technique: Myths and realities. J Adv Nurs 2003; 41:376–382.
32. Streiner DL, Norman GR. Health measurement scales—a practical guide to their development and use. New York: Oxford University Press; 2003.
33. Fayers PM, Machin D. Quality of life—assessment, analysis and interpretation. Chichester: Wiley; 2000.
34. Yorkston KM, Johnson K, Boesflug E, Skala J, Amtmann D. Communication about the experience of pain and fatigue in disability. Qual Life Res 2010;19:243–251.
35. De Wit M. From orphan outcome to darling pet—patient participation at OMERACT and the emergence of fatigue as an important outcome for research. Ann Rheum Dis 2010;69(Suppl3):23.
36. Nikolaus S, Bode C, Taal E, van de Laar MAFJ. Which dimensions of fatigue should be measured in patients with rheumatoid arthritis?—a Delphi study. Musculoskeletal Care. Published online 11 November 2011. doi:10.1002/msc.222.
37. Nikolaus S, Bode C, Taal E, van de Laar MAFJ (in press). Experts' evaluations of fatigue questionnaires used in rheumatoid arthritis—a Delphi study among patients, nurses and rheumatologists in the Netherlands. Clin Exp Rheumatol.
38. Jones J, Hunter D. Qualitative research: consensus methods for medical and health services research. BMJ 1995;311:376–380.
39. Van der Zee KI, Sanderman R. Het meten van de algemene gezondheidstoestand met de RAND-36—een handleiding. Rijksuniversiteit Groningen: Noordelijk Centrum voor Gezondheidsvraagstukken; 1993.
40. FACIT Functional Assessment of Chronic Illness Therapy. (2009). http://www.facit.org. Accessed 16 July 2009.
41. Cluydts RJG. Gemoedstoestanden en slaap—een experimenteel onderzoek [dissertation]. Brussel: Vrije Universiteit Brussel; 1979.
42. MAPI INSTITUTE. (2009). http://www.mapi-institute.com. Accessed 31 July 2009.
43. Guillemin F, Bombardier C, Beaton D. Cross-cultural adaptation of health-related quality of life measures: Literature review and proposed guidelines. J Clin Epidemiol 1993;46:1417–1432.

44. Nikolaus S, Bode C, Taal E, van de Laar MAFJ. Four different patterns of fatigue in rheumatoid arthritis patients: Results of a Q-sort study. Rheumatology 2010;49:2191–2199.
45. Hjollund NH, Andersen JH, Bech P. Assessment of fatigue in chronic disease: A bibliographic study of fatigue measurement scales. Health Qual Life Outcomes 2007;5:art. no. 12.
46. Lwin CTT, Bishay M, Platts RG, Booth DA, Bowman SJ. The assessment of fatigue in primary Sjögren´s Syndrome. Scand J Rheumatol 2003;32:33–37.
47. Neuberger GB. Measures of fatigue. Arthritis Rheum 2003;49:S175–S183.
48. Dittner AJ, Wessely SC, Brown RG. The assessment of fatigue—a practical guide for clinicians and researchers. J Psychosom Res 2004;56:157–170.
49. Ad hoc committee on systemic lupus erythematosus response criteria for fatigue. Measurement of fatigue in systemic lupus erythematosus: A systematic review. Arthritis Rheum 2007;57:1348–1357.
50. Minton O, Stone P. A systematic review of the scales used for the measurement of cancer-related fatigue (CRF). Ann Oncol 2009;20:17–25.
51. Whithead L. The measurement of fatigue in chronic illness: A systematic review of unidimensional and multidimensional fatigue measures. J Pain Symptom Manage 2009;3:107–128.
52. Patient-Reported Outcomes Measurement Information System (PROMIS). (2009). Preliminary item pools—fatigue (fatigue-experience and fatigue-impact). http://www.nihpromis.org/ Web%20Pages/Preliminary%20Item%20Banks%20[2006]aspx. Accessed 13 November 2009.
53. SurveyMonkey [http://www.surveymonkey.com].
54. Avouac J, Kowal-Bielecka O, Landewe R, Chwiesko S, Miniati I, Czirjak L, et al. European league against rheumatism (EULAR) scleroderma trial and research group (EUSTAR) recommendations for the treatment of systemic sclerosis: Methods of elaboration and results of systematic literature research. Ann Rheum Dis 2009;68:629–634.
55. Lai JS, Cook K, Stone A, Beaumont J, Cella D. Classical test theory and item response theory/Rasch model to assess differences between patient-reported fatigue using 7-day and 4-week recall periods. J Clin Epidemiol 2009;62:991–997.
56. Statistics Netherlands. (2010). http://www.cbs.nl/nl-NL/menu/ themas/arbeid-sociale-zekerheid/publicaties/artikelen/archief/ 2005/2005-1825-wm.htm. Accessed 13 October 2010.

APPENDIX:

Figure 1 Item selection in the two Delphi rounds



Chapter 7

Which Dimensions of Fatigue Should be Measured in Patients with Rheumatoid Arthritis? A Delphi Study

S. Nikolaus
C. Bode
E. Taal
M.A.F.J. van de Laar

Musculoskeletal Care. Published online 11 November 2011. doi:10.1002/msc.222

ABSTRACT

Objective Rheumatoid arthritis (RA) patients experience fatigue as a multidimensional symptom. The aim of the present study was to use health professionals and patients alike to identify which dimensions of fatigue should be measured in RA.

Methods Twelve fatigue dimensions were constructed, based on items from traditional questionnaires and items generated from interviews. Health professionals and patients evaluated these dimensions, related to an initial pool of 294 items, in a Delphi procedure. Dimensions were selected if rated important by at least 80% of the participants.

Results Ten rheumatologists, 20 nurses and 15 patients participated. All fatigue dimensions were selected directly (severity, frequency, duration, changes in fatigue, perceived causes of fatigue, energy, sleep/rest, body feeling, cognition/concentration, coping, negative emotions/mood and consequences). No additional dimensions emerged from participants' comments and suggestions.

Conclusions This study revealed 12 fatigue dimensions. This underlines the multidimensionality of fatigue in RA and the need for comprehensive measurement.

INTRODUCTION

Measuring fatigue in rheumatoid arthritis (RA) provides additional information on disease outcomes which does not overlap with other established measures of RA, and is essential for understanding the patient's perspective.[1] Several uni- and multidimensional scales have been developed to assess fatigue in clinical practice and research. Whereas unidimensional questionnaires are usually brief and provide a single score, multidimensional scales comprise a larger number of items and provide more detailed information, giving insight into different profiles and underlying mechanisms of fatigue.[2]

Fatigue measures with single-item scales, such as visual analogue scales (VAS) or numerical rating scales (NRS), might have some value,[1] but do not correspond with the multidimensional character of fatigue as reported by patients. Qualitative studies have shown that patients experience fatigue as a multidimensional, annoying symptom with far-reaching consequences.[3-6] In line with patients' experiences, it is expected that measurement must also be multidimensional, although it is not yet clear which dimensions should be assessed.

Of the four multi-item fatigue questionnaires for which there is reasonable evidence for validity in RA,[7] only the Multidimensional Assessment of Fatigue scale (MAF)[8] comprises several dimensions: severity, distress, timing and interference. The Functional Assessment of Chronic Illness Therapy Fatigue Scale (FACIT-F)[9] provides separate scores for the experience of fatigue as a symptom and the impact of fatigue on daily functioning, although it is usually applied unidimensionally. The Short Form 36 (SF-36) subscale vitality[10] and the Profile of Mood States (POMS) subscale fatigue/ inertia[11] have only one dimension. Nicklin et al.[12] demonstrated that none of these four scales covers patient-reported concepts of fatigue comprehensively: frequency, duration, energy, sleep, cognition, coping, emotion, impact, social life, planning, relationships and quality of life.[4,12] All of these dimensions were also relevant in our study on fatigue experiences.[13] Recently, the Bristol RA Fatigue Multi-Dimensional Questionnaire (BRAF-MDQ)[14] was developed from the patient's perspective and evaluated in a British RA population. With its four dimensions (physical, living, cognition, emotion), a score for each dimension can be calculated. Its construction and evaluation focused on patients scoring 5 or higher, and 7 or higher on a VAS, reflecting severe fatigue.[4,14]

There is still no measurement instrument available that fully embraces the diversity of the experience of fatigue in RA with its several dimensions. Developing a more extensive and informative instrument in traditional format would imply a higher burden for patients, as they would have to fill in more items, requiring more time and energy.[15] A solution lies in computer-adaptive testing (CAT), which provides the possibility of measuring fatigue

comprehensively with relatively few items selected from an item bank.[16] To explore and select the relevant dimensions for the development of a multidimensional CAT for fatigue in RA, we conducted a Delphi study, to allow patients, rheumatologists and nurses to evaluate current dimensions of fatigue.

PATIENTS AND METHODS

Delphi process

By using the Delphi process,[17] opinions about a certain topic can be collected via a questionnaire, which is sent by (electronic) mail to a panel of potential participants, who then answer it individually. A Delphi study consists of at least two rounds and aims to reach consensus among the participants.

Preparation of the dimensions

For the development of a multidimensional CAT for fatigue in RA, a comprehensive pool of dimensions of the fatigue experience in RA is needed. We started by collating dimensions of the validated Dutch versions of the multi-item questionnaires with reasonable evidence for validity in RA:[7] RAND SF-36,[18] FACIT-F,[19] POMS[20] and MAF.[21] The dimensions of the BRAF-MDQ were also considered. All dimensions, with the items reflecting each dimension, were sorted into a table (e.g. frequency, duration, severity, consequences, etc.) which was subsequently supplemented with additional dimensions and items formulated in the context of our previous Q-sort study[13] and based on interview material on the experience of fatigue in patients with RA.[5] To check for possibly missing aspects, we also looked at fatigue questionnaires which have not been validated in RA, but no additional dimensions of fatigue emerged from this search. This procedure resulted in a table with 12 fatigue dimensions (severity, frequency, duration, changes in fatigue, perceived causes of fatigue, energy, sleep/rest, body feeling, cognition/concentration, coping, negative emotions/ mood and consequences) and their associated items. The dimension with the largest number of items was ‘consequences’, capturing general consequences and those regarding work and daily tasks, household tasks, social life, self-care, physical activity, leisure time and planning. The preparation of the item pool is described in detail elsewhere.[22]

Participants and data collection

Experts (40 rheumatologists, 40 nurses and 31 patients) were informed about the study and asked to participate by email, with a link to an online questionnaire. Email addresses

of patients were collated from the database of Patient Research Partners of the Arthritis Centre Twente, and those of rheumatologists and nurses from the member lists of Dutch professional associations in Rheumatology (NVR, Dutch Association for Rheumatology) and the DREAM registry (Dutch Rheumatoid Arthritis Monitoring, a collaboration between hospitals in the Netherlands to improve the quality of care for patients with RA). Brief descriptions of the dimensions, with corresponding items, were provided. Participants were asked to indicate the importance of each dimension for the measurement of fatigue in RA (four-point Likert scale: very important, rather important, less important, not at all important). Moreover, they had the possibility of adding comments and suggestions in the blank fields available for each dimension, to underpin their evaluation or even add new dimensions.

Analysis: Criteria for selection of dimensions

A dimension of fatigue was selected for inclusion in the CAT if 80% or more of the participants rated it as important (response options 'very important' and 'rather important'). A dimension which had been rated as important by 50% or less of participants was excluded. Dimensions rated as important by 50–80% of the experts were first discussed by the authors and then adapted according to participants' comments and suggestions, and re-evaluated if necessary. If the same comment was raised by at least two participants, it led to an adaptation of the dimension. The dimensions of fatigue included in the Delphi are shown in Table 1.

Table 1 Overview of fatigue dimensions

Dimension	Meaning of the dimension	Sample item
Severity	Intensity of fatigue	Please circle the number that shows your average level of fatigue during the past 7 days.
Frequency	Occurrence of fatigue	Did you feel tired?
Duration	Presence of fatigue over time	How long, on average, has each episode of fatigue lasted during the last 7 days?
Changes in fatigue	Variability and unpredictability of fatigue	To what degree has your fatigue changed during the past week?
Perceived causes of fatigue	Circumstances that, according to patients, influence their fatigue	Exercise brings on my fatigue.
Energy	Amount of energy, for example to undertake an activity	Did you have a lot of energy?
Sleep/rest	Patients´ need for sleep or rest	I need to sleep during the day.
Body feeling	Feeling/manifestation of fatigue in the body	I feel relaxed.
Cognition/ concentration	Impact of fatigue on the ability to concentrate, or on tasks that require thinking	Have you forgotten things because of fatigue?
Coping	Ways of dealing with fatigue, for example used strategies, or thoughts about fatigue	I found it difficult to accept the fatigue.
Negative emotions/ moode	Negative emotions or mood caused by fatigue	I am frustrated by being too tired to do the things I want to do.
Consequences	Impact of fatigue on daily life	I have trouble finishing things because I am tired.

RESULTS

In the first round, 15 patients, 10 rheumatologists and 20 nurses, working at 21 hospitals spread across the Netherlands, returned the completed questionnaire (n = 45).
In the second round, 80% of the respondents from the first round (15 nurses, 13 patients and eight rheumatologists) participated (n = 36). In both rounds, our expert panel clearly consisted of more women than men (2:1), and the mean age was 51.4 (standard deviation [SD] = 11.4) in the first round and 52.4 (SD = 10.5) in the second.
In round 1, all 12 dimensions were evaluated as important by more than 80% of the participants, meaning that they could be selected for the development of the CAT. The dimensions were regarded as important by at least 70% of each expert group (rheumatologists, nurses, patients), as displayed in Table 2. Therefore, it was not necessary to ask about the dimensions again in round 2. As it is the principle of a Delphi study to give feedback about the results of the first round to participants, we presented the dimensions again in round 2, but only asked for re-evaluation regarding the items.

Table 2 Percentages of participants who evaluated a fatigue dimension as important

Dimension	global	patients	nurses	rheumatologists
Severity	91%	93%	90%	90%
Frequency	91%	93%	90%	90%
Duration	93%	87%	95%	100%
Changes in fatigue	89%	87%	90%	90%
Perceived causes of fatigue	91%	87%	100%	80%
Energy	96%	93%	95%	100%
Sleep/rest	93%	93%	100%	80%
Body feeling	82%	93%	80%	70%
Cognition/concentration	89%	87%	95%	80%
Coping	96%	100%	100%	80%
Negative emotions/mood	89%	74%	95%	100%
Consequences	96%	87%	100%	100%

DISCUSSION

All dimensions were rated as important for the measurement of fatigue in RA by at least 80% of our expert panel (rheumatologists, nurses and patients) and could be selected immediately in the first round. There was a large consensus between the three expert groups. The classification of the dimensions turned out to be valid. This was checked because participants were not only asked how important they thought a dimension was, and we also wanted to assess whether each item matched a dimension, as described in our article on item selection for an item bank/CAT for fatigue in RA.[22] Furthermore, the experts were able to comment on each dimension and suggest further dimensions. Only six items had to be placed under another dimension, and no additional dimensions were suggested. These results pointed to a good and comprehensive selection and naming of the dimensions, which were also reflected adequately by the items in our item pool. Hence, the 12 content-valid dimensions will be used for the further development of our multidimensional item bank/CAT. Such a comprehensive measurement becomes feasible owing to the advantage of CAT, which requires fewer items than in a fixed-length questionnaire.[16]

In the next study, factor analyses will further validate the proposed dimensions. The present study showed the importance of including every significant aspect of fatigue in the development of the CAT. However, naming dimensions and classifying items without statistical analyses remains disputable. In the evaluation study of the BRAF-MDQ, the 12 patient-reported categories ended up in four dimensions after factor analysis.[14] Pollard et al.[23] conducted a factor analysis of different fatigue instruments and found five dominant factors: psychological factors anxiety/depression, distress/cognition, fatigue severity, physical interference and social interference. They concluded that currently available fatigue questionnaires measure different aspects of fatigue and highlighted the need for a multidimensional tool comprising questions that can appropriately measure the impact of fatigue and its possible drivers, besides its overall severity.[23] Our validated fatigue dimensions are comprehensive and adequate for the development of such a multidimensional instrument in the form of a CAT.

ACKNOWLEDGEMENTS

We would like to thank the following participants in this study: patients from the forum of patient research partners of the Arthritis Centre Twente and rheumatologists and nurses from: Universitair Medisch Centrum Groningen, Medisch Centrum Leeuwarden, Ziekenhuisgroep Twente, Zorg Groep Twente, Medisch Spectrum Twente, St. Maartenskliniek Nijmegen, TweeSteden Ziekenhuis Locatie Tilburg, Universitair Medisch Centrum Utrecht, Rijnstate Ziekenhuis Arnhem, Meander Medisch Centrum Amersfoort, Ziekenhuis Gelderse Vallei Ede, Medisch Centrum Haaglanden s-Gravenhage, Tweesteden Ziekenhuis Tilburg, Evean Thuiszorg Purmerend, Streekziekenhuis Midden Twente Hengelo, Isala Klinieken Zwolle, Albert Schweitzer Ziekenhuis Dordrecht, Spaarne Ziekenhuis Hoofddorp, Scheperziekenhuis Emmen, Thuiszorg West-Brabant Roosendal and Omring Thuiszorg Hoorn.

REFERENCES

1. Minnock P, Kirwan J, Bresnihan B. Fatigue is a reliable, sensitive and unique outcome measure in rheumatoid arthritis. Rheumatology 2009;48:1533–6.
2. Dittner AJ, Wessely SC, Brown RG. The assessment of fatigue – A practical guide for clinicians and researchers. J Psychosom Res 2004;56:157–70.
3. Belza Tack B. Fatigue in rheumatoid arthritis. Conditions, strategies, and consequences. Arthritis Care Res 1990;3:65–70.
4. Hewlett S, Cockshott Z, Byron M, Kitchen K, Tipler S, Pope D, Hehir M. Patients' perceptions of fatigue in rheumatoid arthritis: Overwhelming, uncontrollable, ignored. Arthritis Care Res 2005;53:697–702.
5. Nikolaus S, Bode C, Taal E, van de Laar MAFJ. New insights into the experience of fatigue among patients with rheumatoid arthritis: A qualitative study. Ann Rheum Dis 2010;69:895–7.
6. Repping-Wuts H, Uitterhoeve R, van Riel P, van Achterberg T. Fatigue as experienced by patients with rheumatoid arthritis (RA): A qualitative study. Int J Nurs Stud 2008;45:995–1002.
7. Hewlett S, Hehir M, Kirwan JR. Measuring fatigue in rheumatoid arthritis: A systematic review of scales in use. Arthritis Rheum 2007;57:429–39.
8. Belza Tack B. Dimensions and correlates of fatigue in older adults with rheumatoid arthritis. Dissertation. University of California; 1991.
9. Cella D, Lai JS, Stone A. Self-reported fatigue: one dimension or more? Lessons from the Functional Assessment of Chronic Illness Therapy Fatigue (FACIT-F) questionnaire. Support Care Cancer 2011;19:1441–50.
10. Ware JE Jr, Sherbourne CD. The MOS 36-Item Short-Form Health Survey (SF-36). I. Conceptual framework and item selection. Med Care 1992;30:473–83.
11. McNair D, Lorr M, Droppelman L. Profile of Mood States Manual. New York, NY: Multi-health Systems Inc; 1992.
12. Nicklin J, Cramp F, Kirwan J, Urban M, Hewlett S. Collaboration with patients in the design of patient reported outcome measures: Capturing the experience of fatigue in rheumatoid arthritis. Arthritis Care Res 2010;62:1552–8.
13. Nikolaus S, Bode C, Taal E, van de Laar MAFJ. Four different patterns of fatigue in rheumatoid arthritis patients: Results of a Q-sort study. Rheumatology 2010;49:2191–9.
14. Nicklin J, Cramp F, Kirwan J, Greenwood R, Urban M, Hewlett S. Measuring fatigue in rheumatoid arthritis: A cross-sectional study to evaluate the Bristol Rheumatoid Arthritis Fatigue Multi-Dimensional questionnaire, visual analogue scales, and numerical rating scales. Arthritis Care Res 2010;62:1559–68.

15. Whitehead L. The measurement of fatigue in chronic illness: A systematic review of unidimensional and multidimensional fatigue measures. J Pain Symptom Manage 2009;37:107–22.
16. Rose M, Bezjak A. Logistics of collecting patientreported outcomes (PROs) in clinical practice: An overview and practical examples. Qual Life Res 2009;18:125–36.
17. Jones J, Hunter D. Qualitative research: Consensus methods for medical and health services research. BMJ 1995;311:376–80.
18. van der Zee KI, Sanderman R. Het meten van de algemene gezondheidstoestand met de RAND-36 – een handleiding. Groningen in the Netherlands: Noordelijk Centrum voor Gezondheidsvraagstukken, Rijksuniversiteit Groningen; 1993.
19. FACIT :Functional Assessment of Chronic Illness Therapy. Available at http://www.facit.org [Accessed 16 July 2009].
20. Cluydts RJG. Gemoedstoestanden en slaap – een experimenteel onderzoek. Dissertation. Vrije Universiteit Brussel; 1979.
21. MAPI Institute. Available at http://www.mapi-institute. com [Accessed 31 July 2009].
22. Nikolaus S, Bode C, Taal E, van de Laar MAFJ. Selection of items for a computer adaptive test to measure fatigue in patients with rheumatoid arthritis – A Delphi approach. Qual Life Res 2011 [Epub ahead of print], DOI: 10.1007/s11136-011-9982-8
23. Pollard LC, Scott DL, Donaldson N, Choy EH. Fatigue in rheumatoid arthritis, what are we measuring? A factor analysis of fatigue instruments. Rheumatology 2009;48:i135.

