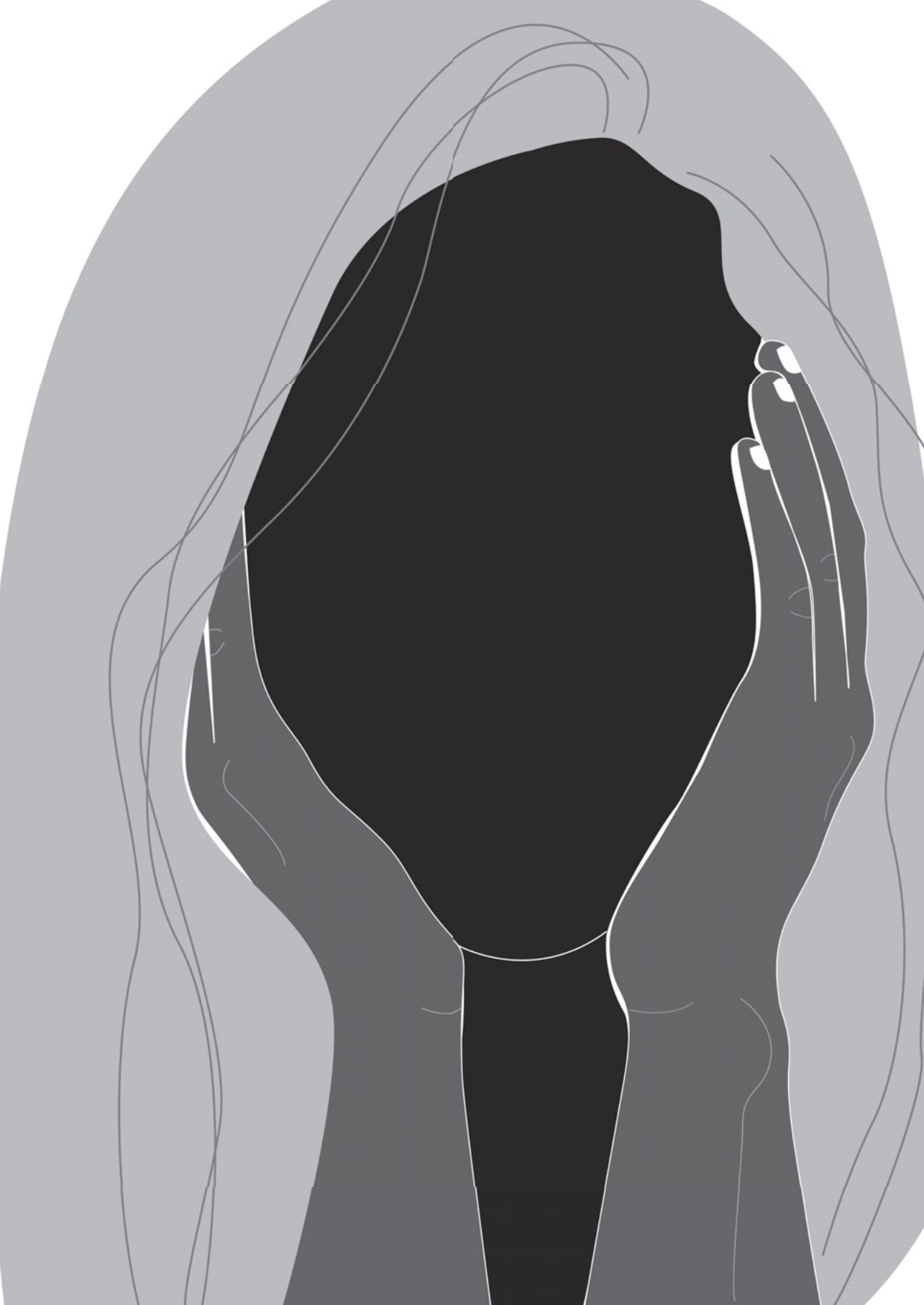


Chapter 1

General introduction

"The symptoms suddenly appeared half a year ago. Now, I always have this tingling sensation in my body, all day long. Last night I couldn't even sleep because of it, so I stayed up all night. I even have it in my fingertips, they become warm, even itch sometimes. You can't see anything, it's really inside! And I don't know what to do about it, and that's so frustrating. It leads to problems at home; I have noticed that I can be very snappy, very curt. I never used to be like this! Things annoy me very quickly. At home too, if I'm watching TV, I just feel it and it's so distracting. It takes up all of your attention. I don't understand what's causing it. I'm just trapped inside my own body." – Female patient, 48 years old, with 'medically unexplained physical symptoms' (MUPS) according to her general practitioner (GP).

"So you go get a blood test, and another one, all those examinations, you go to an internist, and so on. And each time, nothing comes out. And at some point you're just like, I'm not going again! It was the exact same thing last time. I'm a single parent, I have to be there [for her]. And she knows something's wrong with mummy, that mummy's always tired. And she keeps asking me 'why are you always tired?' I avoid many things. There are already so many things in my life that have changed, but are normal to me now. Sometimes I try not to socialize on purpose, like, I'm there but I'm not there. Please don't talk to me. I have an 11-year-old, she's more independent now, but a couple of years ago I found myself standing outside of the school on my own [in order to avoid other parents]. And then it hit me how I had changed, what I was doing, I just didn't know what to do about it."– Female MUPS patient, 50 years old.

"I wonder, do I even have a disease? Is there even anything wrong [with my body] at all? Or am I just imagining it?" – Male MUPS patient, 43 years old.

This thesis focuses on medically unexplained physical symptoms (MUPS) in adult patients in primary care. In this first chapter general background information on MUPS is provided and the topic of measuring MUPS using self-report questionnaires is introduced. Subsequently, we describe the current management of MUPS in primary care in the Netherlands. We then explain why we designed a new intervention for MUPS and describe this intervention. Then we focus on implementation considerations of the intervention. This chapter ends with an outline of the thesis. Throughout the chapter, the research questions addressed in this thesis are formulated.

First, a side note on terminology. We have chosen to use the descriptive term 'MUPS' in this thesis, to reflect our primary care focus on patients presenting symptoms. Also, this is the term used in the Dutch guideline for MUPS for general practitioners (1). In the

scientific literature, various other terms, such as 'psychosomatic symptoms', 'persistent physical symptoms', 'bodily distress (syndrome)' and 'functional somatic syndromes' have been proposed as alternatives for 'MUPS'; so far an international consensus on terminology and classification of MUPS is lacking (2, 3). The choice of terminology may depend on the medical discipline of the healthcare provider, e.g. the DSM-IV classification 'somatoform disorders' and the more recent DSM-5 classification 'somatic symptom disorder' are used in psychiatry. Despite our focus on MUPS, in our trial, we use the DSM-IV diagnosis undifferentiated somatoform disorder (USD) (the most common somatoform disorder) to operationalize MUPS, as a way to capture patients with longer lasting symptoms which impact daily life, who would likely benefit from an intervention in primary care. We chose to use the, now outdated, DSM-IV diagnosis USD, as our trial was initiated in the transition period of DSM-IV to DSM-5, and a diagnostic interview for the DSM-5 was not available yet.

Medically unexplained physical symptoms (MUPS)

Medically unexplained physical symptoms (MUPS), i.e. physical symptoms that last longer than several weeks and for which the doctor cannot find a fitting medical explanation in spite of sufficient examination, are very common in all healthcare settings (4-7). In fact, up to 25-50% of GP consultations concern symptoms that are unexplained (8). Nevertheless, this is not much reason for concern, as most symptoms are transient and do not require treatment. However, 2.5% of the general population has persisting and (more) severe MUPS (8). At this more severe end of the spectrum, persisting MUPS can be classified as a somatoform disorder according to the DSM-IV or somatic symptom disorder according to the more recent DSM-5. The most common DSM-IV somatoform disorder is the so-called 'undifferentiated somatoform disorder' (USD), including patients who suffer from at least one impairing unexplained physical symptom lasting longer than 6 months. Somatoform disorders are associated with functional impairment, poor quality of life, psychological distress and mental health problems, such as anxiety and depression (9, 10). Additionally, they are associated with high societal costs due to direct healthcare use costs and indirect costs due to absence from work and productivity loss (11, 12). Healthcare providers, such as GPs, may find patients with MUPS challenging, as they may feel they cannot sufficiently reassure the patient and may at the same time be afraid to miss a serious disease (13-15). Patients in turn may feel misunderstood, not taken seriously and afraid that they are not receiving the right diagnosis and/or optimal treatment (16, 17). All of this may lead to mutual feelings of frustration, a poor doctor-patient relationship and adverse clinical patient outcomes (18). MUPS therefore constitute a problem for patients and society and there is a need

to find suitable treatment options that reach MUPS patients, are easy to implement by healthcare providers and reduce societal costs.

Measuring MUPS

Timely detection of MUPS can help acknowledgement and labelling by doctors and patients. It could also tailor management, prevent symptoms from progressing, keep patients away from receiving unnecessary and possibly harmful referrals and treatment such as the incidentalomas that are found on MRIs, and prevent patients' and doctors' relationships deteriorating and costs increasing (19). Therefore, a tool that is also suitable for use in daily primary care settings could be useful in timely identifying patients with a higher risk of MUPS. Particularly, as self-report questionnaires are a quick, useful and non-invasive tool, they might be able to help people identify and acknowledge MUPS in the context of an ongoing clinical dialogue.

At the same time, measurement of MUPS can be complex due to high levels of clinical uncertainty (i.e., finding a 'hidden' physical explanation after all, such as for example multiple sclerosis), a lack of formal diagnostic criteria for MUPS and the fact that the current classifications are mainly based on negative features, such as absence of a somatic explanation (20). Diagnosing MUPS therefore requires an expert clinical assessment. Furthermore, psychological aspects such as cognitions, emotions and behaviours (21) also may contribute to MUPS, which may be difficult to incorporate into a user-friendly measurement instrument for use in clinical settings.

Instruments for directly measuring MUPS as a multidimensional, core concept are lacking. Current questionnaires often focus on measuring the number of physical symptoms, and purport to measure the related concept of somatization. Somatization has been defined by Lipowski et al. as "a tendency to experience and communicate somatic distress and symptoms unaccounted for by pathological findings, to attribute them to physical illness, and to seek medical help for them" (22). As with MUPS, this definition of somatization encompasses more than just the number of physical symptoms. However, previous research found that a higher number of reported physical symptoms is associated with a functional somatic syndrome diagnosis (23). There is a relationship between the experienced number of physical symptoms, or the symptom count, and somatization. Therefore, instruments measuring 'somatization' may be useful as a proxy for MUPS.

Different questionnaires have been proposed to measure somatization, such as the Patient Health Questionnaire 15-item somatic symptom severity subscale (PHQ-15)

(24), the 4-Dimensional Symptom Questionnaire (4DSQ) somatization subscale (25), the Symptom Checklist-90-R (SCL-90-R) somatization subscale (26), and also some newer ones such as the Schedule for Evaluating Persistent Symptoms (SEPS) (20) and the Somatic Symptom Disorder - B criteria scale (SSD-12) (27). However, few questionnaires have been specifically developed for and validated in primary care. We wanted to find out which currently available questionnaire has the best test characteristics for detection of somatization in primary care.