Chapter 8

Experts' evaluations of fatigue questionnaires used in rheumatoid arthritis - A Delphi study among patients, nurses and rheumatologists in the Netherlands

S. Nikolaus
C. Bode
E. Taal
M.A.F.J. van de Laar

Clin Exp Rheumatol (in press)

ABSTRACT

Objective Evaluating fatigue items from traditional questionnaires and a new scale (BRAF-MDQ) by experts in rheumatoid arthritis (RA). This evaluation was part of a study to select fatigue items to develop an item bank for a Dutch computer-adaptive test (CAT) for RA. Experts' opinions were incorporated since they are essential for content validity of measurement instruments.

Methods The 60 items of the SF-36 subscale vitality, FACIT-F, POMS subscale fatigue/inertia, MAF and the recently developed BRAF-MDQ were evaluated by rheumatologists, nurses and RA patients in a Delphi procedure. Items were selected for development of the item bank/CAT if rated as adequate by at least 80% of the participants (when 50% or less they were excluded). On the basis of participants' comments, remaining items were re-worded and re-evaluated in the following round. The procedure stopped when all items were selected or rejected.

Results Ten rheumatologists, 20 nurses and 15 RA patients participated. After the first round, 40% of the traditional items and 60% of the BRAF-MDQ items were directly selected and 3 items of the traditional questionnaires and 1 item of the BRAF-MDQ were directly excluded. Remaining items were re-worded, eight of which were presented for re-evaluation in the second round. Finally, 90% of the items from the traditional questionnaires and 95% of the items from the new BRAF-MDQ were included in our item pool.

Conclusion Fifty-five of the 60 items (92%) from fatigue questionnaires proved to have good content validity and were feasible for use in the Netherlands, some after adaptation.

Measuring fatigue in rheumatoid arthritis (RA) provides information on a prominent symptom of RA, and is essential for the understanding of the patients´ perspective.[1] Patients experience fatigue as a multidimensional, annoying symptom with far-reaching consequences.[2-5]

Only four multi-item fatigue questionnaires have shown reasonable evidence for validity in RA[6]: Short Form 36 subscale vitality (SF-36),[7] Functional Assessment of Chronic Illness Therapy Fatigue Scale (FACIT-F),[8] Profile of Mood States subscale fatigue/inertia (POMS),[9] and the Multidimensional Assessment of Fatigue scale (MAF).[10] However, none of these scales met all criteria for validity nor was the perspective of patients included in their development.[6] The Bristol RA Fatigue Multi-Dimensional Questionnaire (BRAF-MDQ)[11] was recently constructed and evaluated in a British RA population. Its development encompasses the patients' perspective with interviews, focus groups and cognitive testing. The development focuses on patients scoring 7 or higher on a VAS fatigue and the evaluation of the questionnaire on patients scoring 5 or higher.[2,11]

There is still no measurement instrument that incorporates the patients' perspective and that is validated in the Netherlands. In this study, the BRAF-MDQ items will be tested in a Dutch population for the first time. However, we do not intend to develop another fatigue questionnaire; instead we will use innovative technology to measure fatigue in RA more precisely with fewer items. Computer-adaptive testing (CAT) allows comprehensively measuring fatigue with relatively few items.[12] Items are respectively selected and based on the patient's previous answer, thus enabling precise measurement at individual level with few items. The working mechanism of a CAT can be illustrated by the following example: if a patient disagreed with the question "Are you too tired to take exercise?", this patient would not also get a question reflecting more severe fatigue as "Are you too tired to go for a short walk?". For the development of a CAT for fatigue in RA, adequate items have to be selected for inclusion in an item bank.

Items of the aforementioned fatigue questionnaires[7-10] are available in Dutch. Though, the patients' perspective was not part of their development and they were not specifically designed for patients with RA. Therefore, it is of particular importance to check content validity.

We conducted a Delphi study to allow patients, rheumatologists and nurses to evaluate fatigue items. It is essential to include the opinion of patients and professionals in the development of questionnaire items to ensure content validity.[13] Only patients can report on the subjective experience of fatigue, and clinicians have the most experience with the outward manifestation of a symptom or condition.[13]

This paper focuses on the evaluation of existing fatigue items: how do experts evaluate the items from traditional fatigue questionnaires (SF-36 subscale vitality, FACIT-F, POMS

subscale fatigue/inertia, MAF)? Do the items of the BRAF-MDQ meet the meaning of fatigue in the Netherlands?

PATIENTS AND METHODS

Delphi process

With the Delphi process,[14] opinions about a certain topic can be collected by a questionnaire, which is sent by (electronic) mail to a panel of potential participants who then fill it out individually. In our case, they rated the fatigue items on their own. Contrary to group discussion, no attention has to be paid to prevent that especially the more dominant participants express their opinion, so that the phenomenon of "group think" cannot obstruct the viewing of different opinions. A Delphi study consists of at least two rounds and aims to reach consensus among the participants. After each Delphi round, systematic feedback on the results of the previous round are provided to the participants.[15] By doing so, the experts are informed about the opinions that are present in the group. The extent of agreement is determined by statistical measures as no in person meetings are conducted.

Preparation of the item pool

In our aim to develop a CAT for fatigue in RA, we need a large and comprehensive item pool to cover all dimensions of the fatigue experience in RA. We included 40 items of the validated Dutch versions of the RAND SF-36 subscale vitality,[16] FACIT-F,[17] POMS subscale fatigue/inertia,[18] and MAF,[19] henceforth referred to as traditional questionnaires. We also included all 20 items of the BRAF-MDQ, which we translated from English into Dutch using the recommended cross-cultural translation procedure.[20] Additional items developed from interview material were enclosed (not reported in this paper, the entire process of the Delphi study is described elsewhere[21]).

Participants and data collection

We contacted the experts (40 rheumatologists, 40 nurses and 31 patients) by e-mail, informing them about the study and requesting their participation. The e-mails contained a link to an online questionnaire in which we presented the fatigue items. The e-mail addresses of patients were collated from the database of Patient Research Partners of the Arthritis Centre Twente and those of rheumatologists and nurses from the member list of a Dutch professional association in Rheumatology (NVR) and the DREAM registry. The participants were asked to indicate on a 4-point Likert scale (very appropriate, rather

appropriate, less appropriate, not at all appropriate) how appropriate they thought each item was. Participants also had the possibility to comment on the items and suggest improvements in the blank fields below each item.
After each round, the data were analysed and summarized. For the items which had to be re-evaluated in round 2, information about its evaluation in round 1 was provided by presenting the relevant item, a brief summary of the comments given, and finally the re-worded item. Items had to be evaluated in the same way as in round 1.

Analysis and criteria for item selection

In Delphi studies, different rules for defining sufficient consensus are applied,[15] ranging from 55 - 80%. To establish the rule for our study, we searched the literature for other Delphi studies in health research. Repeatedly, the criterion of 80% was used for selection of an item or topic, as for EULAR recommendations.[22] On the basis of these examples, we also applied the conservative criterion of 80% agreement. Further properties of our rule were developed in discussion among the authors to meet our special situation with different expert groups.
An item was directly selected for inclusion in the CAT item pool if 80% or more of the participants rated it as appropriate (response options 'very appropriate' and 'rather appropriate'). Agreement by 50% or less of the participants led to the item's rejection.
Besides the global percentage, we also calculated the percentages per expert group. If the mean percentage was sufficiently high (80%) or low (50%), but one group differed by more than 10% from the criterion (below 70% or above 60%), the item was not directly selected or removed. They, and also those items rated as appropriate by 50%-80% of the experts, were discussed by the authors and re-worded according to the participants' comments and suggestions. If the same comment was raised by at least two participants, it led to an adaptation of the item. The re-worded items were judged again in round 2 by the panel. After round 2, comments were discussed among the researchers and the final decision about in- or exclusion of the items was made.

RESULTS

In the first round, 15 patients, 10 rheumatologists and 20 nurses, working at 21 hospitals spread over the Netherlands, returned the completed questionnaire (N=45). In the second round, 80% of the participants of the first round (15 nurses, 13 patients and 8 rheumatologists) participated (N=36). In both rounds, our expert panel clearly consisted of

more women than men (2:1) and the mean age was 51.4 (SD=11.4) in the first round and 52.4 (SD=10.5) in the second round.

Evaluation of the items

Round 1

An immediate decision could be made on 18 of the 40 items of the traditional questionnaires (SF-36 subscale vitality, FACIT-F, POMS subscale fatigue/inertia, MAF). Sixteen items (40%) were rated as appropriate by at least 80% of the participants so that they were directly selected for the development of the CAT. Two items were rejected because they were rated as appropriate by 50% or less of the participants. Item 7 of the POMS "Over the past 7 days I felt bushed" was excluded. Participants rated this item as too negative, dramatic, extreme or exaggerated. Moreover they noted that the item might be confusing because it could also reflect physical circumstances such as a loss of strength in the joints, or have a psychological component such as depressive feelings. The second excluded item was question 2 of the MAF "How severe is the fatigue which you have been experiencing during the past week? 1=mild 2 3 4 5 6 7 8 9 10=severe". In the official Dutch questionnaire, fatigue was translated as '*uitputting*' which actually means 'exhaustion' and it was criticized that exhaustion is always assumed to be present. Furthermore, participants said that it was not possible to measure exhaustion with the used scale since you are either exhausted or not, i.e., that exhaustion is always serious.

Regarding the BRAF-MDQ, 12 out of the 20 items (60%) were directly selected and 1 item (5%) was excluded. This was question 12 "Over the past 7 days, have you felt embarrassed because of fatigue?", which participants found ambiguous and too difficult.

The authors discussed the remaining items (>50% and <80% experts agreed) of the five scales, whereby the participants' comments and suggestions were considered. Items that did not receive comments on their content but on their formulation or the response options, for example, were adjusted (phrase or word replaced, deleted or added) and included in the item pool. This applied to 17 (42.5%) items of the traditional questionnaires. A major concern was again the use of the Dutch word for exhaustion in the MAF. Participants found this term inappropriate and advised using another word that reflects the word 'fatigue' better. Accordingly, we replaced the term '*uitputting*/exhaustion' with '*vermoeidheid*/fatigue' in all concerned MAF items and as respondents found the formulation of most MAF items cumbersome, we shortened them. Furthermore, experts rated the response options of the FACIT-F as unclear, so we re-worded these too. Table 1 shows examples of these adaptations. Item 5 of the POMS 'Over the past 7 days I felt sluggish' was excluded after discussion, the reason being that respondents rated this item as unclear and no alternatives were found for reformulation.

The authors selected 3 (15%) BRAF-MDQ items after minor adaptations, such as the replacement or deletion of a word.

Four items (10%) of the traditional questionnaires needed more extensive adaptations (e.g., inclusion of examples, reformulation of an entire item) and were presented for re-evaluation in round 2. Four BRAF-MDQ items (20%) were presented for re-evaluation either because examples had to be inserted or because only criticism but no suggestion for improvement had been given.

Table 1 Examples for adaptations and selection of items by the authors

Item in round 1	Comments	Adapted and selected item
MAF item 5		
In the past week, to what degree has fatigue interfered with your ability to cook? 1 = not at all 2 3 4 5 6 7 8 9 10 = a great deal	Formulate item and response options easier, replace the Dutch word for exhaustion (*uitputting*) with a word that refers to fatigue better (*vermoeidheid*) because exhaustion is too extreme, it is possible that cooking is not applicable to a patient.	During the past 7 days, how did your fatigue prevent you from cooking? 1=not at all, 2=a little, 3=rather, 4= to a great extent, 5=does not apply to me
FACIT-F item 4		
During the last 7 days, I felt tired. 0=not at all, 1=a little bit, 2=somewhat, 3=quite a bit, 4=very much	The response options are unclear (translated in Dutch as *0=helemaal niet, 1=een beetje, 2=enigzins*, *3=in vrij hoge* mate, 4=*in zeer hoge mate*); the formulation of option 3 and 4 is too formal, and the difference between option 1 and 2 is not clear.	During the past 7 days, I felt tired. 1=not at all, 2=a little, 3=rather, 4=to a great extent (Dutch: *1=helemaal niet, 2=een beetje, 3=nogal*, *4=in sterke mate*)

Round 2

Table II shows examples of items included in round 2. After the second round, 3 adapted items of the traditional questionnaires were evaluated as adequate by more than 80% of the participants and could immediately be selected. Participants had no complaints about the general adaptations made to these items (e.g., different word in MAF, response options FACIT-F). Of the 4 BRAF-MDQ items included in round 2, two were rated as adequate by at least 80% of the participants and immediately selected.

Based on the comments given, the authors decided to exclude one item of the traditional questionnaires, namely question 1 of the POMS (see table 2). This item was adjusted

according to the comments; the Dutch translation for 'worn out' (*doodop*) was replaced with a synonym (*bekaf*) and the response options were adapted. Despite this, many participants still found this item unclear and too negative. Two BRAF-MDQ items were included with adapted response options after discussion between the authors.

Table 2 Examples of adapted items presented for re-evaluation in round 2.

Item origin	Original item	Adapted item
Finally selected		
FACIT-F item 2	During the last 7 days I felt weak all over. 0=not at all, 1=a little bit, 2=somewhat,. 3=quite a bit, 4=very much	During the past 7 days, my fatigue made me feel weak. 1=not at all, 2=a little, 3=rather, 4=to a great extent
MAF item 3	In the past week to what degree has fatigue caused you distress? 1-10 / 1 = no distress, 10 = a great deal of distress	Was your fatigue during the past 7 days aggravating? 1=not at all, 2=a little, 3=rather, 4=to a great extent
MAF item 14	In the past week, to what degree has fatigue interfered with your ability to exercise, other than walking? 1 = not at all 2 3 4 5 6 7 8 9 10 = a great deal	During the past 7 days, how did your fatigue prevent you from moving (e.g. waling, cycling, playing sports)? 1=not at all, 2=a little, 3=rather, 4=to a great extent
BRAF-MDQ item 12	Have you lacked mental energy because of fatigue? 1=not at all, 2=a little, 3= quite a bit, 4=very much	Have you lacked mental energy (e.g. energy to think properly about sth., to make plans, to be creative) because of fatigue? 1=not at all, 2=a little, 3=quite a bit, 4=very much
BRAF-MDQ item 20	Have you felt down or depressed because of fatigue? 1=not at all, 2=a little, 3=quite a bit, 4=very much	Have you felt down or dejected because of fatigue? 1=not at all, 2=a little, 3=quite a bit, 4=very much
Finally excluded		
POMS item 1	Over the past 7 days I felt worn out (Dutch translation: *doodo*p). 1=not at all, 2=a little, 3=moderate, 4=quite a bit, 5=extremely	Over the past 7 days I felt worn out (Dutch: *bekaf*). 1=never, 2=rarely, 3=sometimes, 4=usually, 5=always

Finally, 36 items (90%) from the traditional questionnaires and 19 items (95%) from the BRAF-MDQ were included in our item pool. Table 3 (as provided in the appendix at the end of this article) shows the percentage of items per scale in each selection step.

DISCUSSION

This Delphi study provided insights into the perspective of patients and professionals on fatigue items of four traditional scales (SF-36 subscale vitality, FACIT-F, POMS subscale fatigue/inertia and MAF) and a new questionnaire (BRAF-MDQ) developed in the UK.

It appeared that only 40% of the items from traditional questionnaires were clear and adequate in their original style for at least 80% of the participants. This finding underlines the relevance of the perspective of professionals and patients in the development of measurement instruments for fatigue in RA. However, most of the items could be selected after adaptations, with only 4 of the 40 items being excluded. Remarkably, three of these items originated from the POMS. One item was rated as too extreme and confusing, because it could also be related to circumstances other than fatigue; the other items were also rated as too extreme or as unclear. The fourth excluded item came from the MAF and had to be directly excluded due to insufficient percentages of ratings as adequate. The item would probably have received more positive evaluations, though, with the word '*vermoeidheid*' that we used to replace '*uitputting*' in all of the selected MAF items. The reaction of our experts to the term '*uitputting*' suggests that the Dutch version of the MAF might sooner measure exhaustion than fatigue. This finding points to the need to carefully check and re-check cross-cultural translations to ensure that the intended measurement aim is met. However, our results cannot be generalized to the items in the original English questionnaire since it was the official Dutch translations that were evaluated in this study. An evaluation of the original English items by rheumatologists, nurses and patients could provide more insight.

The BRAF-MDQ turned out to suit the fatigue experience in a Dutch population well. This may be attributed to its advantage of including the patients' perspective in the development process. Only one item had to be excluded and was about feeling embarrassed. It is difficult to fully explain why it received low percentages of agreement since only few comments were given. A possible explanation might be that British patients might feel more embarrassed about fatigue than Dutch patients since they live in a so-called 'stiff upper lip', masculine culture in which it is not so common to show weakness.[2,23] In contrast, the Netherlands is considered a more feminine and openly nurturing culture.[23]

This study was limited to the Netherlands, so we are not able to draw conclusions about the quality of the examined items in other countries. Furthermore, there were no generally accepted criteria we could apply for our in- and exclusion of items. However, we carefully deliberated our criteria and chose for the conservative criterion of 80% agreement as inclusion criterion that was frequently used in other Delphi studies.[22] A considerable amount of the criticism on the fatigue items was related to translation difficulties rather than to their content. To conclude, 90% of the traditional questionnaire items and 95% of the BRAF-MDQ items showed good content validity and feasibility (partly in an adapted version) for use in the Netherlands and will be included in our item pool for the development of a CAT for fatigue in RA. This item pool contains further items that were validated by our expert panel, for example items based on interview material.[21] We assume that our item pool is comprehensive and adequate for the measurement of fatigue in RA. In a next step, the items have to be scaled according to item response theory (IRT). With IRT, item parameters as the difficulty level can be assessed for each item independently.[24] This information is required to ideally match the items to the patient's individual level and construct the CAT for fatigue in RA.

ACKNOWLEDGEMENTS

We would like to thank the following participants in this study: patients from the forum of Patient Research Partners of the Arthritis Centre Twente and rheumatologists and nurses from Universitair Medisch Centrum Groningen, Medisch Centrum Leeuwarden, Ziekenhuisgroep Twente, Zorg Groep Twente, Medisch Spectrum Twente, St. Maartenskliniek Nijmegen, TweeSteden Ziekenhuis Locatie Tilburg, Universitair Medisch Centrum Utrecht, Rijnstate Ziekenhuis Arnhem, Meander Medisch Centrum Amersfoort, Ziekenhuis Gelderse Vallei Ede, Medisch Centrum Haaglanden s-Gravenhage, Tweesteden Ziekenhuis Tilburg, Evean Thuiszorg Purmerend, Streekziekenhuis Midden Twente Hengelo, Isala Klinieken Zwolle, Albert Schweitzer Ziekenhuis Dordrecht, Spaarne Ziekenhuis Hoofddorp, Scheperziekenhuis Emmen, Thuiszorg West-Brabant Roosendal, Omring Thuiszorg Hoorn.

REFERENCES

1. Minnock P, Kirwan J, Bresnihan B. Fatigue is a reliable, sensitive and unique outcome measure in rheumatoid arthritis. Rheumatology 2009;48:1533-1536.
2. Hewlett S, Cockshott Z, Byron M, Kitchen K, Tipler S, Pope D, et al. Patients' perceptions of fatigue in rheumatoid arthritis: overwhelming, uncontrollable, ignored. Arthritis Care Res 2005;53:697-702.
3. Nikolaus S, Bode C, Taal E, van de Laar MAFJ. Four different patterns of fatigue in rheumatoid arthritis patients: results of a Q-sort study. Rheumatology 2010;49:2191-2199.
4. Van Oers ML, Bossema ER, Thoolen BJ, Hartkamp A, Dekkers JC, Godaert GL, et al. Variability of fatigue during the day in patients with primary Sjögren's syndrome, systemic lupus erythematosus, and rheumatoid arthritis. Clin Exp Rheumatol 2010;28:715-21.
5. Strand V, Khanna D. The impact of rheumatoid arthritis and treatment on patients' lives. Clin Exp Rheumatol 2010;28:S32-40.
6. Hewlett S, Hehir M, Kirwan JR. Measuring fatigue in rheumatoid arthritis: a systematic review of scales in use. Arthritis Rheum 2007;57:429-439.
7. Ware JE Jr, Sherbourne CD. The MOS 36-Item Short-Form Health Survey (SF-36). I. Conceptual framework and item selection. Med Care 1992;30:473-483.
8. Cella D, Yount S, Sorensen M, Chartash E, Sengupta N, Grober J. Validation of the Functional Assessment of Chronic Illness Therapy Fatigue Scale relative to other instrumentation in patients with rheumatoid arthritis. J Rheumatol 2005;32:811-819.
9. McNair D, Lorr M, Droppelman L. Profile of Mood States Manual. New York: Multi-health Systems Inc; 1992.
10. Belza Tack B. Dimensions and correlates of fatigue in older adults with rheumatoid arthritis [dissertation]. San Francisco: University of California; 1991.
11. Nicklin J, Cramp F, Kirwan J, Greenwood R, Urban M, Hewlett S. Measuring fatigue in rheumatoid arthritis: A cross-sectional study to evaluate the Bristol Rheumatoid Arthritis Fatigue Multi-Dimensional questionnaire, visual analogue scales, and numerical rating scales. Arthritis Care Res 2010;62:1559-1568.
12. Rose M, Bezjak A. Logistics of collecting patient-reported outcomes (PROs) in clinical practice: an overview and practical examples. Qual Life Res 2009;18:125-136.
13. Streiner DL, Norman GR, eds. Health measurement scales – a practical guide to their development and use. New York: Oxford University Press; 2003.
14. Jones J, Hunter D. Qualitative research: consensus methods for medical and health services research. BMJ 1995;311:376-380.
15. Powell, C. (2003). The Delphi technique: myths and realities. Journal of Advanced Nursing, 41, 376-382.

16. Van der Zee KI, Sanderman, R. Het meten van de algemene gezondheidstoestand met de RAND-36 – een handleiding. Noordelijk Centrum voor Gezondheidsvraagstukken, Rijksuniversiteit Groningen. 1993.
17. FACIT Functional Assessment of Chronic Illness Therapy. Retrieved July 16, 2009, from www.facit.org
18. Cluydts, RJG. Gemoedstoestanden en slaap – een experimenteel onderzoek [dissertation]. Brussel: Vrije Universiteit Brussel; 1979.
19. MAPI INSTITUTE. Retrieved July 31, 2009, from www.mapi-institute.com
20. Guillemin F, Bombardier C, Beaton D. Cross-cultural adaptation of health-related quality of life measures: literature review and proposed guidelines. J Clin Epidemiol 1993;46:1417-1432.
21. Nikolaus S, Bode C, Taal E, van de Laar MAFJ. Selection of items for a computer-adaptive test to measure fatigue in patients with rheumatoid arthritis – A Delphi approach. Qual Life Res. Published online 31 July 2011. doi: 10.1007/s11136-011-9982-8.
22. Avouac J, Kowal-Bielecka O, Landewe R, Chwiesko S, Miniati I, Czirjak L, et al. European League Against Rheumatism (EULAR) Scleroderma Trial and Research group (EUSTAR) recommendations for the treatment of systemic sclerosis: methods of elaboration and results of systematic literature research. Ann Rheum Dis 2009;68:629-634.
23. Geert Hofstede™ Cultural Dimensions. Retrieved July 22, 2010, from www.geert-hofstede.com/hofstede_dimensions.php?culture1=62&culture2=94
24. Hambleton RK, Swaminathan H, Rogers HJ. Fundamentals of Item Response Theory. Sage Publications; 1991.

APPENDIX: Table 3 Selection of fatigue items from the four traditional scales and the BRAF-MDQ

Scale/ Evaluation	RAND-36 4 items	FACIT-F 13 items	POMS 7 items	MAF 16 items	BRAF-MDQ 20 items	Summary 60 items
Round 1:						
Directly selected	3 (75%)	9 (69.2%)	1 (14.3%)	3 (18.8%)	12 (60%)	28 (46.6%)
Selected after adaptation	1 (25%)	3 (23.1%)	3 (42.9%)	10 (62.5%)	3 (15%)	20 (33.3%)
Directly rejected			1 (14.3%)	1 (6.25%)	1 (5%)	3 (5%)
Rejected after discussion			1 (14.3%)			1 (1.7%)
Adapted and selected for re-evaluation in round 2	0	1	1	2	4	8
Round 2:						
Directly selected		1 (7.7%)		2 (12.5%)	2 (10%)	5 (8.3%)
Selected after discussion					2 (10%)	2 (3.3%)
Directly rejected						
Rejected after discussion			1 (14.3%)			1 (1.7%)
Finally **excluded** items	0	0	3	1	1	5 (8%)

Chapter 9

Calibration of a multidimensional item bank to measure fatigue in rheumatoid arthritis patients

S. Nikolaus
C. Bode
E. Taal
A. Oostveen
C.A.W. Glas
M.A.F.J. van de Laar

In preparation for submission

ABSTRACT

Objective Goal of this study was the calibration of an item pool to measure fatigue in rheumatoid arthritis (RA). It was based on the patients´ perspective captured by an interview study, and examined for face and content validity by a previous Delphi study with patients and professionals. The fit of the items with the underlying dimensions was assessed with item response theory (IRT) and the dimensionality structure of the item pool was examined by factor analysis and multidimensional IRT.

Methods Participants were 551 patients with RA from three hospitals in the Netherlands. Obviously, it was not feasible to let each patient score all 245 items of our item pool, so we used an item administration design to construct seven different questionnaire versions. Each patient completed one version of the questionnaire, maximally containing 126 items. IRT analysis using the generalized partial credit model (GPCM) was conducted for each dimension of fatigue. Poorly fitting items were removed. Consecutively, exploratory and confirmatory factor analysis was performed on the remaining items and a multidimensional IRT model was fitted.

Results In the IRT analysis, 49 items showed insufficient item characteristics. Items with a discriminative ability <0.60 and/or model misfit effect sizes >0.10 were removed. Exploratory and confirmatory factor analysis on the 196 remaining items revealed three dimensions of fatigue named: severity, impact and variability of fatigue. The dimensions were further confirmed in multidimensional IRT model analysis.

Conclusion This study provided an initially calibrated multidimensional item bank and has shown which dimensions and items that came forward from previous studies are important for the development of a multidimensional computerized adaptive test (CAT) for fatigue in RA.

Many patients with rheumatoid arthritis (RA) complain about fatigue.[1,2] It is not yet understood which mechanisms cause fatigue in RA.[3] Most patients try to manage their fatigue by trial and error and report to receive no adequate medical or para-medical support.[4] Patients experience fatigue as a multidimensional, annoying symptom with far-reaching consequences.[5-8] They mention not only physical experiences, but also cognitive and emotional elements of fatigue. Patients describe their fatigue being different from normal tiredness, as it is often more extreme, not always caused by high levels of activity and therefore unpredictable.[6]

The issue of fatigue was raised by patients and thoroughly discussed on OMERACT meetings (international group of experts on outcome measures in rheumatology).[9] Thereupon, fatigue was recommended to be included in the core set of outcome variables in clinical trials.[10,11] The number of studies on RA using fatigue as outcome rapidly increased over the last years.

Measuring fatigue provides important information for the understanding of the patients´ perspective on disease impact and outcome.[12] The measurement of fatigue with single item scales has some but limited value,[12] since it does not correspond to the multidimensional character of fatigue.[5-8]

Existing multi-item measurement instruments that include scales about aspects of fatigue, as the Short Form 36 subscale vitality (SF-36),[13] Functional Assessment of Chronic Illness Therapy Fatigue Scale (FACIT-F),[14] Profile of Mood States subscale fatigue/inertia (POMS),[15] and the Multidimensional Assessment of Fatigue scale (MAF),[16] have the shortcoming of unsatisfactory content validity for fatigue in RA as they were not developed from the perspective of patients with RA.[17] The opinion of patients and professionals in the field is essential in the development of questionnaire items to ensure content validity.[18,19] Only patients can report on the experience of fatigue,[9,20] where clinicians have the most experience with the outward manifestation of a symptom or condition.[20]

The recently constructed Bristol Rheumatoid Arthritis Fatigue Multi-Dimensional Questionnaire (BRAF-MDQ)[21] includes the patients´ perspective and is multidimensional, but has as previously mentioned questionnaires, a traditional format. That means the instruments are fixed length-questionnaires. Using them is time consuming because patients have to answer questions that may not apply to their personal situation. Also, fixed lengths-questionnaires suffer from repeated testing-effects because a patient always fills in the same questions and so can easily remember the questions and previously given answers. Furthermore, existing questionnaires do not capture all aspects of fatigue that should be measured.[22] So there is a need for more appropriate and efficient ways of measurement.

Computerized adaptive testing (CAT) is a promising technology as it provides the possibility to comprehensively measure patient reported outcomes (PROs) with relatively few items.[23] The burden for patients decreases since it is no longer necessary that everybody answers the same questions and consequently all items of a questionnaire. Using a CAT increases measurement precision since items are sequentially selected from a calibrated item bank, based on the previous answer of this patient. In this way precise measurement at individual level with few items becomes possible. For the computerized selection of the best matching items, a large item bank is needed that contains much more items than are finally presented to an individual patient.[24] Before a CAT can be developed, an item bank has to be scaled. In contrast with classical test theory, with item response theory (IRT), item parameters such as the difficulty level can be estimated for each individual item as well as the scale values for fatigue levels.[24] In our case this means that we can estimate the level of fatigue reflected by the item. All items are placed on this continuum, ranging from no fatigue to severe fatigue.

Furthermore, with IRT it can be calculated how well an item discriminates between more or less fatigued patients. This information is required to optimally match the items to the patient's individual level and this information supports inter-individual comparisons on the measured construct even if patients filled in different items. Primarily, CAT was used for ability and achievement testing, and it is an upcoming and innovative domain in medical settings.[25] A joined initiative is developing a Patient-Reported Outcome Measurement Information System (PROMIS) aiming to construct a large item bank and CAT system for the assessment of PROs in chronic diseases.[26] CATs for depression, anxiety and stress perception turned out to be reliable, valid and efficient instruments that measure more precisely than traditional questionnaires.[27-29] Moreover, they are less time consuming so that the burden for patients is reduced.

We aim to develop a CAT fatigue in RA. So, we constructed an item pool to measure fatigue that is based on the patients´ perspective. To ensure that all relevant aspects of fatigue are included, we made several steps. First, we conducted in-depths interviews to describe the individual experience of fatigue.[8] Secondly, we examined whether groups of patients can be distinguished that experience fatigue in different ways. For this end we used Q-methodology[30] whereby patients sorted cards with statements about their fatigue. These statements informed the new item formulations. In a third step, a preliminary item pool was constructed. Aim was to build up a large and comprehensive item pool to cover all dimensions of fatigue experience in RA. Therefore we included all items and dimensions of validated fatigue scales in RA, and supplemented them with items from interview material and Q-sort study. To check for possibly missing aspects we also examined fatigue questionnaires not validated for RA. This procedure resulted in 294

items distributed over 12 dimensions (severity, frequency, duration, changes in fatigue, perceived causes of fatigue, energy, sleep/rest, body feeling, cognition/concentration, coping, negative emotions/mood, consequences).

In a next step, the item pool was evaluated in a Delphi study with experts from the Netherlands (patients, nurses, rheumatologists) to select adequate items and dimensions to measure fatigue in RA.[31-33] Items were included if 80% of participants rated items as adequate, reject if 50% or less rated an item as adequate, and remaining items were discussed and adjusted according to suggestions and comments of the participants and presented for re-evaluation. After two rounds all items were included or rejected. The final content valid item pool consisted of 245 items and 12 dimensions.

This consecutive study is intended to calibrate and factor analyze these items to finally build the CAT RA fatigue. We do not intend to develop one CAT for each dimension of fatigue, but to produce a multidimensional instrument. Multidimensional adaptive testing provides not only information about the level of a participant on each dimension, but also about the amount of association between dimensions in the population.[34] The cross-information gained from items of correlated dimensions can be effectively modeled by multidimensional IRT. It facilitates CAT by supporting the selection of optimal, informative items, and by supporting the estimation of ability with optimal precision. According to Segall,[34] multidimensional adaptive testing offers equal or even higher precision with approximately one third fewer items than would be needed in unidimensional adaptive testing. So, the increased measurement efficiency manifests itself by both, greater precision as well as reduced test lengths.[34] With this innovative method, measuring fatigue in RA can become more precise and at the same time more user-friendly.

PATIENTS AND METHODS

Patients

Altogether 999 patients with RA of the Ziekenhuis Groep Twente (Almelo and Hengelo) and the Arthritis Center Twente at Medical Spectrum Twente (Enschede) were invited for participation. The potential participants were invited via a letter from their rheumatologist, informing them about the aim of the study and details about participation. Each letter was accompanied by a questionnaire consisting of one version of the fatigue questionnaire, an informed consent form and a paid return envelope.

Questionnaire

Patient characteristics

The questionnaire started with questions about demographic characteristics (gender, date of birth, marital status, level of education, work status) and disease characteristics (year of disease onset, comorbidities).

Disease related measures

Patients filled in eleven-point numerical rating scales (NRS) for pain and impact of the disease, and three fatigue NRS[21] for severity ('Please circle the number which shows your average level of fatigue during the past 7 days' (anchors: no fatigue—totally exhausted); impact ('Please circle the number which shows the effect fatigue has had on your life during the past 7 days' (anchors: no effect—a great deal of effect); and coping ('Please circle the number which shows how well you have coped with fatigue over the past 7 days' (anchors: not at all well—very well). As these three NRSs were developed in the UK, they were translated into Dutch according to the recommended translation procedure for cross-cultural adaptation.[35] Furthermore, patients completed the 10-item version of the HAQ-II.[36]

Fatigue items

Due to the large amount of items in our item pool, it was not feasible to let each participant fill in all fatigue items of our item pool. The burden of answering 245 questions about fatigue would be too high, so we prepared different versions of questionnaires containing between 103 and 126 items. This resulted in a common-item linking design.[37] That means that each questionnaire version consists of a different composition of dimensions and items, but also has some sets of items in common (see appendix 1 at the end of this article for the item administration design). In this way, items and dimensions of the different questionnaire versions can be related to each other in the IRT analysis.[37]

Our previous Delphi study revealed 12 content valid dimensions of fatigue;[32] "severity", "frequency", "duration", "changes in fatigue", "perceived causes of fatigue", "energy", "sleep/rest", "body feeling", "cognition/concentration", "coping", "negative emotions/mood", "consequences". However, some of the dimensions are closely related to each other (e.g. severity, frequency and duration are all about the manifestation of fatigue) and are measured as one dimension in other fatigue questionnaires. They were combined in the further analyses. In contrast, other dimensions, as came forward from the Delphi study, are not yet frequently covered by fatigue scales (e.g. changes in fatigue, perceived causes of fatigue and coping with fatigue) and therefore separately included in

the further analysis. By this means we wanted to find out how these dimensions, consisting of many newly constructed items, would fit in the IRT model. An overview of the original and newly formed dimensions ("severity", "physical", "mental", "changes in fatigue", "perceived causes of fatigue", "coping" and "consequences") is given in Table 1.

Table 1 Dimensions of fatigue, summarized for IRT analysis

Original dimensions	**Larger dimensions**
-Severity, frequency, duration	SEVERITY
-Sleep/rest, body feeling, energy (items of the energy dimension were split up into physical and mental, depending on their content)	PHYSICAL
-Cognition/concentration, negative emotions/mood, energy	MENTAL
-Changes in fatigue - "new" dimension	CHANGES IN FATIGUE
-Perceived causes of fatigue – "new" dimension	PERCEIVED CAUSES
- Coping – "new" dimension	COPING
-Consequences - was already a very large dimension	CONSEQUENCES

Analyses

Aim of the analyses was to build a between-items-multi-dimensional IRT model to create an initial calibration for an item bank to run a CAT. The objective was to define the dimensions using the seven previously described dimensions, but to aim at a lower dimensionality. Four steps were executed.

Step 1: Initial item selection

For each dimension, the fit to a unidimensional IRT model and the measurement quality was investigated. Because of the incomplete item administration design, this was done using IRT and not using classical test theory. For each of the seven dimensions, an IRT analysis was conducted with public domain software MIRT[38] under the generalized partial credit model (GPCM).[39] This two-parameter logistic model is applicable to analyze polytomous items, meaning items with more than two response options, and allows the items within a scale to differ in discrimination parameter values.[40]

The discrimination parameter is highly correlated with the item/rest-score.[41] The item/rest-score-correlation is the correlation between a specific item response and the total score without the specific item. In classical test theory it is used as an indication for the contribution of the item to the reliability of the test. We dismissed all items with a discrimination parameter below 0.60. Such items contribute little to the overall reliability

and will not be selected in the CAT anyway. A low discrimination parameter means that an item does not discriminate well between more or less fatigued persons. Further, we excluded items with an item misfit-index larger than 0.10.[42] Misfit is present if the probability to respond in the different categories of the item varies across subgroups of patients or the same level of fatigue.[43]

Step 2: IRT analysis of the fit for the separate original dimensions after removal of the unreliable items.

We repeated the IRT analyses carried out in step 1, but without the excluded items. The number of significant model tests and reliabilities as estimated under the GPCM, are provided in the result section. If the data strictly fit the model, the percentage of model tests which are significant at the 5% level should also be approximately 5%.

Step 3: determining the dimensionality of the item pool

A problem was that the dataset has too many items (K=196) to run exploratory analyses in standard software such as Mplus. To solve this problem, further analyses were conducted on IRT-based person parameter estimates. To account for measurement error, the seven dimensions were split up into two parts, containing odd and even items, and plausible values of the person parameters were drawn for both sets.[44] So the input for MPlus consisted of 14 variables for each patient. For the interpretation of model fit, the root mean square error of approximation (RMSEA) was used as criterion. A RMSEA value smaller or equal to 0.05 indicates a close fit between the observed correlation matrix and the correlation matrix expected under the model. Values between 0.05 and 0.08 suggest a reasonable approximation.

Step 4: Confirmative MIRT analysis using the results of step 3.

In step 4, the results of the exploratory analysis in step 3 were further confirmed using multdimensional IRT to end up with a model suited for item bank calibration and CAT. To evaluate the fit to the IRT model, counts of significant item tests were used.

Participants

We received 551 completed questionnaires, a response rate of 55%. The sample consisted of 367 women, 164 men and 20 persons who did not fill in their sex. Mean age was 63.38 (SD=12.7) and the mean disease duration was 15.15 years (SD=11.22). The level of fatigue, measured with the NRS, was 4.89 (SD=2.3), ranging from 0 to 10, showing that a broad

range of fatigue was represented among the participants. More information about the sample is shown in tables 2 and 3.

Table 2 Sample characteristics (N=551)

	Mean (S.D.)	Range
Age in years	63.38 (12.70)	24 - 92
Disease duration in years	15.15 (11.22)	0 - 67
NRS General health	4.54 (2.11)	0 - 10
NRS Pain	4.38 (2.38)	0 - 10
NRS Fatigue severity	4.89 (2.30)	0 - 10
NRS Impact of fatigue	4.62 (2.53)	0 - 10
NRS Coping with fatigue	6.50 (2.02)	0 - 10
HAQ-II score	1.00 (0.65)	0 - 3

NRS= Numerical Rating Scale, HAQ=Health Assessment Questionnaire

Table 3 Sample characteristics (N=551)

	N
Sex	
Women	367
Men	164
Marital status	
Single	27
Living with partner/married	412
Widow/widower	73
Divorced	33
Level of education	
Low (≤ 12 years of education)	362
Moderate (13-14 years of education)	109
High (≥ 14 years of education)	71
Work status	
Working full-time	58
Working part-time	89
Household/unemployed	105
Disabled/Retired	290
Co-morbidities	
Yes	260
No	291

RESULTS

Step 1: Initial item selection

We excluded 49 items due to unsatisfactory item characteristics; having a discriminative ability less than 0.60 and/or an absolute item fit index larger than 0.10. This applied to items in all dimensions. Table 1 shows the number of excluded items per dimension.

Step 2: IRT analysis of the fit for the separate original dimensions after removal of the unreliable items

We counted the number of significant model tests among the different questionnaire versions per dimension as provided by Lagrange tests for GPCM. They are provided in table 4. Also the reliabilities as estimated under the GPCM are shown per dimension.

Table 4 Significant model test and reliability per dimension (estimated with GPCM)

Dimension	significant model tests at 5%	reliability
Severity	24 of 104 (23.1 %)	0.959
Physical	85 of 169 (50.3%)	0.975
Mental	19 of 224 (8.5%)	0.974
Consequences	19 of 307 (6.2%)	0.978
Change	4 of 45 (8.9%)	0.701
Perceived causes	13 of 79 (16.5%)	0.646
Coping	12 of 88 (13.7%)	0.878

Note that the percentages of significant model tests are too high for the “severity” and “physical” dimensions, so for these two dimensions unidimensionality was not supported. The percentages of significance probabilities for the dimensions “mental”, “consequences” and “change” were quite close to the nominal significance probability of 5%, so here unidimensionality was considered acceptable.