Our first research question is therefore:

What is the best self-report measurement instrument to measure somatization in primary care?

Current management of MUPS in primary care

In 2013, the Dutch MUPS guideline for GPs was published. It provides diagnostic and therapeutic recommendations (1). One of the key points for diagnosis is symptom exploration, not only of the somatic dimension (S), but also of the cognitive (C), emotional (E), behavioural (B) and social (S) aspects of the patient's symptoms. As MUPS is assumed to be influenced by bio-psycho-social factors, it is important to explore all these dimensions to find the most fitting treatment for the patient (28-30). Therapeutic recommendations include a stepped care model in which GPs treat MUPS patients themselves when symptoms are mild, refer patients to a mental health nurse practitioner (MHNP) or another primary care healthcare provider such as a psychosomatic physiotherapist when symptoms are moderate, and refer to a multidisciplinary secondary care treatment center when symptoms are severe and functioning Is impaired.

Despite the existence of the guideline, it is currently unknown whether GPs actually comply with it. Due to various barriers such as insufficient communication skills, knowledge, confidence and time constraints, GPs may find it difficult to work with a bio-psycho-social approach (31). We were therefore interested in finding out what the current management of MUPS patients entails and whether the recommendations from the guideline are being followed by Dutch GPs. Additionally, investigating current management of MUPS offers a comparison to our own intervention.

Our second research question is:

What does the current management of MUPS patients in Dutch primary care entail and to what extent is it in line with the national guideline for persistent MUPS?

Rationale for our intervention

Various treatment forms, such as pharmacological, non-pharmacological treatments, and enhanced care have been investigated for their effectiveness versus control conditions on MUPS and been described in various meta-analyses (21, 32, 33). Pharmacological interventions did not succeed in reducing symptom severity versus placebo or treatment as usual (33). The effectiveness of medication such as different types of antidepressants, antipsychotics, a combination of antidepressant and an antipsychotic, and natural products was investigated versus placebo, treatment as usual and other medication. Only (very) low quality evidence was found for new generation anti-depressants and natural products when compared with a placebo. However, these studies had serious shortcomings, such as a high risk of bias (particularly allocation concealment and blinding), strong heterogeneity in the data and small sample sizes. Also, none of the trials included follow-up assessments longer than 12 weeks. Therefore, firm conclusions cannot be drawn from these results. Additionally, adverse side effects of the medication may have amplifying effects on symptom perception, which is particularly detrimental to MUPS patients (33).

The most effective treatment thus far seems to be cognitive behavioural therapy (CBT), a psychological treatment. With CBT, unhelpful cognitions and behaviours concerning the symptoms are addressed and modified into more helpful ones. As a result, symptom reduction can be achieved. Previous studies of low to moderate quality found that CBT significantly, but modestly, improves physical functioning and reduces somatic, anxiety and depressive symptoms in patients with MUPS (32, 34). The effects of CBT also seem to last at least up to one year follow-up.

Research investigating CBT for MUPS has been conducted in primary and secondary healthcare settings, and CBT has been delivered by different healthcare providers such as GPs, psychologists and psychotherapists. Whereas CBT interventions carried out by psychologists and psychotherapists mostly seem to be effective over control conditions such as waiting lists, GPs seem to achieve poorer results (21, 35-37). Possible reasons for psychotherapists achieving better results could be that psychotherapists take more time to explain the additional factors influencing the symptoms and provide more sessions than the GP. On the other hand, patients seeking help from a psychotherapist may differ in their motivation for change from patients seeking help from a GP, and may be in a different stage of accepting their symptoms and coping with them.

Results of a recent literature review investigating treatment of physical complaints by GPs and nurses suggest that interventions delivered by primary care nurse-practitioners

have at least equal and sometimes even more beneficial effects on patient satisfaction and quality of life in primary care patients than interventions delivered by GPs (38). However, similar research has not yet been conducted for MUPS. Also, there is little research on effectiveness of interventions delivered by nurses who have a mental health background.

The Dutch multidisciplinary guideline for MUPS and somatoform disorders (2011) contains several recommendations for scientific research with respect to MUPS (39). One of these recommendations is that research should be done in primary care in particular, and should investigate effectiveness of various types of CBT-based interventions, such as classic CBT and problem-solving treatment (PST). Interventions carried out by GPs, primary care psychologists, (psychosomatic) physiotherapists and nurses/mental health nurse practitioners should be studied.

In the Netherlands, every citizen registers with a general practice. The GP is the first healthcare provider patients turn to with physical or mental symptoms. The GP has a gatekeeper's role and, in general, patients have no access to secondary care without a referral from the GP. When treatment in secondary care has ended, the patient is referred back to the GP. The GP therefore always remains in charge of the management of the patient's healthcare.

In 2014 a reform in Dutch healthcare took place in order to make mental healthcare more accessible in primary care and to reduce increasing mental healthcare costs in the secondary care (40). The position of a MHNP already existed since 2007 in some practices, but was not mandatory or especially stimulated by the Dutch government and healthcare insurance companies. As part of the healthcare reform in 2014, all general practices received extra money to employ a MHNP for one day a week per fulltime GP. In 2016, 87% of the general practices in the Netherlands had a MHNP and these seem to fulfill a need, as can be seen by the increasingly growing number of patient consultations with the MHNP and even waiting lists (41, 42).

MHNPs work under the supervision of the GP and are mostly trained social psychiatric nurses or psychologists (40, 42). A smaller percentage are nurses, social workers or remedial educationalists (42). The GP typically sees the patient first and may then decide to refer the patient to the MHNP. The MHNPs' main tasks are to perform a diagnostic work-up, deliver short-term counselling for patients with psychosocial problems and assist in referrals to other mental health caregivers if needed (40).

General practice provides asset-based generalist support for people with distressing and persistent physical and mental health symptoms but generally does not provide specific treatments for them. The introduction of the MHNP in Dutch primary care offers community-based opportunities for more specific forms of support for such patients, but that would require further evidence of their (cost-)effectiveness. It would also be necessary to better understand how to further tailor such treatments to what people with MUPS need in primary care and how to implement them if they work.

With this in mind, one of the aims of our research is to develop and evaluate a new CBT-based intervention suitable for MHNPs in a randomized controlled trial called the Cognitive-behavioural Intervention in PRimary care for Undifferentiated Somatoform disorder (CIPRUS) study. The intervention consists of a combination of the consequences model of MUPS, which establishes the consequences (e.g. emotional, behavioural, functional) of MUPS experienced in one's life, and cognitive behavioural problem-solving techniques according to the steps outlined in problem-solving treatment (PST). Using PST, patients will learn to tackle the various consequences and develop general problem-solving skills.

The third research question addressed in this thesis is:

> *What is the effectiveness of a CBT-based intervention for patients with undifferentiated somatoform disorder carried out by mental health nurse practitioners in Dutch primary care?*

Costs

Additionally, MUPS are burdensome and costly for the patient and society (11). Previous studies showed that patients with MUPS generate substantial healthcare use costs (12), possibly because patients do not receive a sufficient explanation for their symptoms and keep returning and seeking more help. Patients are also often subjected to unnecessary secondary care referrals and examinations, sometimes causing poorer clinical outcomes and more anxiety for the patient (43). Furthermore, non-healthcare use costs are high in this patient group as well. Patients with MUPS often feel very limited in their functioning, and therefore make high indirect societal costs in the form of loss of productivity at work, sickness absence, and receiving paid and unpaid help with their daily activities (44).

A recent systematic review provided an overview of economic evaluations of different interventions for MUPS (45). CBT interventions, being as yet the most effective intervention for MUPS, were included in this review as well. Results show that CBT is not only more effective but is also generally cost-effective compared to other types of

interventions or usual care. Furthermore, group interventions are often considered to be the most cost-effective, obviously because the costs of the treatment can be divided by the number of patients in the treatment group.

However, group interventions in primary care may be complicated to organize and not pragmatic, as the current Dutch primary healthcare system primarily allows for individual sessions between patient and MHNP. The MHNP is in a perfect position to carry out a CBT-based intervention within general practice, but it is unknown whether such an intervention would be cost-effective.

Our fourth research question is therefore:

Is the new CBT-based intervention cost-effective compared to current usual care?

If our intervention turns out to be effective in treating patients with undifferentiated somatoform disorder, we are also interested in finding out *how* the intervention works, i.e. which mechanisms are responsible for change. As we opt for a problem-solving intervention, we hypothesize that the intervention will bring about change by first altering problem-solving skills of patients with MUPS. Problem-solving skills will in turn improve patients' functioning and quality of life. Other factors, such as dysfunctional somatic attributions, cognitions, behavioural responses and health anxiety, are thought to be important maintaining factors of MUPS as well. We investigate these factors as potential mediators in our trial.

Implementation considerations

The main question in randomized controlled trials is usually "does the intervention work?" In our research project we aimed to provide an answer to this question. Moreover, we also report on another important question, whether the intervention is cost-effective. However, apart from establishing whether the intervention is (cost-)effective, it is also important to evaluate the process of the trial and how results may be implemented or 'normalised' into the daily 'work' of a practice. This can be done by conducting process evaluations with participants, patients as well as MHNPs. Process evaluations help gather information on topics such as implementation, receipt and setting of an intervention and help in the interpretation of results (46, 47). More detailed information on delivery of the intervention can also be gathered. It then becomes clearer to which parts of the intervention effects can be attributed. Patients and MHNPs deal with an intervention in the real world and therefore may have valuable information on its strengths and drawbacks. Involving participants and exploring their views can help distinguish between components that

are useful or faulty, and barriers and facilitators to implementation. Eventually, this type of crucial information can help improve an intervention.

Therefore, our final research question is:

How did patients and MHNPs feel about and evaluate the CBT-based intervention for undifferentiated somatoform disorder?