Step 3: determining the dimensionality of the item pool

Four factor solutions were taken into account. The following results were found for the four analyses:

1. A factor solution with one dimension was rejected. The analysis resulted in a Root Mean Square Error of Approximation of 0.098, that is, RMSEA=0.098. The maximal bound for the RMSEA is usually taken as 0.05. The test of the hypothesis that the RMSEA is smaller than 0.05 was highly significant, that is, $p(RMSEA \leq 0.05) < 0.001$.

2. A model with two dimensions showed a good fit (RSMEA= 0.049, $p(RMSE \leq 0.05) = 0.485$), but the dimensions were hard to interpret, and the subsequent confirmatory MIRT analysis did not support between-items multidimensionality. All items loaded to some degree on both dimension, resulting in an un-interpretable within-items MIRT model.

3. Three dimensions fitted well and lead to RMSE=0.041 ($p(RMSE \leq 0.05) = 0.534$). This solution can also be interpreted in theoretical terms. The first dimension (factor) consists of the original dimension 1 (severity of fatigue), the second dimension (factor) consists of the original dimension 2 (physical), 3 (mental), 4 (consequences) and 7 (coping) – all dimensions referring to the impact/consequences of fatigue in a broader sense, and the third dimension (factor) consists of the original dimensions 5 (changes) and 6 (perceived causes). These two latter dimensions contain several new formulated items and refer to aspects of the variability of fatigue.

4. Four dimensions (with the third dimension splitted into two dimensions: 5 (changes) and 6 (perceived causes)) did not result in better model fit. That is, the likelihood ratio test of a model with three dimensions against a model with four dimensions had a Chi-square value of 2.133 with 3 degrees of freedom. That is, using four dimensions did not significantly improve model fit.

Step 4: Confirmative MIRT analysis using the 3-dimensional model

The three dimensional IRT model was compared with a one-dimensional GPCM using a likelihood-ratio test. The value of the chi-square was 148, with 2 degrees of freedom, so the unidimensional model was clearly rejected. Analogous to the test of model fit of the original dimensions, fit to the IRT model was evaluated using counts of significant item tests. This resulted in 90 tests significant at 5% out of 597 tests conducted (15%). The conclusion was that the model fit was reasonable. The correlations of the latent variables in the multidimensional GPCM are shown in table 5. Note that the correlations are moderate.

Table 5 Estimated correlations between the three dimensions

	Severity (1)	Impact (2)	Variability of fatigue (3)
1	1.000	0.495	0.247
2	0.495	1.000	0.580
3	0.247	0.580	1.000

Conclusion of the data analysis

The initially calibrated multidimensional item pool consists of 196 items, spread among three dimensions; severity (severity), impact (physical, mental, consequences, coping) and variability (change, perceived causes). These items are displayed in Appendix 2 of this article.

DISCUSSION

This study provided the first calibrated item pool for the multidimensional testing of fatigue in patients with RA, a necessity to develop a multidimensional CAT for fatigue in RA.

The strength of the item pool lies in its stepwise development from the patients´ perspective and the thorough selection of meaningful items and dimensions for the measurement of fatigue in RA. Before the statistical analyses described in this paper, our item pool consisted of 245 items and 12 dimensions that were qualitatively evaluated by our expert panel.[31-33]

These dimensions were already summarized into larger categories for the construction of the linking design that we used for the composition of the different versions of questionnaires that patients completed in this study. In IRT analysis, items with insufficient item characteristics were omitted from the pool and consecutive factor analysis clearly showed that the model with three factors fitted our item pool best.

This solution also makes sense in theoretical terms. The first factor is the dimension "severity" (see table 2), containing items about the intensity, frequency and duration of fatigue. The second factor embraces several relatively large dimensions; "physical", "mental", "consequences", and "coping". All the items have in common that they are about impact of fatigue, on physical and mental level and impact directly related to different aspects of daily life as already included in our previous dimension called "consequences". That the dimension "coping" also belongs to the second factor is of special interest. It is one of the relative small dimensions, we did not summarize to a larger dimension for the IRT analysis because it contains items that are not frequently included in other fatigue instruments yet. Items about what people did or did not do to cope with their fatigue can also be regarded as a consequence or impact of fatigue because these behaviors are resulting from the fatigue. The two other "new" dimensions, namely "change" and "perceived causes" form the third factor. They both do not contain items about the impact of fatigue, but refer to the changing character described by patients[8] and the reasons patients attribute to their fatigue. The third factor is clearly less stable than the first two in psychometric terms. However it reflects important aspects of the patient perspective on fatigue. In our aim to include the patients´ perspective into measurement instruments, it is important to trade psychometric results against information gained from experiences of patients. To ensure that items from this less stable third factor will be drawn in the adaptive testing process, it could be a possibility to place accordant restrictions on the CAT.[45]

The results of this study are also important for the ongoing discussion about the need for multidimensional measurement of fatigue in RA. Our findings clearly underline the importance of multidimensional measurement. In exploratory factor analysis, the one-dimensional model had to be rejected immediately. Also when comparing the fit of the three-factor model with the one-dimensional model, the multidimensional model turned out to be better. This finding supports the view that fatigue in RA should be measured multidimensionally. For this aim, single item instruments as frequently used VAS fatigue scales are not appropriate. Adequate multidimensional measurement is not only essential for scientific purposes, as science needs precise measures of fatigue to get more insight into fatigue and its causes and impact and to be able to evaluate interventions or treatments to reduce fatigue.[46] Also in clinical practice, appropriate measurement of fatigue is important. Only by detailed measuring, understanding fatigue in clinical practice is possible and in the end might stimulate the development of interventions.

In conclusion, the calibrated item pool is thoroughly developed from the patients' perspective and includes not only already established aspects of fatigue, but also newer aspects that were brought up by patients. Furthermore, it consists of three dimensions what matches the multidimensional character of fatigue that was repeatedly described by patients.[5-8] This multidimensional item pool forms the basis for the development of a multidimensional CAT. With this innovative measurement approach, it will be possible to measure fatigue in RA more precise and with few items.[34]

However, also a limitation of the calibration has to be mentioned. The study is conducted with a relatively small sample compared to other samples used for calibration studies.[25] Therefore, it has to be found out how robust the results of this study are. Possibly, the third dimension (variability of fatigue) will work out better in an analysis with more patients. This study has to be regarded as first, explorative approach to form the basis for the development of a multidimensional CAT for fatigue in RA. The product of this study is an initial calibrated item pool that will undergo further statistical examination in the future process of the multidimensional CAT development.

ACKNOWLEDGEMENTS

We would like to thank the participants in this study: patients from Ziekenhuisgroep Twente (Almelo and Hengelo) and Arthritis Center Twente at Medical Spectrum Twente. Furthermore our thanks to rheumatologists and nurses for supporting the recruitment of participants.

REFERENCES

1. Wolfe F, Hawley DJ, Wilson K. The prevalence and meaning of fatigue in rheumatic disease. J Rheumatol 1996;23(8):1407-1417.
2. Hewlett S, Carr M, Ryan S, Kirwan J, Richards P, Carr A, et al. Outcomes generated by patients with rheumatoid arthritis: how important are they? Musculoskeletal Care 2005;3:131-142.
3. Hewlett S, Chalder T, Choy E, Cramp F, Davis B, Dures E, et al. Fatigue in rheumatoid arthritis: time for a conceptual model. Rheumatology 2011;50:1004-1006. doi: 10.1093/rheumatology/keq282
4. Repping-Wuts H, van Riel P, van Achterberg T. Fatigue in patients with rheumatoid arthritis: what is known and what is needed. Rheumatology 2009;48:207-209.
5. Belza Tack B. Fatigue in rheumatoid arthritis. Conditions, strategies, and consequences. Arthritis Care Res 1990;3:65-70.
6. Hewlett S, Cockshott Z, Byron M, Kitchen K, Tipler S, Pope D, et al. Patients' perceptions of fatigue in rheumatoid arthritis: overwhelming, uncontrollable, ignored. Arthritis Care Res 2005;53:697-702.
7. Repping-Wuts H, Uitterhoeve R, van Riel P, van Achterberg T. Fatigue as experienced by patients with rheumatoid arthritis (RA): a qualitative study. Int J Nurs Stud 2008;45:995-1002.
8. Nikolaus S, Bode C, Taal E, van de Laar MAFJ. New insights into the experience of fatigue among patients with rheumatoid arthritis: a qualitative study. Ann Rheum Dis 2010;69:895-897.
9. De Wit M. From orphan outcome to darling pet – patient participation at OMERACT and the emergence of fatigue as an important outcome for research. Ann Rheum Dis 2010;69(Suppl3):23.
10. Kirwan, J.R., Hewlett, S. Patient perspective: reasons and methods for measuring fatigue in rheumatoid arthritis. J Rheumatol 2007;34:1171-1173.
11. Kirwan JR, Minnock P, Adebajo A et al. Patient perspective: fatigue as a recommended patient centered outcome measure in rheumatoid arthritis. J Rheumatol 2007;34:1174-1177.
12. Minnock P, Kirwan J, Bresnihan B. Fatigue is a reliable, sensitive and unique outcome measure in rheumatoid arthritis. Rheumatology 2009;48:1533-1536.
13. Ware JE Jr, Sherbourne CD. The MOS 36-Item Short-Form Health Survey (SF-36). I. Conceptual framework and item selection. Med Care 1992;30:473-483.

14. Cella D, Yount S, Sorensen M, Chartash E, Sengupta N, Grober J. Validation of the Functional Assessment of Chronic Illness Therapy Fatigue Scale relative to other instrumentation in patients with rheumatoid arthritis. J Rheumatol 2005;32:811-819.
15. McNair D, Lorr M, Droppelman L (1992) Profile of Mood States manual. New York: Multi-health Systems Inc.
16. Belza Tack B. Dimensions and correlates of fatigue in older adults with rheumatoid arthritis (dissertation). San Francisco: University of California; 1991.
17. Hewlett S, Hehir M, Kirwan JR. Measuring fatigue in rheumatoid arthritis: a systematic review of scales in use. Arthritis Rheum 2007;57(3):429-439.
18. Streiner DL, Norman GR. Health measurement scales—a practical guide to their development and use. New York: Oxford University Press; 2003.
19. Fayers PM, Machin D. Quality of life—assessment, analysis and interpretation. Chichester: Wiley; 2000.
20. Yorkston KM, Johnson K, Boesflug E, Skala J, Amtmann D. Communication about the experience of pain and fatigue in disability. Qual Life Res 2010;19:243–251.
21. Nicklin J, Cramp F, Kirwan J, Greenwood R, Urban M, Hewlett S. Measuring fatigue in rheumatoid arthritis: A cross-sectional study to evaluate the Bristol Rheumatoid Arthritis Fatigue Multi-Dimensional questionnaire, visual analogue scales, and numerical rating scales. Arthritis Care Res 2010;62(11):1559-1568.
22. Nicklin J, Cramp F, Kirwan J, Urban M, Hewlett S. Collaboration with patients in the design of patient reported outcome measures: Capturing the experience of fatigue in rheumatoid arthritis. Arthritis Care Res 2010;62:1552-1558.
23. Rose M, Bezjak A. Logistics of collecting patient-reported outcomes (PROs) in clinical practice: an overview and practical examples. Qual Life Res 2009;18:125-136.
24. Wainer H. Computerized adaptive testing: a primer. Lawrence Erlbaum Associate: Hillsdale, New Yersey; 1990.
25. Walter OB. Adaptive tests for measuring anxiety and depression. In: WJ van der Linden, CAW Glas eds. Elements of adaptive testing. New York: Springer; 2010:123-136.
26. Cella D, Yount S, Rothrock N, Gershon R, Cook K, Reeve B, et al. The patient-reported outcomes measurement information system (PROMIS): Progress of an NIH Roadmap Cooperative Group during its first two years. Med Care 2007;45:I3-I11.
27. Fliege H, Becker J, Walter OB, Rose M, Bjorner JB, Klapp BF. Evaluation of a computer-adaptive test for the assessment of depression (D-CAT) in clinical application. Int J Methods Psychiatr Res 2009;18(1):23-36.
28. Becker J, Fliege H, Kocalevent RD, Bjorner JB, Rose M, Walter OB, et al. Functioning and validity of a computerized adaptive test to measure anxiety (A-CAT). Depress Anxiety 2008;25:182-194.

29. Kocalevent RD, Rose M, Becker J, Walter OB, Fliege H, Bjorner JB, et al. An evaluation of patient-reported outcomes found computerized adaptive testing was efficient in assessing stress perception. J Clin Epidemiol 2009;62:278-287.
30. Nikolaus S, Bode C, Taal E, van de Laar MAFJ. Four different patterns of fatigue in rheumatoid arthritis patients: results of a Q-sort study. Rheumatology 2010;49(11):2191-2199.
31. Nikolaus S, Bode C, Taal E, van de Laar MAFJ. Selection of items for a computer-adaptive test to measure fatigue in patients with rheumatoid arthritis – A Delphi approach. Qual Life Res. Published online 31 July 2011. doi: 10.1007/s11136-011-9982-8
32. Nikolaus S, Bode C, Taal E, van de Laar MAFJ. Which dimensions of fatigue should be measured in patients with rheumatoid arthritis? – A Delphi study. Musculoskeletal Care. Published online 11 November 2011. doi: 10.1002/msc.222
33. Nikolaus S, Bode C, Taal E, van de Laar MAFJ (in press). Experts' evaluations of fatigue questionnaires used in rheumatoid arthritis – A Delphi study among patients, nurses and rheumatologists in the Netherlands. Clin Exp Rheumatol.
34. Segall DO. Principles of Multidimensional Adaptive Testing. In: WJ van der Linden, CAW Glas eds. Elements of Adaptive Testing. Statistics for Social and Behavioral Sciences, Springer; 2010:57-75.
35. Guillemin F, Bombardier C, Beaton D. Cross-cultural adaptation of health-related quality of life measures: literature review and proposed guidelines. J Clin Epidemiol 1993;46:1417–.
36. Wolfe F, Michaud K, Pincus T. Development and validation of the Health Assessment Questionnaire II. Arthritis Rheum 2004;50:3296-305.
37. Reckase MD. Linking and scaling. In: MD Reckase, Multidimensional Item Response Theory (pp. 275-310). Statistics for Social and Behavioral Sciences. Springer; 2009.
38. Glas CAW (2010). Software program Multidimensional Item Response Theory (MIRT). http://www.utwente.nl/gw/omd/afdeling/temp_test/mirt-manual.pdf
39. Muraki E. A generalized partial credit model: application of an EM algorithm. Appl Psychol Meas 1992;16:159-176.
40. Embretson SE, Reise SP. Polytomous IRT models. In: Item response theory for psychologists (pp.95-124). Lawrence Erlbaum Associates, Publishers. London; 2000.
41. Lord FM. Applications of item response theory to practical testing problems. Hillsdale, NJ: Erlbaum; 1980.
42. Van Groen MM, ten Klooster PM, Taal E, van de Laar MAFJ, Glas CAW. Applications of the health assessment questionnaire disability index to various rheumatic diseases. Qual Life Res 2010;19:1255-1263.

43. Hambleton RK, Swainathan H, Rogers HJ. Identification of potentially biased test items. In: Fundamentals of item response theory. Sage Publications; 1991.
44. Thomas N, Gan N. Generating multiple imputations for matrix sampling data analyzed with item response models. J Educ Behav Stat 1997;22,425-445.
45. Van der Linden. Constrained adaptive testing with shadow tests. In: WJ van der Linden, CAW Glas eds. Elements of Adaptive Testing. Springer; 2010:31-56.
46. Hewlett S, Dures E, Almeida C. Measures of Fatigue. Arthritis Care Res, 2011;63,S263-S286. Doi 10.1002/acr.20579.

APPENDIX 1: Table A1 Item administration design

Booklet	Severity 19 items	Physical 37 items	Mental 54 items	Conse-quences I 43 items	Conse-quences II 42 items	Change 9 items	Perceived causes I 9 items	Perceived causes II 9 items	Coping I 12 items	Coping II 11 items	Items	N
1	■	■				■	■	■	■	■	106	80
2	■		■			■	■		■		103	80
3	■			■		■	■	■	■	■	112	79
4		■	■			■	■		■		121	85
5		■			■	■		■		■	108	77
6			■	■		■		■		■	126	81
7	■			■	■	■					113	69
Larger dimen-sion	1	2	3	4		5	6		7		245	551
Number excluded items	6	6	2	10		5	8		12		49	

Columns refer to the seven dimensions as used for the analyses. The shaded areas refer to administered items. The blank cells refer to not administered items. The last row refers to the number of items removed due to misfit or unreliability.

APPENDIX 2

Preliminary fatigue item pool in Dutch

Dimension “Severity” (13 items)

Item	Response options	Origin
Kruis het hokje aan dat uw gemiddelde mate van vermoeidheid gedurende de laatste 7 dagen weergeeft.	0 – 10 / 0 geen vermoeidheid, 10 totaal uitgeput	BRAF-MDQ
Gedurende de laatste 7 dagen voelde ik me moe.	1 helemaal niet, 2 een beetje, 3 nogal, 4 in sterke mate	FACIT-F
Gedurende de laatste 7 dagen voelde ik me vermoeid.	1 helemaal niet, 2 een beetje, 3 nogal, 4 in sterke mate	POMS
Gedurende de laatste 7 dagen was vermoeidheid de ergste klacht als gevolg van mijn reuma.	1 - 5 / 1 helemaal niet mee eens, 5 helemaal mee eens	Interviews
Voelde u zich moe gedurende de laatste 7 dagen?	1 altijd, 2 meestal, 3 soms, 4 zelden, 5 nooit	SF-36
Op hoeveel dagen ervoer u vermoeidheid in de afgelopen week (7 dagen)?	0 1 2 3 4 5 6 elke dag	BRAF-MDQ
Gedurende de laatste 7 dagen was ik op … vooral ´s avonds moe.	0 dagen, 1-2 dagen, 3-4 dagen, 4-5 dagen, 5-6 dagen, elke dag	Interviews
Voelde u zich vermoeid gedurende de laatste 7 dagen?	1 altijd, 2 meestal, 3 soms, 4 zelden, 5 nooit	SF-36
Hoe vaak bent u in de afgelopen 7 dagen vermoeid geweest?	1 elke dag, 2 meestal, maar niet elke dag, 3 af en toe, maar de meeste dagen niet, 4 bijna nooit	MAF
Gedurende de laatste 7 dagen was ik moe ook als ik niets buitengewoons had gedaan.	1 altijd, 2 meestal, 3 soms, 4 zelden, 5 nooit	Interviews
Gedurende de laatste 7 dagen werd ik gauw moe als ik met iets bezig was.	1 altijd, 2 meestal, 3 soms, 4 zelden, 5 nooit	Interviews
Hoe lang duurde iedere periode van vermoeidheid gemiddeld in de afgelopen 7 dagen?	1 minder dan een uur, 2 meerdere uren, 3 de hele dag	BRAF-MDQ
Als u denkt aan de laatste 7 dagen; hoe veel uur was u per dag gemiddeld vermoeid?	1 = minder dan een uur, 2 = ongeveer 1-4 uur, 3 = ongeveer 5-8 uur, 4 = meer dan 8 uur	Interviews

Dimension “Impact” (169 items)

Item	Response options	Origin
Had u veel energie gedurende de laatste 7 dagen?	1 altijd, 2 meestal, 3 soms, 4 zelden, 5 nooit	SF-36
Gedurende de laatste 7 dagen voelde ik me actief.	1 altijd, 2 meestal, 3 soms, 4 zelden, 5 nooit	Lee fatigue scale
Gedurende de laatste 7 dagen voelde ik me fit.	1 altijd, 2 meestal, 3 soms, 4 zelden, 5 nooit	CIS
Gedurende de laatste 7 dagen voelde ik me afgemat.	1 altijd, 2 meestal, 3 soms, 4 zelden, 5 nooit	POMS
Had u een gebrek aan LICHAMELIJKE energie door uw vermoeidheid gedurende de laatste 7 dagen?	1 helemaal niet, 2 een beetje, 3 nogal, 4 in sterke mate	BRAF-MDQ
Gedurende de laatste 7 dagen voelde ik me uitgerust.	1 altijd, 2 meestal, 3 soms, 4 zelden, 5 nooit	CIS/MFSI
Gedurende de laatste 7 dagen kon ik de hele dag wel slapen.	1 helemaal niet, 2 een beetje, 3 nogal, 4 in sterke mate	Lee fatigue scale
Gedurende de laatste 7 dagen voelde ik me sloom door de vermoeidheid.	1 helemaal niet, 2 een beetje, 3 nogal, 4 in sterke mate	Interviews
Gedurende de laatste 7 dagen had ik behoefte aan slaap overdag.	1 helemaal niet, 2 een beetje, 3 nogal, 4 in sterke mate	FACIT-F
Gedurende de laatste 7 dagen moest ik eerst even rusten nadat ik iets gedaan had.	1 altijd, 2 meestal, 3 soms, 4 zelden, 5 nooit	Interviews
Gedurende de laatste 7 dagen had ik langere rustperiodes nodig door de vermoeidheid.	1 helemaal niet, 2 een beetje, 3 nogal, 4 in sterke mate	Interviews
Gedurende de laatste 7 dagen moest ik vroeg naar bed om te kunnen blijven functioneren.	1 altijd, 2 meestal, 3 soms, 4 zelden, 5 nooit	Interviews
Hoe vaak had u gedurende de laatste 7 dagen genoeg energie om de dag door te komen zonder te rusten of te slapen?	1 altijd, 2 vaak, 3 soms, 4 zelden, 5 nooit	Interviews
Gedurende de laatste 7 dagen was ik slaperig.	1 altijd, 2 meestal, 3 soms, 4 zelden, 5 nooit	Interviews
Gedurende de laatste 7 dagen moest ik door de vermoeidheid tijdens mijn bezigheden pauzes nemen om weer bij te komen.	1 altijd, 2 meestal, 3 soms, 4 zelden, 5 nooit	Interviews
Gedurende de laatste 7 dagen moest ik meer dan normaal slapen (´s nachts en mogelijk over dag) om te kunnen blijven functioneren.	1 altijd, 2 meestal, 3 soms, 4 zelden, 5 nooit	Interviews