Outline of this thesis

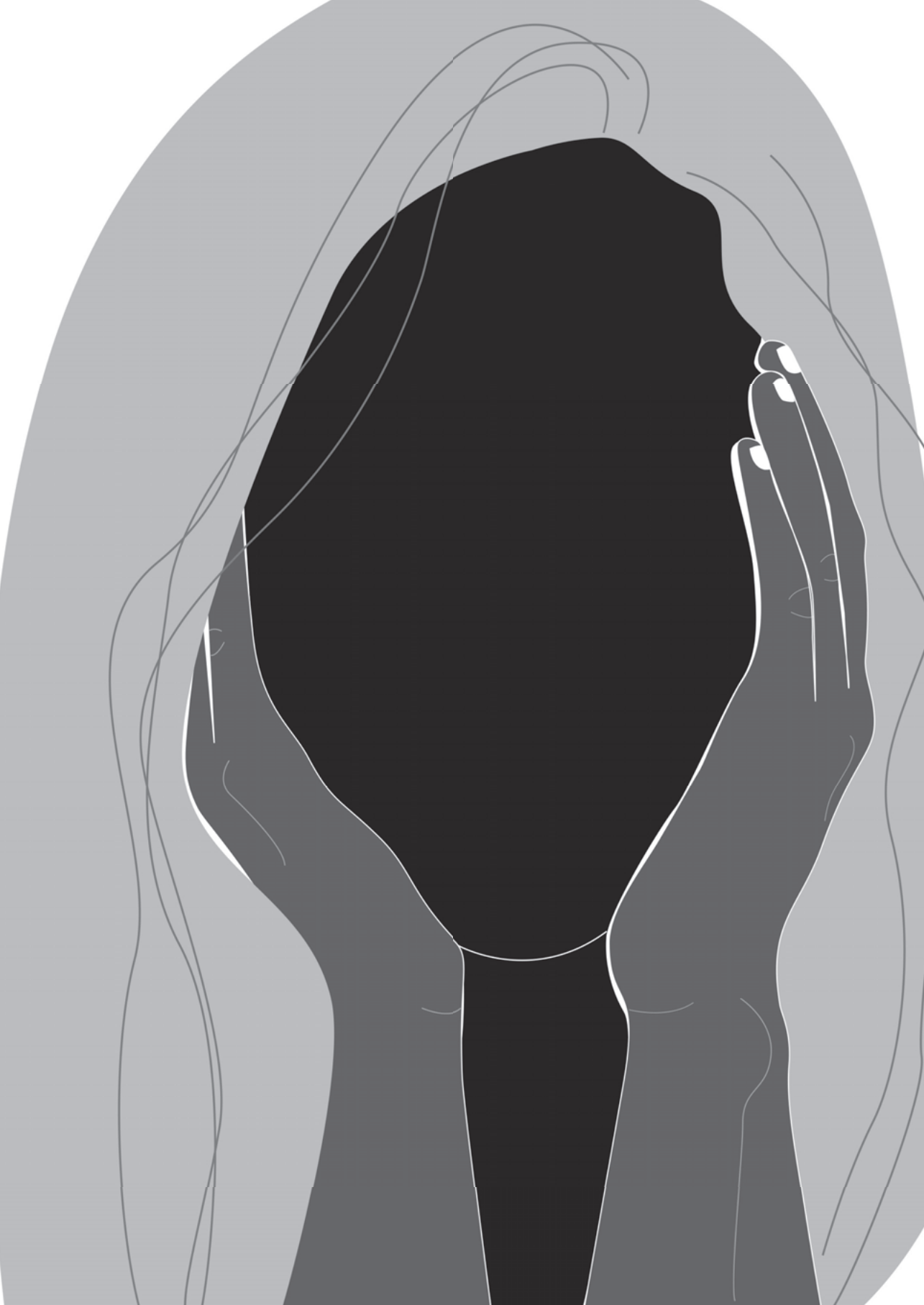
In **chapter 2** we present results of a systematic review on clinimetric properties of questionnaires that measure somatization. In this chapter we aimed to provide an overview of the best measurement instruments according to the state-of-the-art COSMIN criteria. **Chapter 3** presents the study design of the randomized controlled trial of the CIPRUS study. In this chapter the methods of the effectiveness trial and the cost-effectiveness study are described in detail. In **chapter 4** a descriptive study of the current usual care in general practices in the Netherlands is presented. For this study we investigated current management of patients with MUPS in the 'usual care practices' of the CIPRUS study. **Chapter 5** presents the results of the effectiveness of the CIPRUS study over a period of 12 months. In **chapter 6** we present the results of our intervention's cost-effectiveness compared to usual care. **Chapter 7** presents the results of a process evaluation among patients and MHNPs who respectively received and provided the cognitive behavioural intervention. Finally, in **chapter 8** we summarize the main overall findings of this thesis, provide a methodological reflection, reflect on the strengths and limitations and give recommendations for practice and future research.

REFERENCES

1. Olde Hartman T, Blankenstein A, Molenaar A, Bentz van den Berg D, van der Horst H, Arnold I, et al. [NHG guideline on medically unexplained symptoms (MUS)] (in Dutch). Huisarts Wet. 2013;56(5):222-30.
2. Henningsen P, Zipfel S, Sattel H, Creed F. Management of Functional Somatic Syndromes and Bodily Distress. Psychother Psychosom. 2018;87(1):12-31.
3. Budtz-Lilly A, Schroder A, Rask MT, Fink P, Vestergaard M, Rosendal M. Bodily distress syndrome: a new diagnosis for functional disorders in primary care? BMC Fam Pract. 2015;16:180.
4. Haller H, Cramer H, Lauche R, Dobos G. Somatoform disorders and medically unexplained symptoms in primary care. Dtsch Arztebl Int. 2015;112(16):279-87.
5. Steinbrecher N, Koerber S, Frieser D, Hiller W. The prevalence of medically unexplained symptoms in primary care. Psychosomatics. 2011;52:263-71.
6. Nimnuan C, Hotopf M, Wessely S. Medically unexplained symptoms, an epidemiological study in seven specialities. J Psychosom Res. 2001;51:361-7.
7. Reid S, Wessely S, Crayford T, Hotopf M. Medically unexplained symptoms in frequent attenders of secondary health care: retrospective cohort study. BMJ. 2001;322:1-4.
8. Verhaak PF, Meijer SA, Visser AP, Wolters G. Persistent presentation of medically unexplained symptoms in general practice. Fam Pract. 2006;23(4):414-20.
9. Haftgoli N, Favrat B, Verdon F, Vaucher P, Bischoff T, Burnand B, et al. Patients presenting with somatic complaints in general practice: depression, anxiety and somatoform disorders are frequent and associated with psychosocial stressors. BMC Fam Pract. 2010;11(67).
10. de Waal MWM, Arnold IA, Eekhof JAH, van Hemert AM. Somatoform disorders in general practice: prevalence, impairment and comorbidity with anxiety and depressive disorders. Br J Psychiatry. 2004;184(6):470-76.
11. Konnopka A, Schaefert R, Heinrich S, Kaufmann C, Luppa M, Herzog W, et al. Economics of medically unexplained symptoms: a systematic review of the literature. Psychother Psychosom. 2012;81(5):265-75.
12. Barsky AJ, Orav EJ, Bates DW. Somatization increases medical utilization and costs independent of psychiatric and medical comorbidity. Arch Gen Psychiatry. 2005;62:903-10.
13. Sirri L, Grandi S, Tossani E. Medically unexplained symptoms and general practitioners: a comprehensive survey about their attitudes, experiences and management strategies. Fam Pract. 2017;34(2):201-5.
14. Howman M, Walters K, Rosenthal J, Ajjawi R, Buszewicz M. "You kind of want to fix it don't you?" Exploring general practice trainees' experiences of managing patients with medically unexplained symptoms. BMC Med Educ. 2016;16:27.
15. Johansen ML, Risor MB. What is the problem with medically unexplained symptoms for GPs? A meta-synthesis of qualitative studies. Patient Educ Couns. 2017;100(4):647-54.
16. Werner A, Malterud K. It is hard work behaving as a credible patient: encounters between women with chronic pain and their doctors. Soc Sci Med. 2003;57(8):1409-19.
17. Nettleton S. 'I just want permission to be ill': towards a sociology of medically unexplained symptoms. Soc Sci Med. 2006;62(5):1167-78.
18. Stone L. Managing the consultation with patients with medically unexplained symptoms: a grounded theory study of supervisors and registrars in general practice. BMC Fam Pract. 2014;15:192.
19. Rosendal M, Olde Hartman TC, Aamland A, van der Horst H, Lucassen P, Budtz-Lilly A, et al. "Medically unexplained" symptoms and symptom disorders in primary care: prognosis-based recognition and classification. BMC Fam Pract. 2017;18(1):18.
20. Tyrer H, Ali L, Cooper F, Seivewright P, Bassett P, Tyrer P. The Schedule for Evaluating Persistent Symptoms (SEPS): a new method of recording medically

unexplained symptoms. Int J Soc Psychiatry. 2013;59(3):281-7.
21. Rosendal M, Blankenstein AH, Morriss R, Fink P, Sharpe M, Burton C. Enhanced care by generalists for functional somatic symptoms and disorders in primary care. Cochrane Database Syst Rev. 2013(10):CD008142.
22. Lipowski ZJ. Somatization: the concept and its clinical application. Am J Psychiatry. 1988;145:1358–68.
23. Kingma EM, de Jonge P, Ormel J, Rosmalen JG. Predictors of a functional somatic syndrome diagnosis in patients with persistent functional somatic symptoms. Int J Behav Med. 2013;20(2):206-12.
24. Kroenke K, Spitzer RL, Williams JBW. The PHQ-15: Validity of a new measure for evaluating the severity of somatic symptoms. Psychosom Med. 2002;64(2):258-66.
25. Terluin B, van Marwijk HW, Ader HJ, de Vet HC, Penninx BW, Hermens ML, et al. The Four-Dimensional Symptom Questionnaire (4DSQ): a validation study of a multidimensional self-report questionnaire to assess distress, depression, anxiety and somatization. BMC Psychiatry. 2006;6:34.
26. Derogatis LR. The Symptom Checklist-90-revised. Minneapolis, MN: NCS Assessments; 1992.
27. Toussaint A, Lowe B, Brahler E, Jordan P. The Somatic Symptom Disorder - B Criteria Scale (SSD-12): Factorial structure, validity and population-based norms. J Psychosom Res. 2017;97:9-17.
28. Tanaka Y, Kanazawa M, Fukudo S, Drossman DA. Biopsychosocial model of irritable bowel syndrome. J Neurogastroenterol Motil. 2011;17(2):131-9.
29. Van Houdenhove B, Egle UT. Fibromyalgia: a stress disorder? Piecing the biopsychosocial puzzle together. Psychother Psychosom. 2004;73(5):267-75.
30. Johnson SK. The biopsychosocial model and chronic fatigue syndrome. Am Psychol. 1998;53(9):1080-1.
31. Murray AM, Toussaint A, Althaus A, Lowe B. The challenge of diagnosing non-specific, functional, and somatoform disorders: A systematic review of barriers to diagnosis in primary care. J Psychosom Res. 2016;80:1-10.
32. van Dessel N, den Boeft M, van der Wouden J, Kleinstäuber M, Leone S, Terluin B, et al. Non-pharmacological interventions for somatoform disorders and medically unexplained physical symptoms (MUPS) in adults. Cochrane Database Syst Rev. 2014;11, CD01114.
33. Kleinstäuber M, Witthöft M, Steffanowski A, van Marwijk H, Hiller W, Lambert MJ. Pharmacological interventions for somatoform disorders in adults Cochrane Database Syst Rev. 2014;11:CD010628.
34. Jing L, Gill NS, Teodorczuk A, Li ZJ, Sun J. The efficacy of cognitive behavioural therapy in somatoform disorders and medically unexplained physical symptoms: A meta-analysis of randomized controlled trials. J Affect Disord. 2018;245:98-112.
35. Gerger H, Hlavica M, Gaab J, Munder T, Barth J. Does it matter who provides psychological interventions for medically unexplained symptoms? A meta-analysis. Psychother Psychosom. 2015;84:217-26.
36. Sumathipala A, Siribaddana S, Abeysingha MR, De Silva P, Dewey M, Prince M, et al. Cognitive-behavioural therapy v. structured care for medically unexplained symptoms: randomised controlled trial. Br J Psychiatry. 2008;193(1):51-9.
37. Zonneveld LN, van Rood YR, Timman R, Kooiman CG, Van't Spijker A, Busschbach JJ. Effective group training for patients with unexplained physical symptoms: a randomized controlled trial with a non-randomized one-year follow-up. PLoS One. 2012;7(8):e42629.
38. Laurant M, van der Biezen M, Wijers N, Watananirun K, Kontopantelis E, van Vught AJ. Nurses as substitutes for doctors in primary care. Cochrane Database Syst Rev. 2018;7:CD001271.
39. Trimbos instituut. [Multidisciplinary guideline for MUPS and somatoform disorders] (In Dutch). 2011.
40. Magnee T, de Beurs DP, de Bakker DH, Verhaak PF. Consultations in general practices with and without mental health nurses: an observational study from 2010 to 2014. BMJ Open. 2016;6(7):e011579.
41. Verhaak PFM, Nielen M, de Beurs D. Factsheet 1: [Demand for care - aid