Gedurende de laatste 7 dagen had ik de behoefte om mijn ogen dicht te doen als ik moe was.	1 altijd, 2 vaak, 3 soms, 4 zelden, 5 nooit	Interviews
Gedurende de laatste 7 dagen moest ik regelmatig slapen om te kunnen functioneren (bijvoorbeeld steeds ongeveer op dezelfde tijd naar bed, ongeveer hetzelfde aantal uren slapen).	1 helemaal niet, 2 een beetje, 3 nogal, 4 in sterke mate	Interviews
Gedurende de laatste 7 dagen had ik pijn als ik moe was.	1 altijd, 2 meestal, 3 soms, 4 zelden, 5 nooit	Interviews
Gedurende de laatste 7 dagen voelde ik me lichamelijk in een slechte conditie.	1 altijd, 2 meestal, 3 soms, 4 zelden, 5 nooit	CIS
Gedurende de laatste 7 dagen voelde ik me lichamelijk vermoeid.	1 helemaal niet, 2 een beetje, 3 nogal, 4 in sterke mate	Interviews
Gedurende de laatste 7 dagen voelde ik me krachteloos door mijn vermoeidheid.	1 helemaal niet, 2 een beetje, 3 nogal, 4 in sterke mate	FACIT-F
Door mijn vermoeidheid zat ik gedurende de laatste 7 dagen niet lekker in mijn vel.	1 altijd, 2 meestal, 3 soms, 4 zelden, 5 nooit	Interviews
Had u door de vermoeidheid gedurende de laatste 7 dagen moeite om met andere reumaklachten (bijvoorbeeld pijn of stijfheid) om te gaan?	1 helemaal niet, 2 een beetje, 3 nogal, 4 in sterke mate	Interviews
Gedurende de laatste 7 dagen voelde de vermoeidheid als een lichamelijke last.	1 helemaal niet, 2 een beetje, 3 nogal, 4 in sterke mate	Interviews
Gedurende de laatste 7 dagen had ik minder kracht in mijn spieren door de vermoeidheid.	1 helemaal niet, 2 een beetje, 3 nogal, 4 in sterke mate	FQ/ProF
Gedurende de laatste 7 dagen kon ik moeilijk op mijn benen staan door mijn vermoeidheid.	1 helemaal niet, 2 een beetje, 3 nogal, 4 in sterke mate	Interviews
Gedurende de laatste 7 dagen had ik een zwaar gevoel in mijn hoofd door mijn vermoeidheid.	1 helemaal niet, 2 een beetje, 3 nogal, 4 in sterke mate	Interviews
Gedurende de laatste 7 dagen was ik misselijk van vermoeidheid.	1 helemaal niet, 2 een beetje, 3 nogal, 4 in sterke mate	Interviews
Gedurende de laatste 7 dagen was ik duizelig van vermoeidheid.	1 helemaal niet, 2 een beetje, 3 nogal, 4 in sterke mate	Interviews
Gedurende de laatste 7 dagen had ik coördinatieproblemen (bijvoorbeeld struikelen, naast voorwerpen grijpen, voorwerpen laten vallen) door mijn vermoeidheid.	1 helemaal niet, 2 een beetje, 3 nogal, 4 in sterke mate	Interviews
Gedurende de laatste 7 dagen voelde ik me lusteloos (futloos).	1 helemaal niet, 2 een beetje, 3 nogal, 4 in sterke mate	FACIT
Gedurende de laatste 7 dagen voelde ik me energiek.	1 helemaal niet, 2 een beetje, 3 nogal, 4 in sterke mate	FACIT

Gedurende de laatste 7 dagen had ik door de vermoeidheid minder zin in bepaalde activiteiten.	1 altijd, 2 meestal, 3 soms, 4 zelden, 5 nooit	Interviews
Gedurende de laatste 7 dagen was mijn energie gauw op als ik iets deed.	1 altijd, 2 meestal, 3 soms, 4 zelden, 5 nooit	Interviews
Gedurende de laatste 7 dagen had ik genoeg energie om leuke dingen te doen.	1 altijd, 2 meestal, 3 soms, 4 zelden, 5 nooit	CIS
Had u een gebrek aan MENTALE energie (bijvoorbeeld energie om goed over iets na te denken, plannen te maken, creatief te zijn) door uw vermoeidheid gedurende de laatste 7 dagen?	1 helemaal niet, 2 een beetje, 3 nogal, 4 in sterke mate	BRAF-MDQ
Gedurende de laatste 7 dagen had ik te weinig energie om mijn plannen uit te kunnen voeren.	1 altijd, 2 meestal, 3 soms, 4 zelden, 5 nooit	CIS
Gedurende de laatste 7 dagen was ik nergens in geïnteresseerd door mijn vermoeidheid.	1 altijd, 2 meestal, 3 soms, 4 zelden, 5 nooit	CFS
Gedurende de laatste 7 dagen voelde ik me duf door de vermoeidheid.	1 helemaal niet, 2 een beetje, 3 nogal, 4 in sterke mate	Interviews
Heeft u gedurende de laatste 7 dagen dingen vergeten door uw vermoeidheid?	1 helemaal niet, 2 een beetje, 3 nogal, 4 in sterke mate	BRAF-MDQ
Maakte uw vermoeidheid het gedurende de laatste 7 dagen moeilijk helder na te denken?	1 helemaal niet, 2 een beetje, 3 nogal, 4 in sterke mate	BRAF-MDQ
Maakte uw vermoeidheid het gedurende de laatste 7 dagen moeilijk u te concentreren?	1 helemaal niet, 2 een beetje, 3 nogal, 4 in sterke mate	BRAF-MDQ
Gedurende de laatste 7 dagen was ik door mijn vermoeidheid minder gemotiveerd om iets te doen waar ik bij na moest denken.	1 altijd, 2 meestal, 3 soms, 4 zelden, 5 nooit	FIS
Gedurende de laatste 7 dagen was ik door mijn vermoeidheid minder goed in staat om taken af te maken waar ik bij na moest denken.	1 altijd, 2 meestal, 3 soms, 4 zelden, 5 nooit	FIS
Gedurende de laatste 7 dagen kostte het me veel moeite ergens mijn aandacht bij te houden als ik moe was.	1 helemaal niet, 2 een beetje, 3 nogal, 4 in sterke mate	Interviews
Gedurende de laatste 7 dagen had ik door mijn vermoeidheid moeite beslissingen te nemen.	1 altijd, 2 meestal, 3 soms, 4 zelden, 5 nooit	FIS
Gedurende de laatste 7 dagen was ik te moe om te lezen.	1 altijd, 2 meestal, 3 soms, 4 zelden, 5 nooit	Interviews
Hoe vaak voelde u zich minder alert door uw vermoeidheid?	1 altijd, 2 vaak, 3 soms, 4 zelden, 5 nooit	Interviews
Gedurende de laatste 7 dagen ging denken langzamer dan normaal door mijn vermoeidheid.	1 altijd, 2 meestal, 3 soms, 4 zelden, 5 nooit	FIS

Gedurende de laatste 7 dagen was het door de vermoeidheid moeilijk mijn gedachten op een rij te zetten.	1 altijd, 2 meestal, 3 soms, 4 zelden, 5 nooit	FIS
Gedurende de laatste 7 dagen was ik sneller prikkelbaar door de vermoeidheid.	1 helemaal niet, 2 een beetje, 3 nogal, 4 in sterke mate	Interviews
Gedurende de laatste 7 dagen kostte denken inspanning door mijn vermoeidheid.	1 helemaal niet, 2 een beetje, 3 nogal, 4 in sterke mate	Interviews
Heeft u gedurende de laatste 7 dagen vergissingen gemaakt door uw vermoeidheid?	1 helemaal niet, 2 een beetje, 3 nogal, 4 in sterke mate	BRAF-MDQ
Gedurende de laatste 7 dagen raakte ik in de war door vermoeidheid.	1 vaak, 2 soms, 3 zelden, 4 nooit	MFSI
Gedurende de laatste 7 dagen vond ik het frustrerend als ik te moe was om de dingen te doen die ik wilde doen.	1 helemaal niet, 2 een beetje, 3 nogal, 4 in sterke mate	FACIT-F
Gedurende de laatste 7 dagen was ik boos over mijn vermoeidheid.	1 helemaal niet, 2 een beetje, 3 nogal, 4 in sterke mate	Interviews
Gedurende de laatste 7 dagen werd ik moedeloos van mijn vermoeidheid.	1 helemaal niet, 2 een beetje, 3 nogal, 4 in sterke mate	Interviews
Gedurende de laatste 7 dagen was ik verdrietig door mijn vermoeidheid.	1 helemaal niet, 2 een beetje, 3 nogal, 4 in sterke mate	Interviews
Gedurende de laatste 7 dagen maakte ik me zorgen over mijn vermoeidheid.	1 helemaal niet, 2 een beetje, 3 nogal, 4 in sterke mate	Interviews
Gedurende de laatste 7 dagen raakte ik snel geïrriteerd door mijn vermoeidheid.	1 helemaal niet, 2 een beetje, 3 nogal, 4 in sterke mate	Interviews
Gedurende de laatste 7 dagen zag ik overal tegen op als ik moe was.	1 helemaal niet, 2 een beetje, 3 nogal, 4 in sterke mate	Interviews
Gedurende de laatste 7 dagen kon ik minder van leuke dingen genieten door de vermoeidheid.	1 altijd, 2 meestal, 3 soms, 4 zelden, 5 nooit	Interviews
Gedurende de laatste 7 dagen heeft de vermoeidheid mijn kwaliteit van leven verminderd.	1 helemaal niet, 2 een beetje, 3 nogal, 4 in sterke mate	Interviews
Gedurende de laatste 7 dagen was ik chagrijnig door mijn vermoeidheid.	1 helemaal niet, 2 een beetje, 3 nogal, 4 in sterke mate	Interviews
Gedurende de laatste 7 dagen voelde ik me nutteloos door mijn vermoeidheid.	1 helemaal niet, 2 een beetje, 3 nogal, 4 in sterke mate	Interviews
Had u gedurende de laatste 7 dagen het gevoel door vermoeidheid minder controle te hebben over uw leven?	1 helemaal niet, 2 een beetje, 3 nogal, 4 in sterke mate	BRAF-MDQ
Ik heb door de vermoeidheid de interesse verloren in dingen die ik eerder deed.	1 helemaal niet, 2 een beetje, 3 nogal, 4 in sterke mate	FQ

Gedurende de laatste 7 dagen voelde ik me gespannen door de vermoeidheid.	1 helemaal niet, 2 een beetje, 3 nogal, 4 in sterke mate	Lee fatigue scale
Gedurende de laatste 7 dagen heeft de vermoeidheid mij stress opgeleverd.	1 helemaal niet, 2 een beetje, 3 nogal, 4 in sterke mate	Interviews
Gedurende de laatste 7 dagen voelde ik me machteloos door de vermoeidheid.	1 helemaal niet, 2 een beetje, 3 nogal, 4 in sterke mate	Interviews
Gedurende de laatste 7 dagen ben ik onverschillig geweest door mijn vermoeidheid.	1 helemaal niet, 2 een beetje, 3 nogal, 4 in sterke mate	CFS
Voelde u zich gedurende de laatste 7 dagen somber of neerslachtig door uw vermoeidheid?	1 helemaal niet, 2 een beetje, 3 nogal, 4 in sterke mate	BRAF-MDQ
Door mijn vermoeidheid had ik gedurende de laatste 7 dagen minder zelfvertrouwen.	1 helemaal niet, 2 een beetje, 3 nogal, 4 in sterke mate	Interviews
Gedurende de laatste 7 dagen zat ik over mijn vermoeidheid te piekeren.	1 helemaal niet, 2 een beetje, 3 nogal, 4 in sterke mate	Interviews
Hoe vaak dacht u gedurende de laatste 7 dagen over uw vermoeidheid na?	1 altijd, 2 vaak, 3 soms, 4 zelden, 5 nooit	Interviews
Gedurende de laatste 7 dagen kon ik minder hebben dan gewoonlijk door mijn vermoeidheid.	1 helemaal niet, 2 een beetje, 3 nogal, 4 in sterke mate	Interviews
Gedurende de laatste 7 dagen voelde ik me door de vermoeidheid minder prettig.	1 helemaal niet, 2 een beetje, 3 nogal, 4 in sterke mate	Interviews
Maakte uw vermoeidheid u gedurende de laatste 7 dagen van streek?	1 helemaal niet, 2 een beetje, 3 nogal, 4 in sterke mate	BRAF-MDQ
Gedurende de laatste 7 dagen voelde ik me huilerig door mijn vermoeidheid.	1 helemaal niet, 2 een beetje, 3 nogal, 4 in sterke mate	Interviews
Hoe vaak was u emotioneel vermoeid gedurende de laatste 7 dagen?	1 altijd, 2 vaak, 3 soms, 4 zelden, 5 nooit	Interviews
Gedurende de laatste 7 dagen voelde ik me schuldig als ik te moe was om die dingen te doen die ik gewoonlijk doe.	1 helemaal niet, 2 een beetje, 3 nogal, 4 in sterke mate	Wu fatigue scale
Hoe vaak was u gedurende de laatste 7 dagen te moe om zich gelukkig te voelen?	1 altijd, 2 vaak, 3 soms, 4 zelden, 5 nooit	Interviews
Kruis een hokje aan om het effect dat vermoeidheid de laatste 7 dagen op uw leven heeft gehad weer te geven.	0-10 / 0 = Geen effect, 10 = Heel veel effect	NRS Nicklin et al
Gedurende de laatste 7 dagen moest ik door de vermoeidheid altijd opletten niet te veel te doen.	1 helemaal niet, 2 een beetje, 3 nogal, 4 in sterke mate	Interviews
Gedurende de laatste 7 dagen kon ik activiteiten minder lang volhouden door de vermoeidheid.	1 helemaal niet, 2 een beetje, 3 nogal, 4 in sterke mate	Interviews

Gedurende de laatste 7 dagen kostte het me moeite om met dingen te BEGINNEN omdat ik zo moe was.	1 helemaal niet, 2 een beetje, 3 nogal, 4 in sterke mate	FACIT-F
Gedurende de laatste 7 dagen kon ik minder doen dan ik zou willen door mijn vermoeidheid.	1 helemaal niet, 2 een beetje, 3 nogal, 4 in sterke mate	Interviews
Gedurende de laatste 7 dagen heeft de vermoeidheid mij herhaaldelijk problemen veroorzaakt.	1 helemaal niet, 2 een beetje, 3 nogal, 4 in sterke mate	Interviews
Gedurende de laatste 7 dagen was het door de vermoeidheid moeilijk om de dag door te komen.	1 helemaal niet, 2 een beetje, 3 nogal, 4 in sterke mate	Interviews
Gedurende de laatste 7 dagen voelde vanwege mijn vermoeidheid alles wat ik deed als een inspanning.	1 helemaal niet, 2 een beetje, 3 nogal, 4 in sterke mate	Interviews
Maakte uw vermoeidheid het gedurende de laatste 7 dagen moeilijk uw werk of andere dagelijkse activiteiten uit te voeren?	1 helemaal niet, 2 een beetje, 3 nogal, 4 in sterke mate	BRAF-MDQ
In de afgelopen 7 dagen, hoe heeft vermoeidheid u belemmerd om te werken?	1 helemaal niet, 2 een beetje, 3 nogal, 4 in sterke mate, 5 niet van toepassing	MAF
Gedurende de laatste 7 dagen moest ik door de vermoeidheid meer tijd nemen voor activiteiten.	1 helemaal niet, 2 een beetje, 3 nogal, 4 in sterke mate	Interviews
Gedurende de laatste 7 dagen kon ik bepaalde verantwoordelijkheden niet aan vanwege mijn vermoeidheid (bijv. op mijn werk of binnen het gezin).	1 altijd, 2 meestal, 3 soms, 4 zelden, 5 nooit	Interviews
Gedurende de afgelopen 7 dagen kwam er minder uit mijn handen door mijn vermoeidheid dan ik had gewild.	1 altijd, 2 vaak, 3 soms, 4 zelden, 5 nooit	Interviews
Ik vind dat ik gedurende de laatste 7 dagen weinig deed op een dag door mijn vermoeidheid.	1 - 5 / 1 helemaal niet mee eens, 5 helemaal mee eens	CIS
In de afgelopen 7 dagen, hoe heeft vermoeidheid u belemmerd om huishou-delijke taken en klussen in huis te doen?	1 helemaal niet, 2 een beetje, 3 nogal, 4 in sterke mate	MAF
Gedurende de afgelopen 7 dagen was het moeilijk te klussen (in huis of in de tuin) door mijn vermoeidheid.	1 helemaal niet, 2 een beetje, 3 nogal, 4 in sterke mate	Interviews
In de afgelopen 7 dagen, hoe heeft vermoeidheid u belemmerd om te koken?	1 helemaal niet, 2 een beetje, 3 nogal, 4 in sterke mate, 5 niet van toepassing	MAF
Gedurende de laatste 7 dagen moest ik mijn sociale activiteiten beperken omdat ik moe was.	1 helemaal niet, 2 een beetje, 3 nogal, 4 in sterke mate	FACIT-F
Gedurende de laatste 7 dagen was ik te moe om anderen te helpen.	1 helemaal niet, 2 een beetje, 3 nogal, 4 in sterke mate	Interviews

Gedurende de laatste 7 dagen had ik door mijn vermoeidheid minder zin in sociale activiteiten.	1 altijd, 2 meestal, 3 soms, 4 zelden, 5 nooit	FIS
In de afgelopen 7 dagen, hoe heeft vermoeidheid u belemmerd om op bezoek te gaan bij of tijd door te brengen met vrienden of familie?	1 helemaal niet, 2 een beetje, 3 nogal, 4 in sterke mate	MAF
Gedurende de laatste 7 dagen was ik door mijn vermoeidheid beperkt in mijn contacten met familie, vrienden en kennissen.	1 helemaal niet, 2 een beetje, 3 nogal, 4 in sterke mate	Interviews
Gedurende de laatste 7 dagen had ik door mijn vermoeidheid niet veel sociale contacten buitenshuis.	1 altijd, 2 meestal, 3 soms, 4 zelden, 5 nooit	FIS
Gedurende de laatste 7 dagen maakte ik me zorgen hoe ik door de vermoeidheid op andere mensen overkwam.	1 helemaal niet, 2 een beetje, 3 nogal, 4 in sterke mate	Interviews
Gedurende de laatste 7 dagen was ik sneller moe dan andere mensen van mijn leeftijd.	1 helemaal niet, 2 een beetje, 3 nogal, 4 in sterke mate	Interviews
Gedurende de laatste 7 dagen was ik te moe om voldoende te eten.	1 altijd, 2 meestal, 3 soms, 4 zelden, 5 nooit	FACIT-F
In de afgelopen 7 dagen, hoe heeft vermoeidheid u belemmerd om een bad / douche te nemen of u te wassen?	1 helemaal niet, 2 een beetje, 3 nogal, 4 in sterke mate	MAF
Had u gedurende de laatste 7 dagen moeite met bewegen door vermoeidheid?	1 helemaal niet, 2 een beetje, 3 nogal, 4 in sterke mate	Interviews
In de afgelopen 7 dagen, hoe heeft vermoeidheid u belemmerd om te lopen?	1 helemaal niet, 2 een beetje, 3 nogal, 4 in sterke mate	MAF
In de afgelopen 7 dagen, hoe heeft uw vermoeidheid u belemmerd om te bewegen (bijvoorbeeld wandelen, fietsen, sporten).	1 helemaal niet, 2 een beetje, 3 nogal, 4 in sterke mate	MAF
Door de vermoeidheid had ik gedurende de laatste 7 dagen minder zin om iets te doen dat lichamelijk inspannend voor me is.	1 altijd, 2 meestal, 3 soms, 4 zelden, 5 nooit	FIS
Gedurende de laatste 7 dagen voelde ik me helemaal uitgeput na lichamelijke inspanning (bijvoorbeeld in verband met klussen, sporten).	1 altijd, 2 meestal, 3 soms, 4 zelden, 5 nooit	Interviews
In de afgelopen 7 dagen, hoe heeft vermoeidheid u belemmerd om aan vrijetijds- en ontspannende bezigheden te doen?	1 helemaal niet, 2 een beetje, 3 nogal, 4 in sterke mate	MAF