regarding psychological and social problems in the general practice 2011-2016] (in Dutch). Utrecht: Netherlands Institute of Health Services Research (NIVEL); 2017.
42. [National Society of General Practitioners]. [Members of the National Society of General Practitioners mental healthcare poll results] (In Dutch). 2017.
43. Burton C, McGorm K, Richardson G, Weller D, Sharpe M. Healthcare costs incurred by patients repeatedly referred to secondary medical care with medically unexplained symptoms: a cost of illness study. J Psychosom Res. 2012;72(3):242-7.
44. Gustavsson A, Svensson M, Jacobi F, Allgulander C, Alonso J, Beghi E, et al. Cost of disorders of the brain in Europe 2010. Eur Neuropsychopharmacol. 2011;21(10):718-79.
45. Wortman MSH, Lokkerbol J, van der Wouden JC, Visser B, van der Horst HE, Olde Hartman TC. Cost-effectiveness of interventions for medically unexplained symptoms: A systematic review. PLoS One. 2018;13(10):e0205278.
46. Oakley A, Strange V, Bonell C, Allen E, Stephenson J. Process evaluation in randomised controlled trials of complex interventions. BMJ. 2006;332:413-16.
47. Liu H, Muhunthan J, Hayek A, Hackett M, Laba TL, Peiris D, et al. Examining the use of process evaluations of randomised controlled trials of complex interventions addressing chronic disease in primary health care-a systematic review protocol. Syst Rev. 2016;5(1):138.



Chapter 2

Systematic review of measurement properties of questionnaires measuring somatization in primary care patients

Kate Sitnikova
Sandra MA Dijkstra-Kersten
Lidwine B Mokkink
Berend Terluin
Harm WJ van Marwijk
Stephanie S Leone
Henriëtte E van der Horst
Johannes C van der Wouden

Published in: Journal of Psychosomatic Research 2017;103:42-62.

ABSTRACT

Objective: The aim of this review is to critically appraise the evidence on measurement properties of self-report questionnaires measuring somatization in adult primary care patients and to provide recommendations about which questionnaires are most useful for this purpose.

Methods: We assessed the methodological quality of included studies using the COnsensus-based Standards for the selection of health Measurement INstruments (COSMIN) checklist. To draw overall conclusions about the quality of the questionnaires, we conducted an evidence synthesis using predefined criteria for judging the measurement properties.

Results: We found 24 articles on 9 questionnaires. Studies on the Patient Health Questionnaire-15 (PHQ-15) and the Four-Dimensional Symptom Questionnaire (4DSQ) somatization subscale prevailed and covered the broadest range of measurement properties. These questionnaires had the best internal consistency, test-retest reliability, structural validity, and construct validity. The PHQ-15 also had good criterion validity, whereas the 4DSQ somatization subscale was validated in several languages. The Bodily Distress Syndrome (BDS) checklist had good internal consistency and structural validity. Some evidence was found for good construct validity and criterion validity of the Physical Symptom Checklist (PSC-51) and good construct validity of the Symptom Check-List (SCL-90-R) somatization subscale. However, these three questionnaires were only studied in a small number of primary care studies.

Conclusion: Based on our findings, we recommend the use of either the PHQ-15 or 4DSQ somatization subscale for somatization in primary care. Other questionnaires, such as the BDS checklist, PSC-51 and the SCL-90-R somatization subscale show promising results but have not been studied extensively in primary care.

INTRODUCTION

Experiencing one or several medically unexplained symptoms without a known underlying somatic explanation is common for all people, especially in stressful situations. However, experiencing many medically unexplained symptoms from various organ systems may imply somatization (1). A widely accepted definition of somatization is: "a tendency to experience and communicate somatic distress and symptoms unaccounted for by pathological findings, to attribute them to physical illness, and to seek medical help for them" (2). If symptoms persist, patients may seek medical help. Due to its generalist nature, primary care is the first port-of-call for people who are worried about such physical experiences, but in all health care settings a substantial number of patients have medically unexplained symptoms (3, 4).

Unexplained physical symptoms in primary care can be aligned across a spectrum of the number, severity and functional impairment of symptoms, with having just one or a few transient symptoms at one end of the spectrum, and having multiple severe symptoms for a long period of time and therefore meeting diagnostic criteria for a somatoform disorder according to the Diagnostic and Statistical Manual of mental disorders 4th, (DSM-IV) (5) or a somatic symptom disorder according to the 5th edition (DSM-5) (6), at the other end (7). In primary care, mostly patients with physical symptoms at the milder end of the spectrum are seen. However, patients with multiple severe symptoms also frequently end up in primary care, usually when after referral to specialized settings further examinations yielded no results and patients are referred back to primary care practice.

The sooner high levels of somatization are signalled and discussed, the sooner patients can learn to make sense of them and the sooner appropriate care can be provided. As a result, otherwise potentially unnecessary, costly, medical procedures with possible side-effects can be avoided. Considering the general practitioners' (GP) and nurse practitioners' time-restrictions, self-report questionnaires can be a useful, quick, non-invasive tool to assist GPs in detecting symptoms of somatization directly from the patient's point of view.

Somatization is a complicated concept to measure, as in addition to the dimension of experienced physical symptoms, after a somatic cause has been ruled out, it also has cognitive and behavioural dimensions (8, 9). It is particularly difficult to operationalize cognitions, attributions, worries and behavioural aspects, such as seeking medical help, and incorporate these dimensions in one measurement instrument with the experienced

physical symptoms (10). As previous research found that the number of symptoms predicts the course of the medically unexplained symptoms and health status (11, 12), we use the experienced physical symptoms, or the symptom count, as a proxy for somatization in this review, which is also common in other studies (13–16). Therefore, we restrict our definition of 'somatization' to having multiple physical symptoms, unaccounted for by somatic pathology, at the same time and look into questionnaires that quantify these symptoms, their severity and impairment caused by the symptoms as a proxy for somatization. We acknowledge the various possible explanatory factors and consequences that somatization can have, but do not focus on these in the current study.

Research comparing the quality of various available questionnaires to measure somatization in primary care has not yet been done. Therefore, to date, it remains unclear which questionnaire can be used best for this purpose.

Two previous articles (17, 18) provided overviews of measurement instruments, one for common somatic symptoms (17) and the other for somatoform disorders (18). However, neither was specifically focused on use in primary care and neither used the state-of-the-art COnsensus-based Standards for the selection of health Measurement Instruments (COSMIN) methodology (19, 20) for conducting systematic reviews on measurement instruments.

The aim of this review is to critically appraise the evidence on the measurement properties of (subscales of) self-report questionnaires measuring somatization in adult primary care patients and to provide recommendations about which questionnaires are most useful for this purpose.

METHODS

This review is reported according to the Preferred Reporting Items for Systematic Reviews and Meta-Analyses (PRISMA) guidelines (21).

Literature search

A search was performed on August 13, 2015 in PubMed/Medline, Embase, Psycinfo and Cinahl from inception. No time period restrictions were used. In all databases search terms for construct, population, measurement properties and setting were combined using the Boolean operator 'AND'. In PubMed a validated search filter was used for finding articles investigating measurement properties (22). In the other databases, adapted versions of this search filter were used. The adaptations were performed by

a scientific information specialist. The full search strategies for each database can be found in Appendix A. A second updated search was performed on October 31, 2016 following the same procedure, in order to include articles published after our initial search. Reference lists of the included articles and reviews found during the searches, were searched to identify additional relevant articles. Authors of articles were contacted in case manuscripts were not available online.

Inclusion and exclusion criteria

Inclusion criteria were:

1. The questionnaire or subscale aims to measure somatization defined as having multiple physical symptoms.
2. The study population is adults (age 18 and above) who are patients in primary care.
3. The instrument of study is developed as a paper or online self-report questionnaire.
4. The aim of the study is the development of a questionnaire or the evaluation of one or more of its measurement properties.
5. The article is published as a full text original article.

Exclusion criteria were:

1. The article is published in languages other than English or Dutch.
2. The study measures somatization as a personality or character trait.
3. The study investigates a specific functional syndrome (e.g. fibromyalgia, irritable bowel syndrome, chronic pain syndrome).
4. The questionnaire includes items on somatization among other items, but without a separate subscore for somatization.

Selection procedure

The selection of articles based on titles and abstracts was independently performed by two reviewers (KS and SDK). Afterwards, these two reviewers separately checked whether the full text articles met the inclusion criteria. In case of disagreement or doubt, a third reviewer (JW/BT) was consulted in order to make the decision regarding inclusion of the article.

Data extraction

Two reviewers (KS and SDK) independently extracted and evaluated the general characteristics of the questionnaires, the characteristics of the studies, and information on generalizability and interpretability, using a structured form. When not enough information could be obtained from the included articles, original development articles

were consulted. Disagreement between reviewers was discussed until consensus was reached. In case of disagreement or doubt, a third reviewer (JW) was consulted.

Assessment of the methodological quality of the included studies

The methodological quality of the studies was assessed using the COSMIN checklist (19). The COSMIN checklist has been developed in an international Delphi Study and can be used to evaluate the methodological quality of studies on measurement properties. The COSMIN checklist consists of 12 boxes. Nine boxes contain standards for quality of the methodological properties reliability, measurement error, content validity, structural validity, hypotheses testing, cross-cultural validity, criterion validity and responsiveness. One box contains standards for studies on interpretability. One box contains general requirements for articles using item response theory (IRT), and one box contains general requirements for the generalizability of results.