Heeft u gedurende de laatste 7 dagen vermeden plannen te maken door uw vermoeidheid? Bijvoorbeeld plannen om uit te gaan, of klussen rondom huis of in de tuin te doen?	1 helemaal niet, 2 een beetje, 3 nogal, 4 in sterke mate	BRAF-MDQ
Om gedurende de laatste 7 dagen een ongewone activiteit (bijvoorbeeld winkelen, naar een feest gaan) te kunnen doen moest ik ter compensatie extra rust inplannen.	1 helemaal niet, 2 een beetje, 3 nogal, 4 in sterke mate	Interviews
Gedurende de laatste 7 dagen was het door de vermoeidheid moeilijk van tevoren iets te plannen (omdat ik niet wist hoe het op een ander moment met de vermoeidheid zou gaan).	1 helemaal niet, 2 een beetje, 3 nogal, 4 in sterke mate	Interviews
Gedurende de laatste 7 dagen was vermoeidheid een probleem.	1 helemaal niet, 2 een beetje, 3 nogal, 4 in sterke mate	Interviews
Gedurende de laatste 7 dagen moest ik keuzes maken welke dingen ik wel of niet zou doen omdat ik te moe was om alles te doen.	1 helemaal niet, 2 een beetje, 3 nogal, 4 in sterke mate	Interviews
Gedurende de laatste 7 dagen maakte de vermoeidheid het moeilijk om aan iets nieuws te beginnen.	1 helemaal niet, 2 een beetje, 3 nogal, 4 in sterke mate	Interviews
Gedurende de laatste 7 dagen kostte het me moeite om dingen AF TE MAKEN omdat ik zo moe was.	1 helemaal niet, 2 een beetje, 3 nogal, 4 in sterke mate	FACIT-F
Gedurende de laatste 7 dagen had ik last van vermoeidheid.	1 helemaal niet, 2 een beetje, 3 nogal, 4 in sterke mate	FACIT-F
Gedurende de laatste 7 dagen zag ik er erg tegen op om iets te doen door de vermoeidheid.	1 helemaal niet, 2 een beetje, 3 nogal, 4 in sterke mate	MFI
Had u gedurende de laatste 7 dagen moeite het huis te verlaten door vermoeidheid?	1 helemaal niet, 2 een beetje, 3 nogal, 4 in sterke mate	Interviews
Gedurende de laatste 7 dagen moest ik me dwingen om´smorgens op te staan en iets te doen, door de vermoeidheid.	1 helemaal niet, 2 een beetje, 3 nogal, 4 in sterke mate	Interviews
Gedurende de laatste 7 dagen had ik door de vermoeidheid hulp nodig bij het doen van mijn gebruikelijke activiteiten.	1 helemaal niet, 2 een beetje, 3 nogal, 4 in sterke mate	FACIT-F
Gedurende de laatste 7 dagen moest ik mijn werk verminderen of veranderen vanwege vermoeidheid.	1 helemaal niet, 2 een beetje, 3 nogal, 4 in sterke mate	Interviews
Gedurende de laatste 7 dagen was het moeilijk de auto te besturen door mijn vermoeidheid.	1 helemaal niet, 2 een beetje, 3 nogal, 4 in sterke mate, 5 niet van toepassing	CFS-APQ

Gedurende de laatste 7 dagen had de vermoeidheid een negatieve invloed op mijn relatie met collega's.	1 helemaal niet, 2 een beetje, 3 nogal, 4 in sterke mate, 5 niet van toepassing	Interviews
Gedurende de laatste 7 dagen duurde het door de vermoeidheid langer dingen af te maken.	1 helemaal niet, 2 een beetje, 3 nogal, 4 in sterke mate	Interviews
Gedurende de laatste 7 dagen ervoer ik normale dagelijkse activiteiten als stressvol door mijn vermoeidheid.	1 altijd, 2 meestal, 3 soms, 4 zelden, 5 nooit	FIS
Gedurende de laatste 7 dagen was het moeilijk vrijwilligerswerk te doen door mijn vermoeidheid.	1 helemaal niet, 2 een beetje, 3 nogal, 4 in sterke mate, 5 niet van toepassing	Interviews
Gedurende de laatste 7 dagen was het moeilijk de computer te gebruiken door mijn vermoeidheid.	1 helemaal niet, 2 een beetje, 3 nogal, 4 in sterke mate, 5 niet van toepassing	CFS-APQ
Gedurende de laatste 7 dagen was ik te moe om mijn belangrijkste taken goed te doen.	1 helemaal niet, 2 een beetje, 3 nogal, 4 in sterke mate	Interviews
Gedurende de laatste 7 dagen kon ik mijn taken minder goed doen door mijn vermoeidheid.	1 helemaal niet, 2 een beetje, 3 nogal, 4 in sterke mate	Interviews
In de afgelopen 7 dagen, hoe heeft vermoeidheid u belemmerd om boodschappen / een boodschap te doen?	1 helemaal niet, 2 een beetje, 3 nogal, 4 in sterke mate	MAF
Gedurende de afgelopen 7 dagen was het moeilijk het huishouden (schoonmaken, was, etc.) te doen door mijn vermoeidheid.	1 helemaal niet, 2 een beetje, 3 nogal, 4 in sterke mate	Interviews
Gedurende de laatste 7 dagen had mijn vermoeidheid een negatieve invloed op de relatie met mijn partner en/of andere naasten.	1 helemaal niet, 2 een beetje, 3 nogal, 4 in sterke mate	Interviews
Gedurende de laatste 7 dagen maakte de vermoeidheid het moeilijk om het tempo van anderen vol te houden. Bijvoorbeeld tijdens een dagje uit met vrienden.	1 helemaal niet, 2 een beetje, 3 nogal, 4 in sterke mate	Interviews
Hoe werd u door de vermoeidheid belemmerd om seksueel actief te zijn de afgelopen 7 dagen?	1 helemaal niet, 2 een beetje, 3 nogal, 4 in sterke mate	MAF
Heeft uw vermoeidheid uw sociale leven aangetast gedurende de laatste 7 dagen?	1 helemaal niet, 2 een beetje, 3 nogal, 4 in sterke mate	BRAF-MDQ
Gedurende de laatste 7 dagen was het door de vermoeidheid lastig een gesprek te voeren.	1 helemaal niet, 2 een beetje, 3 nogal, 4 in sterke mate	Interviews
Gedurende de laatste 7 dagen was het door de vermoeidheid lastig in een grote groep mee te praten.	1 helemaal niet, 2 een beetje, 3 nogal, 4 in sterke mate	Interviews
Gedurende de laatste 7 dagen was ik te moe om te praten.	1 vaak, 2 soms, 3 zelden, 4 nooit	Interviews

Gedurende de laatste 7 dagen kon ik door de vermoeidheid moeilijk uit mijn woorden komen.	1 helemaal niet, 2 een beetje, 3 nogal, 4 in sterke mate	Interviews
Maakte uw vermoeidheid het gedurende de laatste 7 dagen moeilijk een bad of douche te nemen?	1 helemaal niet, 2 een beetje, 3 nogal, 4 in sterke mate	BRAF-MDQ
Maakte uw vermoeidheid het gedurende de laatste 7 dagen moeilijk uzelf aan te kleden?	1 helemaal niet, 2 een beetje, 3 nogal, 4 in sterke mate	BRAF-MDQ
Hoe vaak was u gedurende de laatste 7 dagen te moe om een korte wandeling te ondernemen?	1 altijd, 2 meestal, 3 soms, 4 zelden, 5 nooit	Interviews
Door de vermoeidheid kon ik gedurende de laatste 7 dagen lichamelijke inspanning niet lang volhouden.	1 altijd, 2 meestal, 3 soms, 4 zelden, 5 nooit	FIS
Gedurende de laatste 7 dagen kon ik door mijn vermoeidheid niet sporten.	1 altijd, 2 meestal, 3 soms, 4 zelden, 5 nooit	Interviews
Gedurende de laatste 7 dagen was het door mijn vermoeidheid moeilijk om meer dan een verdieping met de trap te gaan.	1 helemaal niet, 2 een beetje, 3 nogal, 4 in sterke mate	CFS-APQ
Door de vermoeidheid was ik gedurende de laatste 7 dagen minder goed in staat om lichamelijk inspannende taken af te maken.	1 altijd, 2 meestal, 3 soms, 4 zelden, 5 nooit	FIS
Gedurende de laatste 7 dagen was het door mijn vermoeidheid moeilijk een uur te zitten.	1 helemaal niet, 2 een beetje, 3 nogal, 4 in sterke mate	Interviews
Gedurende de laatste 7 dagen was ik te moe om mijn hobby uit te oefenen.	1 helemaal niet, 2 een beetje, 3 nogal, 4 in sterke mate	Interviews
Heeft u gedurende de laatste 7 dagen plannen afgezegd vanwege uw vermoeidheid? Bijvoorbeeld plannen om uit te gaan, of klussen rondom huis of in de tuin te doen?	1 helemaal niet, 2 een beetje, 3 nogal, 4 in sterke mate	BRAF-MDQ
Gedurende de laatste 7 dagen was het door de vermoeidheid moeilijk met iets onverwachts (buiten de planning om) om te gaan.	1 helemaal niet, 2 een beetje, 3 nogal, 4 in sterke mate	Interviews
Gedurende de laatste 7 dagen kon ik goed met mijn vermoeidheid omgaan.	1 altijd, 2 meestal, 3 soms, 4 zelden, 5 nooit	Interviews
Ik zou mijn vermoeidheid van de afgelopen 7 dagen als normaal / abnormaal beschrijven.	1 - 10 / 1 = normaal, 10 = abnormaal	Piper scale
Gedurende de laatste 7 dagen moest ik altijd rekening houden met mijn vermoeidheid.	1 helemaal niet, 2 een beetje, 3 nogal, 4 in sterke mate	Interviews
Gedurende de laatste 7 dagen heb ik door mijn vermoeidheid taken moeten uitbesteden.	1 altijd, 2 meestal, 3 soms, 4 zelden, 5 nooit	Interviews
Gedurende de laatste 7 dagen ervoer ik de vermoeidheid als een beperking.	1 helemaal niet, 2 een beetje, 3 nogal, 4 in sterke mate	Interviews

Gedurende de laatste 7 dagen ben ik meer gaan rusten vanwege mijn vermoeidheid (bijvoorbeeld gaan zitten of liggen, een ontspannende bezigheid, maar niet slapen).	1 altijd, 2 meestal, 3 soms, 4 zelden, 5 nooit	Interviews
Gedurende de laatste 7 dagen ben ik meer gaan slapen vanwege mijn vermoeidheid.	1 altijd, 2 meestal, 3 soms, 4 zelden, 5 nooit	Interviews
Gedurende de laatste 7 dagen heb ik mijn activiteiten moeten verdelen door mijn vermoeidheid.	1 helemaal niet, 2 een beetje, 3 nogal, 4 in sterke mate	Interviews
Gedurende de laatste 7 dagen heb ik vanwege mijn vermoeidheid situaties vermeden die mij stress opleveren.	1 altijd, 2 meestal, 3 soms, 4 zelden, 5 nooit	Interviews
Gedurende de laatste 7 dagen werd ik ongeduldig als iets niet lukte door de vermoeidheid.	1 helemaal niet, 2 een beetje, 3 nogal, 4 in sterke mate	Interviews
Gedurende de laatste 7 dagen vond ik mijn vermoeidheid erg vervelend.	1 helemaal niet, 2 een beetje, 3 nogal, 4 in sterke mate	Interviews

Dimension “Variability” (14 items)

Item	**Response options**	**Origin**
Hoe is uw vermoeidheid veranderd in de afgelopen 7 dagen?	1 toegenomen, 2 vermoeidheid is op en neer gegaan, 3 hetzelfde gebleven, 4 afgenomen	MAF
Gedurende de laatste 7 dagen was de duur van mijn vermoeidheid wisselend.	1 ja, 2 nee	Interviews
Was het gedurende de laatste 7 dagen wisselend hoe vaak u vermoeid was?	1 ja, 2 nee	Interviews
Gedurende de laatste 7 dagen was de ernst van mijn vermoeidheid wisselend.	1 ja , 2 nee	Interviews
Sporten veroorzaakt mijn vermoeidheid.	1 ja, 2 nee, 3 weet ik niet	FAI
Werk veroorzaakt mijn vermoeidheid.	1 ja, 2 nee, 3 weet ik niet	FAI
Gewone dagelijkse activiteiten veroorzaken mijn vermoeidheid.	1 ja, 2 nee, 3 weet ik niet	FAI
Stress veroorzaakt mijn vermoeidheid.	1 ja, 2 nee, 3 weet ik niet	FAI
De reumatoïde artritis veroorzaakt mijn vermoeidheid.	1 ja, 2 nee, 3 weet ik niet	FAI

Slechte en / of weinig slaap veroorzaakt mijn vermoeidheid.	1 ja, 2 nee, 3 weet ik niet	Interviews
Drukte / te veel activiteiten veroorzaken mijn vermoeidheid.	1 ja, 2 nee, 3 weet ik niet	Interviews
Pijn en / of andere klachten veroorzaken mijn vermoeidheid.	1 ja, 2 nee, 3 weet ik niet	Interviews
Mijn vermoeidheid wordt mede veroorzaakt door de medicijnen die ik moet gebruiken.	1 ja, 2 nee, 3 weet ik niet	Delphi study
Mijn vermoeidheid wordt door mijn reumamedicijnen veroorzaakt.	1 ja, 2 nee, 3 weet ik niet	Delphi study

Chapter 10

Summary and general discussion

Summary

This thesis explored the meaning of fatigue in RA from the perspective of the patients and provided an initially calibrated multidimensional item bank. First, a systematic review about statistical predictors and impact of fatigue was conducted. Then, the difficulties of measuring treatment effects on fatigue in RA were discussed. Second, an item pool for the measurement of fatigue in RA was developed in four consecutive studies (interview study, Q-sort study, Delphi study and calibration study). The sequence of those studies followed a logical order; starting with qualitative research on the experience of patients and ending up with a statistically calibrated item pool that will be used for the construction of a CAT for fatigue in RA.

The first article included in this thesis (**chapter 2**) gives an overview about the state of the art of research on statistical predictors of fatigue in RA and the impact of fatigue on other outcomes.[1] Furthermore it describes differences and similarities in fatigue between different groups of patients. Studies found statistical predictors of fatigue in illness-related aspects (e.g. pain), physical functioning (e.g. disability), cognitive/emotional functioning (e.g. depression) and social aspects (e.g. negative interpersonal events). Additionally, female gender was associated with higher levels of fatigue. Remarkably, inflammatory activity (e.g. ESR, DAS28), showed an unclear relationship with fatigue in RA. Fatigue statistically predicted illness-related aspects (e.g. morning stiffness), physical functioning (e.g. physical quality of life), cognitive/emotional aspects (e.g. psychological distress) and social aspects (e.g. work ability). Patients with RA reported higher severity of fatigue than healthy controls. Fibromyalgia patients, however, reported worse levels of fatigue than other patient groups. The most evidence for a relation between fatigue and other variables was found regarding pain, depression and physical functioning. Many cross-sectional and also longitudinal studies reported that they statistically predicted fatigue.
Although several relations between fatigue and other variables were demonstrated, hardly conclusions about causality could be drawn based on the reviewed literature. Most studies were cross-sectional and not all longitudinal studies controlled for baseline fatigue. Prospective longitudinal research with appropriate designs and analyses is needed to learn about the multicausal relationships of fatigue in RA and other variables.

Chapter 3 discusses the disappointing effects of biotherapies on fatigue in RA found in a recent meta-analysis.[2] Two approaches to explain its results were proposed. First, it is possible that biotherapies are inefficacious to reduce fatigue in RA. We lack sufficient knowledge of the causes and pathways of fatigue in RA. It is not yet clear how

inflammatory processes influence fatigue. Even if biotherapies have effect on these inflammatory processes, fatigue can remain because it is also linked to psychological processes. A second explanation is related to the measurement of fatigue. When we are not able to measure fatigue appropriately, we cannot evaluate intervention effects on fatigue adequately. The problems related to the measurement of fatigue were discussed in detail also in the following chapters.

The central question of **Chapter 4** is how patients with RA experience fatigue. In depth-interviews were conducted to systematically describe the experience of fatigue in patients with all levels of fatigue.[3] Experiences of fatigue were compared between groups of patients (as men vs. women and younger vs. older patients). As documented in previous studies, fatigue was experienced as multidimensional, annoying symptom with far-reaching consequences. However, our study also revealed inter- and intra-individual differences: emotions, consequences and management of fatigue differed with age and gender. Some of these differences were related to the number of daily roles that patients had to fulfil. Especially younger women with multiple daily roles were vulnerable to the negative impact of fatigue. Many patients reported variation in severity, frequency and duration of fatigue, with some even reporting different forms of fatigue. Patients also mentioned positive aspects of fatigue. In this qualitative study we described the meaning of fatigue. In the next step we investigated whether there are certain groups of patients who experience fatigue in a different way, and which patient characteristics are related to such different patterns.

As described in **chapter 5,** we used Q-methodology to find patterns of fatigue experience shared by a certain group of patients.[4] Statements about fatigue were sorted by the participants to reflect their agreement with the statements. Person factor analysis was performed on the Q-sorts to factor patients across items (instead of items across persons as in usual factor analysis). Each factor represented a group of patients with a common perspective on the experience of fatigue. Physical, psychological and social characteristics of patients were associated with those experiences. The groups identified were labelled as: “Little impact of fatigue”, “Good coping and bad sleep”, “Search for balance”, and “High distress”. This analysis provided insight into different perspectives of the experience of fatigue in groups of patients who share certain characteristics. The study showed large inter-individual differences in the experience of fatigue in RA; not all patients really suffered from fatigue, but for many RA patients fatigue was an important and not treated complaint. Moreover, disease activity seemed to be less important in relation to fatigue than daily roles and tasks, what underlines the importance of social aspects for the

experience of fatigue. The Q-sort study provided further information on the different dimensions of fatigue that are important for RA patients; and also on the importance of the individual items to assess certain dimensions.

In **chapter 6, 7 and 8** the development and qualitative evaluation of the item pool for the comprehensive assessment of fatigue in RA is reported. In **chapter 6,** the process of the selection of items for the CAT fatigue RA by using a Delphi approach, is described in detail.[5] For the development of a CAT, a calibrated item bank is needed. Therefore a comprehensive pool of potential items was developed to cover all dimensions of fatigue experience. All translated items of the BRAF-MDQ and all items from those questionnaires with reasonable validity in RA were included. Additional items were constructed based on other questionnaires and material from the interviews and the Q-sort study. In a Delphi procedure, rheumatologists, nurses and patients were asked to judge the selected dimensions and items for relevancy. They were also invited to nominate other dimensions of fatigue and items to be included in a comprehensive measure of fatigue. These experts could also indicate if items reflected other aspects of RA (e.g. items confounded with disability or inflammation) and should therefore not be included. Items were selected for the development of the item bank according to a described decision rule. Remaining items were adjusted based on participants´ comments and re-evaluated in a second round. The procedure stopped when all items were selected or rejected. Items were excluded due to unclear or too extreme formulations or because they overlapped with other aspects than fatigue (e.g. disability). After two rounds 245 items, spread among 12 dimensions, were selected. Many excluded items originated form interview material, indicating that patients´ stories are not enough to construct good items and further evaluation such as a Delphi study is definitely needed. The Delphi procedure turned out to be a beneficial approach to select items for measuring fatigue in RA. Nevertheless, a bottleneck was that patients and professionals tended to evaluate some items in a normative way. In conclusion, the study revealed a qualitatively evaluated pool of potential items for the item bank development.

Chapter 7 focuses on the question which dimensions of fatigue should be used for a comprehensive measurement of fatigue in RA.[6] The experts evaluated our dimensions for the multi-dimensional assessment of fatigue in RA as adequate. All of the 12 proposed dimensions were rated as relevant by the participants; severity, frequency, duration, changes in fatigue, perceived causes of fatigue, energy, sleep/rest, body feeling, cognition/concentration, coping, negative emotions/mood, and consequences. No additional dimensions emerged from the participants' comments, suggesting that our item pool satisfactory covers the phenomenon of fatigue in RA.

Results also provided insight into the content validity of items of frequently used traditional questionnaires (SF-36 subscale vitality, FACIT-F, POMS subscale fatigue/inertia, MAF) and the BRAF-MDQ as reported in **chapter 8**.[7] Only 40% of the items from traditional questionnaires were clear and adequate in their original style. However, most of the other items could be selected after adaptations and were included in our item pool.