We used the COSMIN checklist (20) to determine which measurement properties were evaluated in a study. Two reviewers (KS and SDK) then independently evaluated the quality of the included studies per measurement property, using the COSMIN checklist 4-point rating scale (19, 23) (available from the website www.cosmin.nl). In case of disagreement or doubt, a third reviewer (JW/LM) was consulted.

We modified the COSMIN checklist slightly by omitting the two items on the percentage and handling of missing data. This choice was made because it is unclear how missing data relates to methodological quality and what the best way to handle missing data is, when investigating measurement properties.

Furthermore, the cross-cultural validity box concerns two different aspects: 1) translation of the instrument, and 2) the actual cross-cultural validation analysis between two culturally different groups. To acknowledge these two aspects, we decided to split the box into two sections, i.e. translation score (items 4–11) and the cross-cultural validation score (items 1–3 and 12–15). For criterion validity, we considered validated interviews based on the Diagnostic and Statistical Manual of mental disorders (DSM-IV) (5) criteria for somatoform disorders to be the gold standard (e.g. the Schedules for Clinical Assessment in Neuropsychiatry (SCAN) (24), the Structured Clinical Interview for DSM-IV-TR Axis I Disorders (SCID-I) (25), and the Mini International Neuropsychiatric Interview (MINI) (26)). Studies using other comparison instruments were not considered to address criterion validity but evaluated under hypotheses testing.

Evaluation of the study results against criteria for good measurement properties

The data of each study were extracted and compared to criteria for good measurement properties developed by Prinsen and colleagues in cooperation with the COSMIN initiative (27). We slightly adjusted the criteria described by Prinsen et al. for reliability and criterion validity (Appendix B). The adjustment was done because various studies reported other appropriate coefficients besides the intraclass correlation coefficient (ICC) or weighted kappa when assessing reliability, and other appropriate values than exclusively a correlation were used when assessing criterion validity. For reliability we scored a '+' when ICC or weighted kappa was ≥0.70 and also when Pearson's correlation coefficient was ≥0.80. We scored a "–" when these criteria were not met. For criterion validity we scored a '+' when the correlation with the gold standard was ≥0.70, but also when the area under the curve (AUC) was ≥0.70 or, in case no correlation or AUC was provided, when both sensitivity and specificity were ≥60%. We scored a "–" when none of these criteria were met.

Data syntheses

For each questionnaire, the overall levels of evidence on each measurement property were synthesized using the data on measurement properties from all included studies (28). The levels of evidence were adjusted for the methodological quality of each study according to the criteria provided in Table 1.

Table 1. Levels of evidence for the quality of the measurement properties (21)

Level	Rating	Criteria
Strong	+++ or - - -	Consistent findings in multiple studies of good methodological quality OR in one study of excellent methodological quality
Moderate	++ or - -	Consistent findings in multiple studies of fair methodological quality OR in one study of good methodological quality
Limited	+ or -	One study of fair methodological quality
Conflicting	+/-	Conflicting findings
Unknown	?	Only studies of poor methodological quality

+ = positive rating, ? = indeterminate rating, - = negative rating.

RESULTS

Included studies

The search yielded 5318 hits in total, of which 1326 hits in PubMed, 3029 in Embase, 729 in Psycinfo and 234 in Cinahl. An overview of the searches and selection of articles is presented in Figure 1.

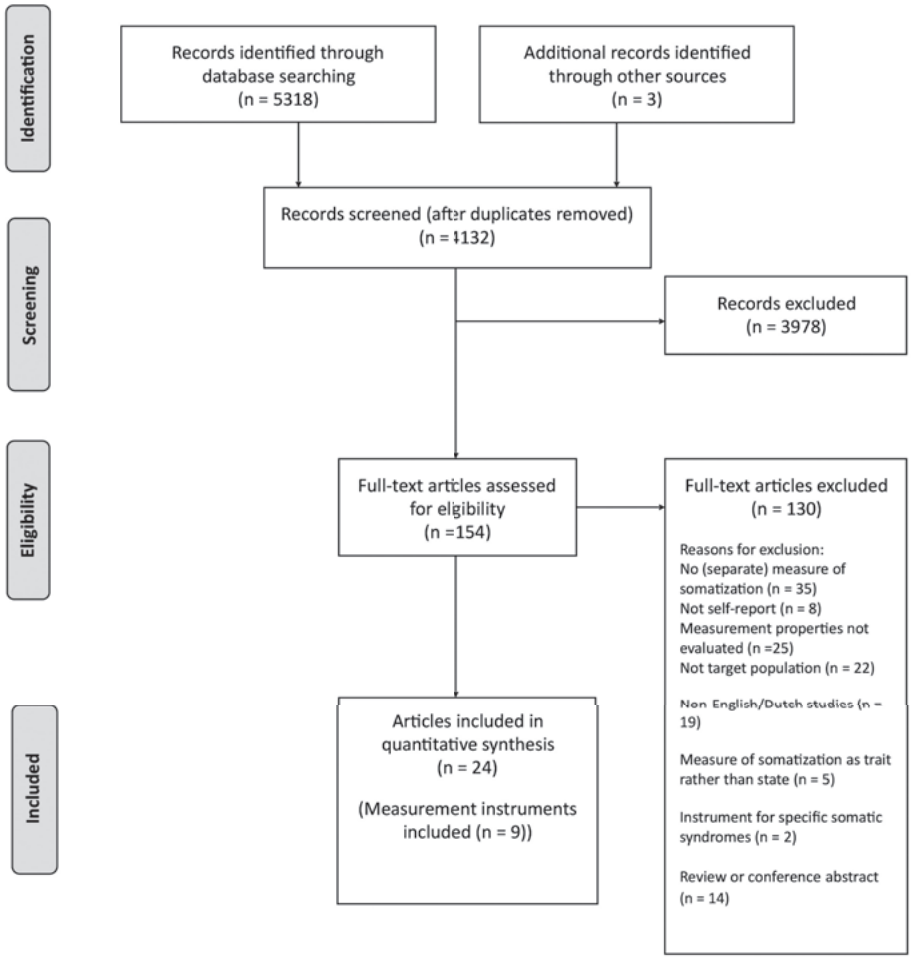


Figure 1. PRISMA Flow Diagram

After removing duplicates, a total of 4129 articles remained. After screening titles and abstracts, 151 articles remained and were assessed for eligibility based on their full-texts. Twenty-one articles were eligible for inclusion. After screening the reference list

of these 21 articles, and of several reviews found in the search, three more eligible articles were identified. This resulted in a total of 24 eligible articles describing various measurement properties of 9 different questionnaires. Most included articles assessed multiple measurement properties of a questionnaire. Internal consistency was assessed in 15 studies, reliability was assessed in 3, measurement error in 1, structural validity in 9, hypotheses testing in 13, cross-cultural validity in 3, criterion validity in 7 and responsiveness in 1. Content validity was not assessed in any of the articles.

The characteristics of the studies are provided in Table 2.

Description of the questionnaires

Table 3 summarizes the general characteristics of the included questionnaires. Seven articles assessed the somatic symptom scale of the Patient Health Questionnaire (PHQ), i.e. the Patient Health Questionnaire-15 (PHQ-15) (29–35), and one study assessed the brief PHQ-r (36), which is an adapted Turkish version of the PHQ. The brief PHQ-r consists of four subscales: somatoform disorder, depressive disorders, panic disorder and functioning of the patient. In our review we only looked at the somatoform disorder subscale. Eight articles assessed the measurement properties of the Four-Dimensional Symptom Questionnaire (4DSQ) somatization subscale (1, 37–43). The 4DSQ consists of 4 subscales: distress, depression, anxiety and somatization. As explained above, we restricted ourselves to the psychometric properties of the subscale somatization. Two articles assessed the Symptom Check-List-90-R (SCL-90-R) somatization subscale (44, 45). The SCL-90-R is a comprehensive questionnaire that aims to measure a broad range of psychological problems. It consists of 9 subscales: somatization, obsessive-compulsive, interpersonal sensitivity, depression, anxiety, anger-hostility, phobic anxiety, paranoid ideation and psychoticism. Again, we only assessed the subscale somatization. The other questionnaires, i.e. the Schedule for Evaluating Persistent Symptoms (SEPS) (46), the Physical Symptom Checklist (PSC-51) (47), the Common Mental Disorders Questionnaire (CMDQ) (48), the Ghent Multidimensional Somatic Complaints Scale (GMSCS) (49), the Screening for Somatoform Symptoms-2 (SOMS-2) (50) and the Bodily Distress Syndrome (BDS) checklist (51) were assessed in one article each. The SEPS records medically unexplained symptoms. The PSC-51 measures somatoform disorders and is based on the DSM-III classification (52), which is an outdated version of the DSM. The CMDQ is a diagnostic tool for common mental disorders and consists of three subscales: somatoform disorder, mental disorder and alcohol dependence. In this review we will only look into the somatoform disorder subscale. The items for the somatoform disorder part are taken from the Symptom Checklist (SCL-90) (53) and the Whiteley index (54). The GMSCS assesses somatic complaints. The SOMS-2 is originally

Table 2. Characteristics of the studies investigating measurement properties of questionnaires measuring somatization

Outcome measure under study	Sample size	Age: mean (SD)	Gender (female)	Setting
PHQ-15 (Becker 2002)	431	Range 18-80	54.1%	Primary care outpatient clinics
PHQ-15 (Kroenke 2002)	3000	46 (17)	66%	Primary care
Brief PHQ-r (Corapcioglu 2004)	1387	28.9 (10.2)	38.2%	Primary health services sites
PHQ-15 (Interian 2006)	172	Hispanic: 39 (±13), non-Hispanic: 43 (±13)	Hispanic: 92%, non-Hispanic: 78%	Primary care: patients with abridged somatization
PHQ-15 (Muramatsu 2007)	131	43.4 (16.4)	59.5%	Primary care and general hospital
PHQ-15 (van Ravesteijn 2009)	904	48	62%	Primary care
PHQ-15 (Körber 2011)	308	47.2 (16.3)	71.4%	Primary care practices
PHQ-15 (Witthöft 2013)	308	47.2 (16.3)	71.4%	Primary care practices
4DSQ somatization subscale (Terluin 1996)	305 (group Y)	Unknown	Unknown	General practices
4DSQ somatization subscale (Terluin 1998)	55	Women: 41.2 (9.9), men; 39.5 (11.4)	53%	Patients with psychological complaints from general practices
4DSQ somatization subscale (Terluin 1999)	382	Women: 40 (18-75), men: 42 (20-72)	61%	Physiotherapy practices