Chapter 9 describes the calibration of the previously evaluated, face and content valid item pool to measure fatigue in RA.[8] The fit of the items with the underlying dimensions was assessed with item response theory (IRT) and the dimensional structure of the item pool was examined by factor analysis. It was not feasible to let each patient fill in all 245 items of the item pool, so we used an item administration design to construct seven different questionnaire versions. 551 patients with RA filled in one of those versions. IRT analysis using the generalized partial credit model (GPCM) was conducted for each dimension of fatigue. Thereby 49 items showed insufficient item characteristics (poor discriminative ability and/or model misfit) and were removed. Exploratory and confirmatory factor analysis on the 196 remaining items revealed three dimensions of fatigue: severity, impact and variability of fatigue. This study provided an initially calibrated multidimensional item bank and has shown which dimensions and items that came forward from previous studies are important for the development of a multidimensional CAT for fatigue in RA.

General discussion and future directions

This thesis provides insights into the meaning of fatigue in RA and offers an initially calibrated, multidimensional item pool to measure fatigue in patients with RA. For its development, several steps were undertaken to ensure the inclusion of the perspective of patients and also the view of professionals.

From patient perspective to measurement

Knowledge about causal relationships between fatigue and other variables did not exist when we started this project. We summarized variables that might have a potential causal relation with fatigue in RA, but hardly conclusions about causes and consequences could be drawn due to the cross-sectional nature of the reviewed studies. It is even unclear whether there are significant relations between fatigue and inflammatory processes in RA. This lack of information underlines the relevance for understanding fatigue from the patients´ perspective. As so little is known about a subjective phenomenon, the

experience of patients is the logical starting point for further studies. What patients tell us about their perspective on fatigue can inform the development of further research questions. Besides this, a precondition for the examination of the complex of causes and consequences of fatigue is appropriate measurement.

This thesis is characterized by the intention to combine the perspective of patients with advanced statistical methods. Both approaches were used for the development of an item pool to measure fatigue in RA.

Especially in the calibration study of the qualitatively evaluated item pool (chapter 9), it was a challenge to balance information from the patient perspective and results from statistical analyses. In our case, the dimension "variability of fatigue" had poorer statistical outcomes than the other two dimensions. However, this variability of fatigue was clearly brought up by patients. In this dilemma, a possibility could be to save this important aspect of the patient perspective by giving this dimension a prominent place in the CAT. It is possible to place restrictions on the mechanisms of the CAT in its future development so that items from this dimension will be chosen (and provided to patients) in the measurement procedure although other items might have somewhat better item characteristics and would therefore be selected with a higher likelihood. Another option would be to use a short traditional questionnaire with questions about the variability of fatigue or to ask about this aspect of fatigue experience in the consultations with the rheumatologist or other health professional. Nevertheless, the tension between incorporating the patients´ perspective and statistical analyses will remain in all attempts of patient centred research. Modern psychometric methods as IRT include the danger of losing face validity of items and the danger of excluding items although they are needed for an adequate reflection of the measured construct.[9] However, it has to be noted that the calibration study statistically supported the multidimensionality of fatigue as reported by patients.

Including the patients´ perspective into research is a hot topic at the moment.[9,10] At first glance it seems simple but patient perspective and statistics are about different philosophies. The perspective of patients alone is not enough and only measurement statistic is not enough too. One without the other makes no sense, but one with another is not perfect either. When experts develop measurement instruments, the patient perspective is missing. However, it is difficult to translate the perspective of patients into a good scale. In this thesis, each study was thoroughly conducted. Several logical steps were undertaken to develop the item pool, but there was still space for interpretation of the authors between the separate studies.[11] The translation between different methods is not always easy and possibly by applying statistical methods, a part of the patients'

perspective gets lost again. There are no guidelines how to come from interviews to items or according to what criteria expert evaluations should be included.
Measurement of fatigue in RA, even as science in itself, is an iterative process. There will never be the final solution. Further validation will take place to check whether items adequately cover everything patients reported. Adaptation to new circumstances is needed since science comes up with new insights and societies are changing, e.g. by aging populations. Over and over again, new medication is developed that possibly will influence the manifestation of fatigue in RA. These developments and changes may make it necessary to adjust measurement instruments accordingly. At least it reminds us that the validation of measurement instruments for fatigue should be re-checked and updated regularly.

Limitations of the thesis

The strength of our approach to develop the item pool is the thorough process from patients´ experience to measurement. Nevertheless, there are some limitations of this thesis.
First, the aim of the systematic review (chapter 2) could not be reached in a satisfactory way. We intended to gain more insight into possible causes and consequences of fatigue in RA by investigating what is reported about possible causes of fatigue in RA, what is reported about possible consequences of fatigue in RA and what is reported about differences in the level of fatigue between different groups, e.g. different diagnoses, patients and healthy controls. We provided an overview about variables that were associated with fatigue in RA and potential causal relations. However, most of the reviewed studies were cross-sectional and did not allow direct conclusions about causal relationships.
Second, we conducted a broad and detailed Delphi study (chapter 6-8) and although this was the best approach for our purposes, it had some bottlenecks. It was very time consuming for participants and researchers and it turned out that participants tended to evaluate some items in a normative way. Possibly, items were evaluated as not appropriate because participants thought that an item was too confronting or that patients gave in fact an answer to the item, reflecting their own fatigue. Cognitive interviews would have been a possibility to gain more clarity about this aspect, but the item pool contained far too many items for applying cognitive interviews. Many items are needed for an item bank to represent the whole measurement range. For the same reason, we started as heterogeneous as possible, including many items and taking into account that some items would be excluded in the Delphi process. In the final

construction of the CAT, we will include cognitive interviews with patients as final check for adequate formulation of the items. An important advantage of the Delphi study, compared to cognitive interviews with patients, was however that also the opinions of nurses and rheumatologists were included. The view of professionals is also essential because they have to accept and use measurement instruments for fatigue.

Excluded items came from both sources; from interview material and also from existing questionnaires. However, the larger number was from interview material. It was a special challenge in the development of the item pool to end up with appropriate items that we developed from our interview material as no guidelines for the process of constructing items from interview material are available at all.

Another limitation of the results presented in this thesis lies in the sample size of the calibration study (chapter 9). For our purposes and the item administration design the number of participants was rather small. For IRT analysis usually larger samples are advised. Therefore the results of the calibration study have to be regarded as preliminary results. In the consecutive project about the further development of the CAT fatigue, more data will be gained and consecutive analyses will be carried out. Our current item pool is a valuable starting point for this future project.

When comparing the three dimensions of our initial item pool (severity, impact, variability of fatigue) with the other fatigue instrument that has been developed from the perspective of patients with RA, the BRAF-MDQ,[12] it turns out that the BRAF-MDQ has different dimensions. Physical fatigue, impact of fatigue on daily living, cognitions and emotions related to fatigue are aspects that are also included in our item pool, but did not come forward from factor analysis in the same way as in the BRAF. With the CAT we intent to develop, it will be possible to measure fatigue more precisely and with few items. However, when choosing a measurement instrument for fatigue, it is important to bear in mind what is aimed to achieve. For a profile of fatigue in an individual patient, probably the BRAF-MDQ is the more appropriate choice as it shows on which area the consequences of fatigue are experienced. For efficient measurement in daily clinical practice and in trials the CAT would be the more favourable instrument to assess fatigue multidimensional and more precise.

Future directions

As pointed out in the introduction, it is not yet understood whether the experience of fatigue in RA is different from other diseases. In many other diseases as MS and cancer, fatigue is one of the main symptoms either. However it might be caused by different processes. Given the absence of evidence for a general, universal fatigue experience

across different chronic, physical conditions, we assume that it is reasonable to develop a fatigue CAT, which is specific for RA.
It has to be examined in the future how fatigue in RA overlaps with fatigue in other diseases. A first step would be a systematic comparison of qualitative research on the experience of fatigue in different rheumatic conditions and other diseases.
Furthermore knowledge about causes and consequences of fatigue in RA is urgently needed. For the investigation of causality, studies are required that use an adequate design. Prospective, longitudinal approaches including control for baseline levels of fatigue would be a good starting point.
With more insight about the causal pathways of fatigue in RA, it could be compared with those of other diseases. Not until then, it will be possible to make a clear statement about the question whether a specific measurement instrument for fatigue in RA is needed. Possibly the CAT can be validated for other (rheumatic) conditions in the future. Independent of the discussion about a need for disease specific measurement instruments, it would also be valuable to validate the CAT in different countries.

In sum, this thesis gained more insight into what fatigue in RA is. By now it is clear that many patients experience fatigue in RA as different from usual fatigue. We know that fatigue has a multidimensional character. Several studies come to relatively consistent results regarding those dimensions. Moreover research has provided ideas about potential factors associated with fatigue in RA and where to search for its causes and pathways. Future studies have to provide more evidence for those. Furthermore, it has to be found out how fatigue in RA can be treated or even prevented. Research is needed to show how patient-education can be provided and how patients having serious problems with fatigue can be adequately supported. Therefore a good measurement instrument is needed. This thesis adds an additional step in this challenging task and provides the thoroughly developed basis for the consecutive construction of a CAT for fatigue in RA.

REFERENCES

1. Nikolaus S, Bode C, Taal E, vd Laar MAFJ (in preparation for submission). Fatigue and factors related to fatigue in rheumatoid arhritis: a systematic review.
2. Nikolaus S, van de Laar MAFJ. Measuring fatigue in rheumatoid arthritis. Nat Rev Rheumatol 2011 Aug 30;7:562-4.
3. Nikolaus S, Bode C, Taal E, van de Laar MAFJ. New insights into the experience of fatigue among patients with rheumatoid arthritis: A qualitative study. Ann Rheum Dis 2010;69(5), 895–897.
4. Nikolaus S, Bode C, Taal E, van de Laar MAFJ. Four different patterns of fatigue in rheumatoid arthritis patients: results of a Q-sort study. Rheumatology 2010;49(11), 2191-2199.
5. Nikolaus S, Bode C, Taal E, van de Laar MAFJ. Selection of items for a computer-adaptive test to measure fatigue in patients with rheumatoid arthritis – A Delphi approach. Qual Life Res. Published online 31 July 2011. doi: 10.1007/s11136-011-9982-8
6. Nikolaus S, Bode C, Taal E, van de Laar MAFJ. Which dimensions of fatigue should be measured in patients with rheumatoid arthritis? – A Delphi study. Musculoskeletal Care. Published online 11 November 2011. doi: 10.1002/msc.222
7. Nikolaus S, Bode C, Taal E, van de Laar MAFJ (in press). Experts' evaluations of fatigue questionnaires used in rheumatoid arthritis – A Delphi study among patients, nurses and rheumatologists in the Netherlands. Clin Exp Rheumatol.
8. Nikolaus S, Bode C, Taal E, van de Laar MAFJ (in preparation for submission). Calibration of a multidimensional item bank to measure fatigue in patients with rheumatoid arthritis.
9. Fries JF. New instruments for assessing disability: not quite ready for prime time. Arthritis Rheum 2004;50:3064-3067.
10. Quest E et al. Patient perspective. J Rheumatol 2003;30:884-885.
11. Bode, C. Qualitative interviews and IRT for developing PRO fatigue. Paper presented at the Annual Meeting of the American College for Rheumatology 5-9 november 2011, Chicago.
12. Nicklin J, Cramp F, Kirwan J, Greenwood R, Urban M, Hewlett S. Measuring fatigue in rheumatoid arthritis: A cross-sectional study to evaluate the Bristol Rheumatoid Arthritis Fatigue Multi-Dimensional questionnaire, visual analogue scales, and numerical rating scales. Arthritis Care Res 2010;62:1559-1568.

Samenvatting

(Dutch summary)

De inhoud van dit proefschrift brengt de betekenis van vermoeidheid bij RA in kaart vanuit het patiënten perspectief en levert een gekalibreerde multidimensionele itembank op.
In het kort gezegd houdt het proefschrift de volgende studies in.
Ten eerste hebben wij een systematische review naar statistische voorspellers van vermoeidheid en de invloed van vermoeidheid op andere uitkomstmaten uitgevoerd. Daarna werden de moeilijkheden van het meten van behandeleffecten op vermoeidheid bij RA bediscussieerd.
Ten tweede hebben wij een item pool voor het meten van vermoeidheid bij RA ontwikkeld door middel van vier op elkaar volgende studies (interview studie, Q-sort studie, Delphi studie en kalibratiestudie). De volgorde van deze studies had een logische opbouw; we zijn begonnen met kwalitatief onderzoek naar de ervaring van patiënten en zijn geëindigd met een statistisch gekalibreerde itempool. Deze zullen wij verder gebruiken om een computer adaptieve test (CAT) voor vermoeidheid bij RA te construeren. Hieronder worden de studies kort samengevat.

Het eerste artikel in het proefschrift (**hoofdstuk 2**) geeft een overzicht over de actuele stand van zaken wat betreft onderzoek naar mogelijke oorzaken en consequenties van vermoeidheid bij RA. Daarnaast hebben wij verschillen en overeenkomsten in vermoeidheid tussen patiënten met verschillende diagnoses beschreven. Studies toonden voorspellers aan voor vermoeidheid onder ziekte gerelateerde aspecten (bijv. pijn), fysiek functioneren (bijv. lichamelijke beperkingen), cognitief/emotioneel functioneren (bijv. depressie) en sociale aspecten (bijv. negatieve interpersoonlijke gebeurtenissen). Bij vrouwen werden over het algemeen hogere vermoeidheidsniveaus gevonden dan bij mannen. Met betrekking tot de relatie tussen ontstekingsactiviteit (bijv. BSE of/en DAS28) en vermoeidheid bij RA kwamen tegenstrijdige resultaten naar voren. Een hogere mate van vermoeidheid was geassocieerd met ziektegerelateerde aspecten (bijv. ochtendstijfheid), fysiek functioneren (bijv. fysieke kwaliteit van leven), cognitieve/emotionele aspecten (bijv. psychologische stress) en sociale aspecten (bijv. werkvermogen). Patiënten met RA rapporteerden een hogere mate van vermoeidheid dan gezonde mensen. Fibromyalgie patiënten daarentegen rapporteerden een hogere mate van vermoeidheid dan andere patiëntengroepen. Het viel op dat de meeste aanwijzingen voor een relatie tussen vermoeidheid en andere variabelen gevonden werden voor pijn, depressie en fysiek functioneren. In veel van de cross-sectionele en ook van de longitudinale studies bleken deze drie variabelen een belangrijke rol te spelen bij het voorspellen van vermoeidheid bij RA.
Hoewel verschillende associaties tussen vermoeidheid en andere variabelen gevonden werden, konden wij op basis van de gereviewde literatuur nauwelijks betrouwbare

conclusies over causaliteit trekken. Het merendeel van de studies was cross-sectioneel en niet alle longitudinale studies controleerden voor vermoeidheid op baseline. Prospectief longitudinaal onderzoek met adequate statistische analyses is nodig om meer over de multicausale relaties tussen vermoeidheid bij RA en andere variabelen aan te tonen.

Hoofdstuk 3 bediscussieert de teleurstellende effecten van biotherapieën op vermoeidheid bij RA zoals aangetoond door een recente meta-analyse. Om deze resultaten te verklaren hebben wij twee scenario's geschetst. Het eerste scenario suggereert dat biotherapieën niet effectief zijn voor het verminderen van vermoeidheid bij RA. We weten nog niet genoeg over de oorzaken en consequenties van vermoeidheid bij RA. Het is nog niet duidelijk hoe vermoeidheid door ontstekingsprocessen beïnvloed wordt. Zelfs als biotherapie effect op vermoeidheid heeft, kan vermoeidheid blijven bestaan omdat deze ook gerelateerd is aan psychologische processen. Het tweede scenario gaat over het meten van vermoeidheid. Als we niet in staat zijn vermoeidheid adequaat te meten, kunnen we interventie-effecten op vermoeidheid ook niet goed evalueren. De problemen rondom het meten van vermoeidheid worden ook in de volgende hoofdstukken besproken.

De centrale vraag van **hoofdstuk 4** is hoe patiënten met RA vermoeidheid ervaren. We hebben diepte-interviews uitgevoerd om de ervaring van vermoeidheid van patiënten met verschillende niveaus van vermoeidheid systematisch te beschrijven. Daarbij hebben wij ook ervaringen van vermoeidheid vergeleken tussen groepen van patiënten (mannen vs. vrouwen en jongere vs. oudere patiënten). Zoals gerapporteerd in voorafgaande studies, werd vermoeidheid ervaren als een multidimensioneel, lastig symptoom met verreikende consequenties. Niettemin liet onze studie ook inter- en intra-individuele verschillen zien: emoties, consequenties en omgang met vermoeidheid varieerden voor leeftijd en geslacht. Sommige van deze verschillen waren gerelateerd aan het aantal dagelijkse rollen van patiënten. Vooral jongere vrouwen met meerdere dagelijkse rollen waren kwetsbaar voor de negatieve invloed van vermoeidheid. Veel patiënten rapporteerden variaties in ernst, frequentie en duur van vermoeidheid, waarbij sommigen zelfs verschillende vormen van vermoeidheid rapporteerden. Een interessant resultaat was dat patiënten ook positieve aspecten van vermoeidheid noemden, bijvoorbeeld bewustere keuzes in het leven maken en leren de (gedwongen) rust te waarderen. Na de betekenis van vermoeidheid te hebben beschreven, wilden wij onderzoeken of er bepaalde groepen van patiënten bestaan die vermoeidheid op een vergelijkbare manier ervaren en welke kenmerken deze groepen gemeen hebben.

In **hoofdstuk 5** wordt beschreven hoe we Q-methodologie gebruikt hebben om patronen in de ervaring van vermoeidheid bij RA te vinden. Deelnemers hebben stellingen over vermoeidheid gesorteerd om aan te geven in hoeverre ze op hen van toepassing waren. Met behulp van personen-factoranalyse op de Q-sorts hebben wij patiënten over items verdeeld (in plaats van items over personen zoals in gebruikelijke factoranalyse). Elke factor stond voor een groep van patiënten met een vergelijkbaar perspectief op de ervaring van vermoeidheid. Fysieke, psychologische en sociale kenmerken van patiënten leken met deze ervaringen samen te hangen. We hebben de groepen benoemd als: "Weinig invloed van vermoeidheid", "Goede omgang en slechte slaap", "Zoek naar balans", en "Veel distress". Deze analyse gaf inzicht in verschillende perspectieven op de ervaring van vermoeidheid in groepen van patiënten en welke kenmerken deze patiëntengroepen gemeen hebben. De studie liet grote inter-individuele verschillen zien in de ervaring van vermoeidheid bij RA; niet alle patiënten hadden last van vermoeidheid, maar voor veel RA-patiënten was vermoeidheid een belangrijk en een (nog) niet behandelde klacht. Bovendien bleek ziekteactiviteit minder belangrijk in relatie tot vermoeidheid dan dagelijkse rollen en taken, wat de relevantie van sociale aspecten voor de ervaring van vermoeidheid onderstreept. De Q-sort studie leverde ook informatie op over welke dimensies van vermoeidheid belangrijk zijn voor RA-patiënten en hoe goed individuele items deze dimensies meten.

Hoofdstuk 6, 7 en 8 gaan over onze aansluitende Delphi-studie waarin wij aan experts (patiënten, verpleegkundigen/reumaconsulenten en reumatologen) gevraagd hebben om items voor het meten van vermoeidheid op hun geschiktheid te beoordelen.
Voor de constructie van een CAT is een gekalibreerde itembank nodig. Om een dergelijk itembank voor het uitgebreid kunnen meten van vermoeidheid bij RA te ontwikkelen, hebben wij een omvangrijke verzameling van potentiële items gemaakt om alle dimensies van de vermoeidheidservaring mee te nemen. Alle vertaalde items van de BRAF-MDQ en alle items van de vragenlijsten met acceptabele validiteit voor RA werden geïncludeerd. We hebben ook aanvullende items geconstrueerd vanuit ons interviewmateriaal en op basis van de Q-sort studie en andere vragenlijsten. In **hoofdstuk 6** wordt het proces van de selectie van items voor de CAT gedetailleerd beschreven. Voor de inhoudelijke evaluatie van onze verzameling van potentiële vermoeidheidsitems wilden wij zowel het perspectief van patiënten als ook het perspectief van professionals meenemen. Daarom hebben wij in een Delphi-procedure aan reumatologen, verpleegkundigen/reumaconsulenten en patiënten gevraagd om de geselecteerde items en dimensies op relevantie te beoordelen. Deze experts werden ook uitgenodigd om andere dimensies van vermoeidheid en items te nomineren om mee te nemen in een omvangrijke meting van vermoeidheid. Zij konden

ook aangeven dat items andere aspecten van RA weergaven (bijv. items die mogelijk beperkingen of ontstekingen reflecteerden) en daarom niet meegenomen zouden moeten worden. We hebben gebruik gemaakt van een vooraf beschreven beslisregel om items voor de ontwikkeling van de itembank te selecteren. De overige items hebben wij aangepast op basis van de commentaren van de deelnemers en in een tweede ronde ter herbeoordeling aangeboden. De procedure eindigde als alle items geselecteerd of geëxcludeerd waren. Redenen voor exclusie waren onduidelijke of te extreme formuleringen of overlap met andere aspecten dan vermoeidheid (bijv. lichamelijke beperkingen). Na twee rondes konden wij 245 items selecteren die verdeeld waren over 12 dimensies. De studie leverde een kwalitatief geëvalueerde itempool op met potentiële items voor de ontwikkeling van de itembank.

De focus van **hoofdstuk 7** ligt op de vraag welke dimensies van vermoeidheid gebruikt zouden moeten worden voor een uitgebreide meting van vermoeidheid bij RA. De experts vonden onze dimensies voor het multidimensioneel meten van vermoeidheid bij RA geschikt. Alle van de 12 voorgestelde dimensies werden door de deelnemers als relevant beoordeeld; ernst, frequentie, duur, verandering in vermoeidheid, ervaren oorzaken van vermoeidheid, energie, slaap/rust, lichaamsgevoel, cognitie/concentratie, omgang, negatieve emoties/stemming en consequenties. Er kwamen geen aanvullende dimensies uit de commentaren van de deelnemers naar voren, wat erop duidt dat ons itempool het fenomeen van vermoeidheid in RA goed afdekt.

De resultaten gaven ook inzicht in de inhoudsvaliditeit van items uit frequent gebruikte traditionele vragenlijsten (SF-36 subschaal vitaliteit, FACIT-F, POMS subschaal fatigue /inertia, MAF) en de BRAF-MDQ zoals gerapporteerd in **hoofdstuk 8**. Slechts 40% van de items van traditionele vragenlijsten was duidelijk en adequaat in hun oorspronkelijke vorm. Echter, de meeste van de overige items werden na aanpassing geselecteerd en konden in ons itempool worden opgenomen.

Hoofdstuk 9 beschrijft de kalibratie van de voorafgaande geëvalueerde, face- en inhoudsvalide itempool voor het meten van vermoeidheid bij RA. We hebben de itemfit met de onderliggende dimensies geëvalueerd door middel van item response theorie (IRT) en de dimensionele structuur van de itempool onderzocht met factoranalyse. Het was niet mogelijk dat elke patiënt alle 245 items van de itempool invulde. Daarom hebben wij een item-afnameschema ontwikkeld om zeven verschillende vragenlijstversies te construeren. Daarbij hebben wij gelet op voldoende overlap tussen de versies voor de analyses. Vervolgens vulden 551 patiënten met RA een van deze versies in. Met behulp van IRT hebben wij de data voor elke dimensie van vermoeidheid geanalyseerd. Daarvoor hebben wij gebruik gemaakt van het generalized partial credit model (GPCM). Het bleek dat 49

items onvoldoende itemkarakteristieken hadden (slecht discriminatief vermogen en/of model misfit). Deze items hebben wij voor de verdere analyse van de itempool verwijderd. Exploratieve en confirmatorische factoranalyse van de 196 overige items brachten drie dimensies naar voren: ernst, impact en variabiliteit van vermoeidheid. Deze studie leverde een voorlopig gekalibreerde multidimensionele itembank op en liet zien welke dimensies en items, die naar voren waren gekomen uit voorafgaande studies, belangrijk zijn voor de ontwikkeling van een multidimensionele CAT voor vermoeidheid bij RA.

Dit proefschrift verschaft inzichten in de betekenis van vermoeidheid bij RA en levert een voorlopig gekalibreerde, multidimensionele itempool op voor het meten van vermoeidheid bij patiënten met RA. Voor zijn ontwikkeling hebben wij verschillende stappen uitgevoerd om ervoor te zorgen dat zowel het perspectief van patiënten als de zienswijze van professionals meegenomen werd.

In de discussie gaan we in op de spanning die tussen de inclusie van het patiëntenperspectief en statistische analyses kan optreden. Moderne psychometrische methoden zoals IRT lopen het risico face-validiteit van items te verliezen en items te excluderen terwijl deze nodig zijn voor een adequate weerspiegeling van het te meten construct. Op het eerste gezicht lijkt het eenvoudig het patiëntenperspectief mee te nemen, maar de statistische benadering daarvan is nogal lastig. In de ontwikkeling van ons itempool hebben wij verschillende logisch-op-elkaar-volgende stappen ondernomen om het patiëntenperspectief te includeren. Echter, tussen de verschillende studies was ruimte voor interpretatie door de onderzoekers. De vertaling tussen verschillende methoden is niet altijd gemakkelijk en mogelijkerwijs gaat het patiëntenperspectief deels verloren als statistische methoden worden toegepast.

We bediscussiëren ook dat het meten van vermoeidheid bij RA, net als wetenschap zelf, een iteratief proces is. Verdere validatie zal nodig zijn om te kijken of items alle aspecten afdekken die door patiënten gerapporteerd zijn. Bovendien zullen nieuwe wetenschappelijke inzichten en maatschappelijke veranderingen eisen dat de validatie van meetinstrumenten voor vermoeidheid regelmatig herhaald wordt om de nodige aanpassingen uit te voeren.

Daarnaast gaan we in op de beperkingen van het proefschrift. Hoofdstuk 2 geeft een overzicht van variabelen die met vermoeidheid bij RA samenhangen en van potentiële causale relaties. Echter, de meeste van de gereviewde studies waren cross-sectioneel en laten geen directe conclusie over causale relaties toe. De Delphi-studie, zoals beschreven in hoofdstuk 6 t/m 8,was zorgvuldig uitgevoerd en hoewel het een goede methode voor ons doeleinden was, kwamen we een paar knelpunten tegen. De studie kostte veel tijd

voor deelnemers en onderzoekers en het bleek dat deelnemers de neiging hadden om sommige items op een normatieve manier te beoordelen. Geëxcludeerde items kwamen uit interviewmateriaal en uit bestaande vragenlijsten, maar het grootste aantal items was afkomstig uit het interviewmateriaal. Het is een grote uitdaging om geschikte items vanuit interviewmateriaal te ontwikkelen, omdat er geen richtlijnen voor dit proces bestaan. Een andere beperking van de resultaten die in dit proefschrift gepresenteerd worden, is de relatief kleine steekproef van de kalibratiestudie (hoofdstuk 9). De bevindingen van deze studie moeten als voorlopige resultaten beschouwd worden. In het aansluitende project over de verdere ontwikkeling van de CAT-vermoeidheid zullen we meer data verzamelen en verdere analyses uitvoeren. In de discussie vergelijken wij bovendien de dimensies van ons voorlopige itempool met de dimensies van het andere meetinstrument voor vermoeidheid bij RA dat werd ontwikkeld vanuit het patiëntenperspectief, de BRAF-MDQ. We gaan daarbij in op doeleinden waarvoor de twee instrumenten het meest geschikt zijn. Afsluitend bediscussiëren wij mogelijke richtingen voor toekomstig onderzoek. Het is nog onduidelijk in hoeverre vermoeidheid bij RA overlap vertoont met vermoeidheid bij andere ziektes. Bovendien is dringend meer kennis nodig over oorzaken en consequenties van vermoeidheid bij RA. Pas met meer inzicht in de etiologie van vermoeidheid zal het mogelijk zijn de vraag te beantwoorden of het nodig is een specifiek meetinstrument voor RA te gebruiken. Mogelijk kan de CAT in de toekomst gevalideerd worden voor andere (reumatische) aandoeningen. Verder is het waardevol om de CAT in verschillende landen te valideren.

De bevindingen in dit proefschrift hebben meer inzicht gegeven in de betekenis van vermoeidheid voor RA. Het is duidelijk dat veel patiënten met RA vermoeidheid als onvoorspelbaar en meer extreem ervaren dan gewone vermoeidheid. Bovendien is bekend dat vermoeidheid een multidimensioneel karakter heeft. Onderzoek heeft ideeën naar voren gebracht over potentiële factoren die met vermoeidheid bij RA geassocieerd zijn. Het is duidelijk geworden op welke oorzakelijke factoren toekomstig onderzoek gericht moet zijn. Verder is het belangrijk dat er onderzoek gedaan wordt naar de manier waarop vermoeidheid behandeld of zelfs voorkomen kan worden. Daarvoor is een goed meetinstrument nodig. De resultaten in dit proefschrift voegen een stap toe aan deze uitdagende taak en bieden een degelijke basis voor de constructie van een CAT voor het meten van vermoeidheid bij RA.

Dankwoord

(Acknowledgements)

Promoveren in je eentje is niet mogelijk. Los daarvan zou het ook niet leuk zijn. Daarom ben ik blij dat er veel mensen zijn die me de afgelopen vier jaar begeleid hebben. Ik wil iedereen danken die me tijdens mijn promotietraject ondersteund heeft. Sommige mensen wil ik hier in het bijzonder noemen.

Ten eerste gaat mijn dank uit naar de patiënten die aan mijn studies hebben meegewerkt. Zonder hun participatie was dit proefschrift niet tot stand gekomen. Het was erg fijn dat ik steeds beroep kon doen op het Patiëntenforum Reuma Research Partners. Ik heb het zeer gewaardeerd dat mensen zo bereidwillig waren vragen over vermoeidheid te beantwoorden terwijl het veel van hun tijd en energie gekost moet hebben. Ook de feedback en ondersteuning van patiënten op studiemiddagen waren belangrijk voor mij omdat ik daardoor mocht ervaren hoe relevant mijn werk voor mensen is.

Ik had het plezier meerdere begeleiders te hebben; Prof.dr. Mart van de Laar, Dr. Erik Taal en Dr. Christina Bode.

Mart, hartelijk dank voor je vertrouwen in mij. Jouw interesse in mijn project, je positieve zienswijze en betrokkenheid heb ik als prettig ervaren. Ook bedankt voor de vrijheid die je me hebt gegeven en in de toekomst wilt geven. Ik weet dat dit niet vanzelfsprekend is.

Erik, het was erg fijn jou als begeleider te hebben. Jouw rustige en nauwkeurige manier en je scherpe blik hebben tot verbeteringen in mijn werk geleid. Ik mocht altijd bij je binnenlopen en je hebt steeds met me meegedacht. Bedankt!

Christina, jij was mijn dagelijkse begeleider en daar ben ik ontzettend blij mee. Je hebt me altijd het gevoel gegeven dat je deur voor mij open staat en ik kon altijd op je rekenen. Tijdens de afgelopen vier jaar heb je me uitdagingen gegeven maar je hebt er ook voor gezorgd dat ik niet te veel hooi op mijn vork zou nemen. Bedankt dat je zo´n betrokken begeleider was! Je bent een enthousiast en inspirerend iemand waarvan ik veel kon en kan leren bijvoorbeeld om te relativeren, dingen vanuit verschillende perspectieven te bekijken, iets niet zomaar voor waar aan te nemen, creatieve oplossingen te bedenken en erop te vertrouwen dat er voor elk probleem een oplossing te vinden is. Veel dank ook voor je acceptatie van mijn privéomstandigheden zoals de keuze van mijn woonplaats. Het was steeds erg leuk om met jou op congressen tussen drie talen te switchen en om het ook eens over niet-werk-gerelateerde onderwerpen te hebben. Thank you! Außerdem finde ich es bewundernswert wie Ihr es schafft eine gleichberechtigte Ehe mit zwei Kindern und Karrieren zu führen.

Bij de laatste fase van mijn project was Prof.dr. Cees Glas intensief betrokken. Cees, bedankt voor je expertise op het gebied van IRT en CAT en voor de snelle beantwoording van dringende vragen.

Mart, Erik, Christina en Cees, ik verheug me erop verder met jullie samen te werken.

Graag wil ik ook de leden van mijn promotiecommissie bedanken. Prof.dr. Ernst Bohlmeijer, Prof.dr. Joost Dekker, Prof.dr. Rinie Geenen, Prof.dr. Piet van Riel, hartelijk dank dat u zitting wilt nemen in mijn promotiecommissie. Prof.dr. Sarah Hewlett, thank you for coming to the Netherlands and being part of my examination committee.
Mijn dank gaat ook uit naar alle andere mensen die me bij de uitvoering van mijn onderzoek geholpen hebben. De teams van de poli reumatologie van het Medisch Spectrum Twente en van de ZGT Almelo/Hengelo; bedankt voor de hulp bij het werven van participanten voor mijn verschillende studies.
Ook wil ik de medewerkers van de secretariaten van de UT bedanken voor hun hulp bij organisatorische taken. In het bijzonder dank ik Marieke Smellink en Ria Stegehuis voor het snel en nauwkeurig regelen van allerlei secretariële taken.
Dr. Peter ten Klooster, bedankt dat ik altijd met vragen over methodologische onderwerpen bij je binnen mocht lopen en je geduldig al mijn vragen over Teleform hebt beantwoord.
Medepromovendi Liseth Siemons en Martijn Oude Voshaar, bedankt voor jullie hulp met onderwerpen rondom IRT en CAT, voor de gezelligheid bij IRT-cursussen en in Chicago.
Ik wil ook alle collega's en medepromovendi van de afdeling PGT danken voor de prettige werksfeer en collegialiteit. Vooral dank aan Laurien Buitinga, Martine Fledderus, Marloes Vermeer, Roos van der Vaart, Ingrid Nota, Sanne Lamers, Jojanneke Korte, Saskia Kelders en Rilana Prenger. Het was een leuke en gezellige tijd met jullie samen op de UT, bij P&H, op congressen en ook bij verschillende etentjes en feestjes.
Martine en Laurien, mijn kamergenootjes, natuurlijk bijzonder grote dank aan jullie. Op onze kamer werd hard gewerkt, maar we hebben ook veel plezier gehad. Bijna ons gehele promotietraject hebben we met elkaar gedeeld. Zonder jullie was het veel minder leuk geweest! Ik hoop dat we ook in de toekomst vriendinnen blijven.
Bovendien wil ik ook nog mijn scriptie begeleiders uit Nijmegen noemen; Dr. Yolande Kuin en Prof.dr. Judith Prins. Jullie hebben mij na mijn praktijkstage de mogelijkheid gegeven mijn scriptie in het kader van een onderzoeksstage te schrijven. Daardoor ben ik erachter gekomen welk werk bij mij past en wat ik na mijn studie wilde doen. Jullie waren inspirerende rolmodellen voor mij en hebben voor mij de deur naar de wereld van de wetenschap geopend, bedankt!
Bedankt ook Kathrin Fuchs en Christina Pawliczek. Danke dass Ihr beiden meine Paranimfen sein wollt und mir an diesem besonderen Tag auf spezielle Art und Weise zur Seite steht. Kathrin, es ist schön dass wir schon so viele Jahre kleine und große Ereignisse in unseren Leben miteinander teilen, so unterschiedlich wir auch sind. Tini, seit unserer Studentenzeit sind mittlerweile ein paar Jahre vergangen und wir haben schon dass eine oder andere miteinander erlebt. Unter anderem Danke für all die Gespräche über die

Arbeit unter Wissenschaftlerinnen. Aber ebenso möchte ich mich bei zwei weiteren Menschen bedanken: Kristina Forsmann und Carola Polat-Tenhaft. Danke für Eure Freundschaft; für gute Gespräche auch mal über ernstere Themen, gemeinsame Unternehmungen und Euer Verständnis auch für Erzählungen über meinen manchmal nicht ganz einfach erklärbaren Beruf. Ich bin froh vier so gute Freundinnen zu haben.
Ein besonderer Dank auch an meine „Mitsportler“ vom Maxmove; Sabine, Michael, Dolly, Svetlana und Holger. Der Sport und die Geselligkeit dabei waren in den letzten vier Jahren von unschätzbarem Wert für mich. Nach einem guten Training ist jeglicher Frust über überfüllte Straßen wieder vergessen. Sabine und Michael, wer hätte gedacht dass wir mal in einer familiären Beziehung zueinander stehen könnten.
Dann möchte ich natürlich zwei besonders wichtigen Menschen danken; meinen Eltern. Mama und Papa die immer hinter mir stehen und an mich glauben, egal welche in ihren Augen möglichen und unmöglichen Entscheidungen ich bislang in meinem Leben getroffen habe. Die gebündelte Menge an Durchhaltevermögen, Ehrgeiz, und Genauigkeit die ich von Euch mitbekommen habe konnte ich in den letzten vier Jahren gut gebrauchen. Mir ist klar dass ich ohne Euch und Euren Rückhalt niemals dort wäre wo ich heute bin. Unendlich vielen Dank Euch beiden!
Danke auch an die Lebenspartner meiner Eltern - Lidia und Horst - für das aufrichtige Interesse an meiner Doktorarbeit. Horst, danke für all Deine Unterstützung in den letzten Jahren, unter anderem beim Einrichten meiner Wohnumgebung.
Auch meine übrige Familie war stets interessiert an Fortschritten in meiner Arbeit. Nochmal zur Deutlichkeit, man bekommt in den Niederlanden keinen Doktorhut.
Steven, auch wenn wir nur den letzten Teil meiner Doktorandenzeit zusammen verbracht haben möchte ich Dir für so vieles danke sagen. In erster Linie für Deine Liebe, Dein Vertrauen und dass Du für mich da bist. Aber auch für praktische Dinge wie der Bereitstellung eines ungestörten Heimarbeitsplatzes und der ausgezeichneten Versorgung mit Kaffee und vegetarischem Essen - nicht nur in der letzten Schreibphase. Mit Dir war es nicht allzu schwer den nötigen Abstand zur Arbeit zu halten und Dingen die richtige Relevanz zuzuordnen. Thank you just for everything – I´m very sure that someone like you did never happen to me before. I hope that we will follow each other and not become tired of our relationship at some days but that it will remain our stable guiding light.

Stephanie Nikolaus,
Februari 2012

Curriculum vitae

Stephanie Nikolaus werd geboren op 2 november 1980 in Kleef (Duitsland). In 2000 behaalde zij haar “Abitur” aan het Johanna-Sebus-Gymnasium in Kleef. Daarna werkte zij bij het districtsbestuur in Kleef waar zij onder andere als administratief medewerkster bij de vakgroep financiën in dienst was. In 2003 begon zij met de studie psychologie aan de Radboud Universiteit Nijmegen. Vier jaar later rondde zij haar studie (afstudeerrichting psychogerontologie) met de Master of Science af en begon in november 2007 als promovenda op de afdeling Psychologie, Gezondheid & Technologie van de Universiteit Twente. Het promotietraject “Meten van vermoeidheid bij patiënten met reumatoïde artritis” werd begeleid door Dr. C. Bode, Dr. E. Taal en Prof. Dr. M. A. F. J. van de Laar. Stephanie volgde het trainingsprogramma van de onderzoeksschool Psychology & Health. De resultaten van haar promotieproject staan beschreven in dit proefschrift. Op het moment heeft Stephanie een functie als post-doc onderzoeker bij de UT om verder te werken aan de ontwikkeling van een CAT voor vermoeidheid bij RA (gefinancieerd door het Reumafonds).

Stephanie Nikolaus was born on November 2, 1980 in Kleve (Germany). After completing her secondary education in 2000, she worked at the local government in Kleve where she was employed as administrator at the financial department. In 2003 she started to study Psychology at the Radboud University Nijmegen in the Netherlands. Four years later, she completed her study (psychogerontology) with the Master of Science. In November 2007 she got employed as a PhD student at the Department of Psychology, Health & Technology of the University of Twente. The PhD project “Measuring fatigue in patients with rheumatoid arthritis” was supervised by Dr. C. Bode, Dr. E. Taal and Prof. Dr. M. A. F. J. van de Laar. Stephanie participated in the PhD training program of the Interuniversity Research Institute for Psychology & Health. The results of her PhD project are described in this dissertation. Currently Stephanie is appointed as post-doc researcher at the UT to develop a CAT for fatigue in RA (financed by the Dutch Arthritis Foundation).

List of Publications

Nikolaus S, Kuin Y, Verhagen S, Prins J. Mantelzorgers van oncologiepatiënten: ervaren belasting in de palliatieve fase en rouw na het overlijden. Gedragstherapie 2008;41:39-50.

Nikolaus S, Bode C, Taal E, van de Laar MAFJ. New Insights into the Experience of Fatigue Among Patients with Rheumatoid Arthritis: A Qualitative Study. Ann Rheum Dis 2010;69(5):895-897.

Nikolaus S, Bode C, Taal E, van de Laar MAFJ. Four different patterns of fatigue in rheumatoid arthritis patients: results of a Q-sort study. Rheumatology 2010;49(11):2191-2199.

Nikolaus S, Bode C, Taal E, van de Laar MAFJ. Selection of items for a computer-adaptive test to measure fatigue in patients with rheumatoid arthritis – A Delphi approach. Qual Life Res. Published online 31 July 2011. doi: 10.1007/s11136-011-9982-8.

Nikolaus S, van de Laar MAFJ. Measuring fatigue in rheumatoid arthritis. Nat Rev Rheumatol 2011;7(10):562-564. doi: 10.1038/nrrheum.2011.124

Nikolaus S, Bode C, Taal E, van de Laar MAFJ. Which dimensions of fatigue should be measured in patients with rheumatoid arthritis? – A Delphi study. Musculoskeletal Care. Published online 11 November 2011. doi: 10.1002/msc.222.

Nikolaus S, Bode C, Taal E, van de Laar MAFJ (in press). Experts' evaluations of fatigue questionnaires used in rheumatoid arthritis – A Delphi study among patients, nurses and rheumatologists in the Netherlands. Clin Exp Rheumatol.