	Country/ language	Method of patient selection	Proportion of missing values
	Saudi Arabia/ Arabic	2-stage design. Patients completed screening instrument and were rated as cases/non-cases, evaluated by primary care physician for somatoform disorder. The first 173 patients (40% of total sample) received psychiatric interviews.	Not reported
	USA/English	Either consecutive or every *n*th patient until intended quota.	Not reported
	Turkey/Turkish	People who presented to the center with a health problem were asked to participate, completed the questionnaire and were interviewed by a psychiatrist.	Not reported
	USA/English	Referred by primary-care physicians and nurses if presented with MUS at least 3 times. Screening and interviews by research staff. First 172 participants.	Not reported
	Japan/Japanese	Patients seeing their physicians for routine medical appointments were asked about participation.	Not reported
	The Netherlands/ Dutch	Patients selected from database or identified by family physician. Patients were approached to participate. Unknown how.	Not reported
	Germany/ German	Consecutive, then at least 25% of patients with minimal somatic symptom severity (SSS), at least 40% with low SSS, at least 70% with medium and at least 80% with high SSS.	Not reported
	Germany/ German	Consecutive.	Not reported
	The Netherlands/ Dutch	On 10 randomly chosen research days, GPs handed out questionnaires to 3460 consecutive patients.	Not reported
	The Netherlands/ Dutch	GPs recruited patients on (a) specific day(s), that was chosen before, on which they had enough time. All patients seen on that day and who met the inclusion criteria were asked to participate.	0.2% on T1 and 0.4% on T2
	The Netherlands/ Dutch	Physiotherapists asked a series of new unselected patients to complete questionnaires as part of their intake.	Not reported

Table 2. Continued

Outcome measure under study	Sample size	Age: mean (SD)	Gender (female)	Setting
4DSQ somatization subscale (Terluin 2006)	Study A: 2127; Study B: 3852; Study C-J: 1424	Study A: 38.5 (11.5), B: 43.9 (8.1), C: 41.9 (8.1), D: 40.4 (10.6), E: 37.8 (12.4), F: 44.8 (15.7), G: 35.0 (8.3), H: 40.7 (12.5), I: 40.2 (10.0), J: 42.5 (12.7)	Study A: 68%, B: 9%, C: 34%, D: 53%, E: 62%, F: 74%, G: 66%, H: 61%, I: 66%, J: 58%	Study A: general practice patients, B: employees responding to a health survey; C: employees with adjustment disorder, D: distressed general practice patients, E: anxious general practice patients, F: GP patients with minor or mild-major depression, G: social work clients, H: physiotherapy patients, I: GP patients with psychological problems, J: GP patients with psychological problems
4DSQ somatization subscale (Czachowski 2012)	516 total. 254 Polish and 262 Dutch subjects. Polish subjects: 142 students and 153 patients	Polish patients: 43.5 (13), Dutch patients: 43.7 (12.1), Polish students: 25.3 (3.9), Dutch students: 24.3 (5.6)	66%	Students and patients of general practice and psychiatric counseling services.
4DSQ somatization subscale (Tebbe 2013)	478 and 478 primary care reference group	30.96 (4.15)	100%	Primary care midwifery practices (versus primary care reference group)
4DSQ somatization subscale (Terluin 2014)	205 English and 302 Dutch	40.2 (13)	75%	Family practices
4DSQ somatization subscale (Chambe 2015)	231 French and 231 Dutch	French: 42.9 (11.7), Dutch: 42.1 (11.6)	71%	General practice patients
SCL-90-R somatization subscale (Schmitz 2000)	447	43 (15)	68.5%	Primary care offices

	Country/ language	Method of patient selection	Proportion of missing values
	The Netherlands/ Dutch	Unknown. Health care providers selected patients.	Not reported
	Poland/ Polish and the Netherlands/ Dutch	Unknown	7.1% respondents incomplete somatization scale, missing values across all questions.
	The Netherlands/ Dutch	Unknown for midwifery group. (General practice patients randomly selected).	No missing data
	Canada/ English and The Netherlands/ Dutch	English-speaking sample approached in the waiting room and invited to complete the questionnaire. The Dutch reference group was randomly selected from a database containing 4DSQ data of patients attending 37 general practitioners.	0.38% (English-speaking group) and 0.42% (Dutch-speaking group)
	France/French and The Netherlands/ Dutch	2 methods. Group 1: questionnaires were given out to consecutive patients; group 2: GPs were asked to recruit patients with suspected psychological problems or MUPS.	0.5% missing values (French), 0.7% missing values (Dutch)
	Germany/ German	Consecutive sample.	Not reported

Table 2. Continued

Outcome measure under study	Sample size	Age: mean (SD)	Gender (female)	Setting
SCL-90-R somatization subscale (Katerndahl 2002)	68	Unknown	80%	Low-income primary care clinic
SEPS (Tyrer 2012)	405 total, of which 73 primary care	Unknown	Unknown	4 clinics in secondary (hospital) care and a small general practice population
PSC-51 (de Waal 2009	1046	Unknown	Unknown	University-affiliated general practices
CMDQ somatoform disorder subscale (Christensen 2005)	1785	38.8	Unknown	General practice
GMSCS (Beirens 2010)	151 (85 primary care and 66 secondary care)	45.6 (range 18-71)	38.4%	General practice (primary care) and pain clinic (secondary care)
(R-)SOMS-2 (Fabiao 2010)	167	43.7 (14.9)	74.3%	General practice
BDS checklist (Budtz-Lilly 2015)	2480	54.3 (17.5)	62.5%	Primary care

4DSQ: Four-Dimensional symptom Questionnaire; BDS checklist: Bodily Distress Syndrome checklist; brief PHQ-r: revised brief Patient Health Questionnaire; CMDQ: Common Mental Disorders Questionnaire; GMSCS: Ghent Multidimensional Somatic Complaints Scale; GP: general practitioner; MUPS: medically unexplained physical symptoms; MUS: medically unexplained symptoms; PHQ-15: the Patient Health Questionnaire 15; PSC-51: Physical Symptom Checklist; (R-)SOMS-2: (revised) Screening for Somatoform Symtptoms-2; SCAN: Schedules for Clinical Assessment in Neuropsychiatry; SCL-90-R: Symptom Check-List-90-R; SCL-SOM: Symptom Check-List Somatization subscale; SD: standard deviation; SEPS: Schedule for Evaluating Persistent Symptoms; SSS: somatic symptom severity

	Country/ language	Method of patient selection	Proportion of missing values
	USA/English	Unknown	Not reported
	UK/English	The patients were not intended to be consecutive ones and research assistants attended on days when available and assessed all patients.	Not reported
	The Netherlands/ Dutch	A random sample of attendees received the screening questionnaires by mail. For each practice the researchers included all consecutive patients on 13-30 arbitrary days within a three-month period using the electronic diaries of the GPs. Then a two-phase selection procedure followed. First patients with a higher risk for depressive/anxiety or somatoform disorders were identified by means of a mental and physical symptom checklist. In the second stage all high-risk patients and a random sample of 15% of the low risk patients were invited for a psychiatric diagnostic interview.	Not reported
	Denmark/Danish	1) GPs asked consecutive patients presenting with a new health problem during a 3-week period to participate; 2) Every ninth patient and all patients with high scores were selected for the SCAN interview.	Relative missing response rate: SCL-SOM = 1.8 (1,4 to 2.4), Whiteley-7 = 2,8 (1,9 to 4,1).
	Belgium/Dutch	Primary care: consecutively recruited. Secondary care: patients filled in questionnaire as part of intake.	Not reported
	Portugal/ Portuguese	During a 10-day period alternating mornings and afternoons all registered persons fulfilling inclusion criteria attending the unit were invited to participate.	Not reported
	Denmark/Danish	Participating GPs registered all patient contacts during one randomly assigned day. In the present study all identifiable patients who had visited their GP because of a health problem were included.	<5%

Table 3. Description of the questionnaires measuring somatization

	Construct	Recall period	Structure of questionnaire
			Dimensions
PHQ-15 (Becker 2002, Kroenke 2002, Interian 2006,Muramatsu 2007, van Ravesteijn 2009, Körber 2011 Witthöft 2013)	Somatic symptom severity	4 weeks	Unidimensional
Brief-PHQ-r (Corapcioglu 2004)	Somatoform disorder	4 weeks	Dimension somatoform disorder = unidimensional (entire questionnaire also consists of dimensions depression and panic disorder, not considered in this review)
4DSQ somatization subscale (Terluin 1996, Terluin 1998, Terluin 1999, Terluin 2006, Czachowski 2012, Tebbe 2013, Terluin 2014, Chambe 2015)	Somatization	1 week	Unidimensional
SCL-90-R somatization subscale (Schmitz 2000, Katerndahl 2002)	Somatization	1 month	Unidimensional
SEPS (Tyrer 2012)	Medically unexplained symptoms	No recall period	Two-dimensional (section 1: general section, section 2: specific section) and a total score
PSC-51 (de Waal 2009)	Physical symptoms	1 week	Unidimensional

	Structure of questionnaire			Time to complete
	Number of items per dimension	**Number and type of response options**	**Scoring**	
	15	3 point Likert scale: 0 (‘not bothered at all’/’not at all’) to 2 (‘bothered a lot’/’more than half the days’)	Each symptom coded as 0, 1, or 2, and the total score ranges from 0 to 30.	Not reported
	17 about somatoform disorder	3-level ordinal scale: ‘‘Not bothered’’, ‘‘bothered a little’’ and ‘‘bothered a lot’’. Also, the history of consulting a doctor for complaints suggestive of an organic disorder and four additional questions concerning previously diagnosed organic disorders.	When at least three of first 13 questions are positive (an answer as ‘‘bothered a lot’’ is accepted as a ‘‘positive’’ response) and there are no organic disorders, a diagnosis of somatoform disorder is made.	Not reported
	16	5: “no”, “sometimes”, “regularly”, “often”, “very often or constantly”.	0 for “no”, 1 for “sometimes” and 2 for the other response categories. The item scores are summated to scale scores (0-32).	Median of 7 min for the entire questionnaire with 4 subscales
	12	5 point Likert scale: 0 ‘not at all’ (o) to ‘extremely’ (4)	Each item is scored 0-4. Item scores within each scale are summed (0-48).	Not reported
	Section 1: 3, section 2: 6, total score: 9	Four point Likert scale (0-3)	Section 1 (SEPS-1): 0-9. Section 2 (SEPS-2): 0-18, total score is the sum of SEPS-1 and SEPS-2, range 0-27.	Not reported
	51 (+4 gender specific items)	The presence of symptoms is rated on a severity scale from 0 to 3 (4-point Likert scale)	A symptom is rated as present for the scores 3 and 3; the total symptom score ranges from 0 to 51. (Gender specific items were excluded).	Not reported

Table 3. Continued

	Construct	Recall period	Structure of questionnaire
			Dimensions
CMDQ, somatoform disorder subscale (Christensen 2005)	Somatoform disorder (the whole questionnaire: common mental disorders)	4 weeks	Two-dimensional (symptom checklist and illness worry)
GMSCS (Beirens 2010)	Somatic complaints	4 weeks	5 dimensions: pain head shoulders, heart chest, stomach abdomen, warm-cold, fatigue
(R-)SOMS-2 (Fabiao 2010)	Somatoform disorders	2 years	Unidimensional
BDS checklist (Budtz-Lilly 2015)	Bodily distress syndrome	4 weeks	Unidimensional

4DSQ: Four-Dimensional symptom Questionnaire; BDS checklist: Bodily Distress Syndrome checklist; brief PHQ-r: revised brief Patient Health Questionnaire; CMDQ: Common Mental Disorders Questionnaire; GMSCS: Ghent Multidimensional Somatic Complaints Scale; PHQ-15: the Patient Health Questionnaire 15; PSC-51: Physical Symptom Checklist; (R-)SOMS-2: (revised) Screening for Somatoform Symtptoms-2; SCL-90-R: Symptom Check-List-90-R; SEPS: Schedule for Evaluating Persistent Symptoms

	Structure of questionnaire			Time to complete
	Number of items per dimension	**Number and type of response options**	**Scoring**	
	Symptom checklist: 12, Illness worry scale: 7	5 point Likert-scale: 0 (‘not at all’) to 4 (‘extremely’)	Answers on single items are dichotomized between 0 (not at all) and 1 (a little) and added to a sum score with corresponding positive predictive values on the separate subscale.	2-5 min (entire questionnaire)
	Pain head shoulders: 3, heart chest: 4, stomach abdomen: 4, warm=cold: 3, fatigue: 4	8-point Likert scale: 0 (never) to 7 (all the time)	Not described	Not reported
	Original SOMS-2: 53 (5 only for women and 1 for men). Portuguese SOMS-2: 46 symptoms, 45 for women and 42 for men. R-SOMS-2: 29	Two response options: yes/no	Not described	Not reported
	30 (version 1) & 25 (version 2)	5 point Likert scale: 0 (‘not at all’) to 4 (‘a lot’)	Not described	Not reported

developed in German (55, 56), but in the article included in this review, the authors investigated an adapted Portuguese version of the SOMS-2 and a shorter Portuguese version, the R-SOMS-2. Finally, the BDS aims to diagnose functional disorders. In six questionnaires somatization was considered to be unidimensional with a single total score, and in three questionnaires it was considered to be multi-dimensional. Three articles on the 4DSQ were written in Dutch (41–43). All other articles were in English. The measurement properties of the PHQ-15, 4DSQ somatization subscale and SCL-90-R somatization subscale were investigated by seven, five and two different research teams, respectively. The measurement properties of the remaining questionnaires were investigated by a single research team each.

Methodological quality of the included studies on measurement properties

The methodological quality of studies investigating internal consistency and criterion validity varied widely from poor to excellent. All studies investigating structural validity and cross-cultural validity were of excellent quality. Most studies investigating hypotheses testing, on the other hand, were of fair quality, because hypotheses were frequently not stated explicitly.

Measurement properties of the questionnaires and evidence rating

Internal consistency

Table 4 provides an overview of the 15 studies that assessed internal consistency (1, 30, 31, 33, 38, 39, 41–46, 49–51). For all questionnaires except the SCL-90-R somatization subscale, SEPS and the (R)-SOMS-2 there was strong evidence for good internal consistency. Results for the SEPS and the (R)-SOMS-2 could not be evaluated due to poor quality of the studies.

Test-retest reliability

Table 5 provides an overview of the 3 studies in which the test-retest reliability was assessed (33, 42, 50). The PHQ-15 and the 4DSQ somatization subscale were the most reliable questionnaires, although the evidence for good test-retest reliability of the PHQ-15 was moderate, whereas it was limited for good test-retest reliability of the 4DSQ. The evidence on the (R)-SOMS-2 could not be interpreted due to poor quality of the study.

Measurement error

Only one study of poor quality, on the 4DSQ somatization subscale, investigated measurement error (1). The study sample consisted of 1424 participants and showed the following results: standard error of measurement (SEM) =2.80, smallest detectable change (SDC) =7.76. However, the SEM was estimated based on the Cronbach's

alpha, which is a method of poor quality according to the COSMIN checklist (19), and consequently, no conclusions about the measurement error could be drawn.

Structural validity

Table 6 summarizes the results for structural validity from the 9 included studies (1, 32, 38, 40, 41, 43, 46, 49, 51). Strong evidence was found for good structural validity of the PHQ-15, the 4DSQ somatization subscale and BDS checklist and for poor structural validity of the GMSCS and the SEPS.

Hypotheses testing

The results from the 13 studies evaluating hypotheses testing (1, 29–31, 37, 41–47, 50) are provided in Table 7. Good construct validity was supported by moderate evidence for the PHQ-15, the 4DSQ somatization subscale and the SCL-90-R somatization subscale. Limited evidence supported good construct validity of the SEPS and the PSC-51. The (R)-SOMS-2 seemed to have poor construct validity due to low sensitivity, however, due to poor quality of the study, the evidence could not be interpreted.

Cross-cultural validity

Cross-cultural validity was assessed in three studies on the 4DSQ somatization subscale (38–40). Translation of the questionnaires was not described in these articles. In the studies English (38), Polish (39) and French (40) versions of the 4DSQ somatization were compared to Dutch samples. Strong evidence supported good cross-cultural validity. In all studies, results showed that the translated versions conveyed the same meaning as the original Dutch version of the questionnaire, i.e. none of the items included in any of these questionnaires showed differential item functioning between language groups. Also, the same cut-off points for determining severity of somatization could be used across language groups.

Criterion validity

Table 8 summarizes results from the 7 studies investigating criterion validity (29, 33–36, 47, 48). Strong evidence was found for good criterion validity of the PHQ(-15) and limited evidence for good criterion validity of the PSC-51. Limited evidence was found for poor criterion validity of the CMDQ somatoform disorder subscale.

Responsiveness

Responsiveness was evaluated in one study (1) of fair quality on the 4DSQ somatization subscale. In this study 86 GP patients with psychosocial problems (age 40.2 (10.0) and 66% female) completed the 4DSQ twice within a, relatively short, time interval of 10 days.

Fifty-nine of these patients answered a 5-point Global Impression (GI) question. The correlation between the somatization change scores and the GI score was weak (r= 0.30 (0.04–0.53)). The patients were then divided into 2 groups: improved and not improved, and receiver operating characteristic (ROC) analyses were performed. The area under the curve (AUC) was 0.69, just below the cut-off of 0.70 for good responsiveness (57). Therefore, there is limited evidence for poor responsiveness.

An overview of the overall evidence rating for all measurement properties of all questionnaires is provided in Table 9.

Table 4. Internal consistency of questionnaires measuring somatization

	Study population	Cronbach's alpha	Methodological quality	Evidence rating*
PHQ-15 (Kroenke 2002)	n=3000	0.80	Poor	+++
PHQ-15 (Interian 2006)	n=172	Hispanic & non-Hispanic both 0.79	Poor	
PHQ-15 (van Ravesteijn 2009)	n=904	0.80	Excellent	
4DSQ somatization subscale (Terluin 1996)	n=305 (group Y)	0.84	Excellent	+++
4DSQ somatization subscale (Terluin 1998)	n=55	T1: 0.88, T2: 0.91	Good	
4DSQ somatization subscale (Terluin 1999)	n=382	0.85	Excellent	
4DSQ somatization subscale (Terluin 2006)	Study C-J: n=1424	0.84	Excellent	
4DSQ somatization subscale (Czachowski 2012)	Total: n=516. Polish: n=254, Dutch: n=262. Polish subjects: n=142 students and n=153 patients	0.82	Fair	
4DSQ somatization subscale (Terluin 2014)	English: n=205, Dutch: n=302	English: 0.82, Dutch: 0.85	Excellent	
SCL-90-R somatization subscale (Schmitz 2000)	n=447	0.83	Poor	?
SCL-90-R somatization subscale (Katerndahl 2002)	n=68	0.81	Poor	
SEPS (Tyrer 2012)	Total: n=405, of which n=73 primary care	0.64	Poor	?
GMSCS (Beirens 2010)	n=151, of which n=85 primary care and n=66 secondary care	Total scale: 0.87. Pain: 0.72, cardio: 0.72, gastro: 0.78, temperature: 0.81, fatigue: 0.87	Excellent	+++
(R-)SOMS-2 (Fabiao 2010)	n=167	0.83	Poor	?
BDS checklist (Budtz-Lilly 2015)	n=2480	All 30 items: 0.92, CP: 0.83, GI: 0.82, MS: 0.86, GS: 0.82	Excellent	+++

4DSQ: Four-Dimensional symptom Questionnaire; BDS checklist: Bodily Distress Syndrome checklist; CP: cardiopulmonary group; GI: gastrointestinal group; GMSCS: Ghent Multidimensional Somatic Complaints Scale; GS: general symptoms; MS: musculoskeletal group; PHQ-15: the Patient Health Questionnaire 15; (R-)SOMS-2: (revised) Screening for Somatoform Symtptoms-2; SCL-90-R: Symptom Check-List-90-R; SEPS: Schedule for Evaluating Persistent Symptoms; * See explanatory Table 1 for levels of evidence.

Table 5. Test-retest reliability of questionnaires measuring somatization

	Study population	Time interval
PHQ-15 (van Ravesteijn 2009)	n=355	14 days
4DSQ somatization subscale (Terluin 1998)	n=51	Mean 1.7 days (0-6), in 90% of the patients 1-3 days
(R-)SOMS-2 (Fabiao 2010)	n=24 (random sample (15%) of total participants)	6 months

4DSQ: Four-Dimensional symptom Questionnaire; ICC: intraclass correlation coefficient; PHQ-15: the Patient Health Questionnaire 15; (R-)SOMS-2: (revised) Screening for Somatoform Symtptoms-2; κ: kappa; * See explanatory table 1

Table 6. Structural validity of questionnaires measuring somatization

	Study population	Model type and type of factor analysis
PHQ-15 (Witthöft 2013)	n=308	Model I: General factor model; CFA Model II: Correlated group factor model; CFA Model III: Hierarchical model; CFA Model IV: Bifactor model; CFA
4DSQ somatization subscale (Terluin 1996)	n=305 (group Y)	4 factor model; EFA
4DSQ somatization subscale (Terluin 1999)	n=382	4 factor model; EFA
4DSQ somatization subscale (Terluin 2006)	Study C-J, n=1424	4 and 5 factor model. 4 factors in accordance with the 4DSQ subscales; CFA
4DSQ somatization subscale (Terluin 2014)	English: n=205, Dutch: n=302	One-factor model (for somatization subscale); CFA
4DSQ somatization subscale (Chambe 2015)	French: n=231, Dutch: n=231	One-factor model (for somatization subscale); CFA

	Reliability coefficient	Methodological quality	Evidence rating*
	ICC=0.83, κ=0.60	Good	++
	r=0.94	Fair	+
	SOMS-2: r=0.67, R-SOMS-2: r=0.69. κ for cut-off points 1-8 ranged from 0.33 to 0.65.	Poor	?

	Factors	Validation results	Methodological quality	Evidence rating*
	Model I: 1: General somatization factor Model II: 4 specific factors: Pain symptoms, Gastroenterological symptoms, Cardio-pulmonary symptoms, Fatigue symptoms Model III: 1 general factor, 4 specific factors Model IV: 1 general factor, 4 specific factors	Model I: $\chi2(65)=356.87$, CFI=0.803, TLI=0.763, RMSEA=0.121, Model fit: poor Model II: $\chi2(59)=160.83$, CFI=0.931, TLI=0.909, RMSEA=0.075, Model fit: poor Model III: $\chi2(61)=160.80$, CFI=0.933, TLI=0.914, RMSEA=0.073, Model fit: poor Model IV: $\chi2(54)=101.39$, CFI=0.968, TLI=0.954, RMSEA=0.053, Model fit: good	Excellent	+++
	4: Depression, Anxiety, Distress, Somatization	15/16 items have a factor loading ≥0.4 on the somatization factor and 4/16 on the anxiety factor.	Excellent	+++
	4: Depression, Anxiety, Distress, Somatization	Factor 1 r=0.15, factor 2 r=0.30, factor 3 r=0.20, factor 4 r=0.91	Excellent	
	4: Depression, Anxiety, Distress, Somatization, and a 5th factor in which items that cross-loaded on Distress and Depression were handled as an additional factor	4 factor model most appropriate. CFI 0.93 (all scales).	Excellent	
	1: Somatization	CFI = 0.985, TLI = 0.983, RMSEA = 0.049	Excellent	
	1: Somatization	CFI = 0.990, TLI = 0.988, RMSEA = 0.046	Excellent	

Table 6. Continued

	Study population	Model type and type of factor analysis
SEPS (Tyrer 2012)	Total: n=405, of which n=73 primary care	2 factor model; EFA
GMSCS (Beirens 2010)	n=151, of which n=85 primary care and n=66 secondary care	Higher order multigroup model; CFA
BDS checklist (Budtz-Lilly 2015)	n=2480	4-factor model; EFA and CFA

4DSQ: Four-Dimensional symptom Questionnaire; BDS checklist: Bodily Distress Syndrome checklist; CFA: confirmatory factor analysis; CFI: comparative fit index; EFA: exploratory factor analysis; GMSCS: Ghent Multidimensional Somatic Complaints Scale; PHQ-15: the Patient Health Questionnaire 15; RMSEA: root mean square error of approximation; SEPS: Schedule for Evaluating Persistent Symptoms; SRMR: standardized root mean square residual; TLI: Tucker-Lewis index; * See explanatory table 1.

	Factors	Validation results	Methodological quality	Evidence rating*
	2: Symptom focus and Attribution	Two components explained 47% of total variance. The first component (symptom focus) explained 29%, the second component (symptom attribution) explained 18%.	Excellent	---
	Higher order multifactorial structure: 1 general somatic complaint factor and 5 specific factors: Pain head shoulders, Heart chest, Stomach abdomen, Warm-cold, Fatigue	χ2=2.34; CFI=0.91; RMSEA=0.061; SRMR=0.073	Excellent	---
	4: Cardiopulmonary / autonomic (arousal) symptoms, Gastrointestinal symptoms, Musculoskeletal symptoms, General symptoms	4 factors were found. 5 of 30 items were removed. Version 2 (25 items): Very good fit. Cut-off 1/2, CFI = 0,958; TLI = 0,954; RMSEA = 0.042 (90%CI 0.039 to 0.046). Cut-off 2/3, CFI = 0,972; TLI = 0,969; RMSEA = 0.028 (90%CI 0.024 to 0.032).	Excellent	+++

Table 7. Hypotheses testing of questionnaires measuring somatization

	Study population	Comparator instrument(s)
PHQ-15 (Kroenke 2002)	n=3000	SF-20 (functional status), disability days, healthcare utilization, symptom related difficulty
PHQ-15 (Interian 2006)	n=172	CIDI Symptom Count
PHQ-15 (Körber 2011)	n=308	n/a
4DSQ somatization subscale (Terluin 1996)	n=305 (group Y)	Other scales of 4DSQ (distress, depression and anxiety) and correlations between 4DSQ dimensions and stress factors and prognosis
4DSQ somatization subscale (Terluin 1998)	n=55	Zung, GHQ, MV, HADS
4DSQ somatization subscale (Terluin 1999)	n=382	Scales Depression, Anxiety and Distress of the 4DSQ, PADS somatization
4DSQ somatization subscale (Terluin 2006)	Study A-J, n=7403 Study A: n= 2127; Somatic reason for encounter: n=1620	4DSQ other scales; SCL-90 Somatization scale; Checklist of psychosocial problems; JCQ; Dutch Neuroticism questionnaire; NEO-FFI; The Pearlin Mastery Scale; the Dutch version of the STAI-State; 5-item social functioning questionnaire; SF-36; single question about sick-leave GPs' diagnoses
4DSQ somatization subscale (Tebbe 2013)	n=478 (and n=478 primary care reference group)	n/a

	Validation results	Methodological quality	Evidence rating*
	SF-20 (all subscales) mean scores decrease as PHQ-15 scores increase (table 3 in Kroenke 2002); Disability days, symptom related difficulty and mean physician visits increase as PHQ-15 scores increase (table 4 in Kroenke 2002).	Fair	++
	Combined: r=0.52, Hispanic: r=0.44, non-Hispanic r=0.68.	Fair	
	PHQ-15 and number of reported symptoms r=0.63. Captures at least 10 of most frequently reported symptoms.	Good	
	Somatization was associated with experiencing health problems r=0.16. Somatization score and somatization score after 2 months (prognosis) r=0,57, somatization score and total 4DSQ score after 2 months r=0.31. Associations with other 4DSQ scales: 0.62 for distress, 0.47 for depression 0.60 for anxiety.	Fair	++
	Zung: r=0.40, GHQ r=0.12, MV r=0.29, HADS Anxiety r=0.35, HADS Depression r=0.22.	Fair	
	Somatization & distress r=0.53, somatization & depression r=0.37, somatization & anxiety r=0.52; PADS somatization r=0.64.	Fair	
	Distress: study A r=0.61, study B r = 0.59, study C-J r=0.46. Depression: study A r=0.43, study B r=0.39, study C-J r=0.35. Anxiety study A r=0.55, study B r=0.50, study C-J r=0.56. SCL-90 somatization r=0.82, p<0.001. Significant relationships with the 4DSQ somatization scale (beta-coefficients): Stress-related measure Decision latitude (JCQ): Study B, -0.07; Neuroticism: Study A 0.22, Study B 0,31; Mastery: Study B: -0.05; SF-36 PCS: Study F: -0.41; Social functioning: Study A: -0.21. AUC = 0.65 (0.63 to 0.68). Cut-off 7: sens=0.60, spec=0.62.	Fair	
	Pregnant and post-partum women scored approximately 1–2 points lower compared to GP-patients (table 4 in Tebbe 2013).	Fair	

Table 7. Continued

	Study population	Comparator instrument(s)
SCL-90-R somatization subscale (Schmitz 2000)	n=409	GHQ-12
SCL-90-R somatization subscale (Katerndahl 2002)	n=68	GHQ-12, SF-36, Duke Scales, Alcohol Use Disorder Identification Test, Brief Social Desirability Scale
SEPS (Tyrer 2012)	Total: n=405, of which n=73 primary care	GP's assessment of unexplained medical symptoms after 18-30 months Health Anxiety Inventory (HAI), HADS, Social Functioning Questionnaire (SFQ)
PSC-51 (de Waal 2009)	n=1046	HADS
(R-)SOMS-2 (Fabiao 2010)	n=167	Diagnosis of clinical somatizers (CS): clinical interview, GP assessment and data form medical records

4DSQ: Four-Dimensional symptom Questionnaire; AUC: area under the curve; CIDI: Composite International Diagnostic Interview; GHQ(-12): General Health Questionnaire (-12); GP: general practitioner; HADS: Hospital Anxiety and Depression Scale; HAI: Health Anxiety Inventory; JCQ: Job Content Questionnaire; MV: Maastrichtse vragenlijst (Maastricht Questionnaire); NEO-FFI: NEO Five Factor Inventory; PADS: PADS klachtenlijst (PADS symptom checklist); PHQ-15: the Patient Health Questionnaire 15; PSC-51: Physical Symptom Checklist; (R-)SOMS-2: (revised) Screening for Somatoform Symtptoms-2; SCL-90(-R): Symptom Check-List-90(-R); Sens: sensitivity; SEPS: Schedule for Evaluating Persistent Symptoms; SF-20: 20-item Short Form Survey; SF-36: Short Form Health Survey; SFQ: Social Functioning Questionnaire; Spec: specificity; STAI-State: State-Trait Anxiety Inventory – State scale; n/a; not applicable; * See explanatory table 1.

	Validation results	Methodological quality	Evidence rating*
	GHQ-12 r=0.52	Fair	++
	GHQ-12 r=0.40 (p≤0.05); SF-36: physical function r=-0.66 (p≤0.05), role-physical r=-0.43 (p≤0.05), pain r=0.43 (p≤0.05), social function r=-0.32 (p≤0.05), mental health r=-0.11, energy r=-0.06, family life changes r=0.47 (p≤0.05), hassles r=0.32 (p≤0.05); Duke Scales stress r=0.23; Alcohol Use Disorder Identification Test r=0.25 (p≤0.05); Brief Social Desirability Scale r=0.05.	Fair	
	AUC 0.63 (0,55 to 0,72). Cut-off 14, with sens = 0.65 and spec = 0.53. HAI r=0.40, HADS anxiety r=-0.35, HADS depression r=0.35, SFQ r=0.26, p<.001.	Fair	+
	HADS - PSC-51: r=0.6, p<.01	Fair	+
	Optimal cut-off = 4/29 symptoms. SOMS-2: sens=57,9, spec=88.2; R-SOMS-2: sens=56.1, spec=93.6.	Poor	?

Table 8. Criterion validity of questionnaires measuring somatization

	Study population	Gold standard
PHQ-15 (Becker 2002)	n=173, 55.5% female	Structured Clinical Interview (SCID-R)
Brief PHQ-r (Corapcioglu 2004)	n=1387	Diagnoses by psychiatrists using the DSM-IV
PHQ-15 (Muramatsu 2007)	n=131	*Mini-International Neuropsychiatric Interview-Plus (MINI-Plus)*
PHQ-15 (van Ravesteijn 2009)	n=904	Structured Clinical Interview for the Diagnostic and Statistical Manual-IV (DSM-Iv) Axis I disorders (SCID-I)
PHQ-15 (Körber 2011)	n=308	Structured diagnostic interview
PSC-51 (de Waal 2009)	n=473	Schedules for Clinical Assessment in Neuropsychiatry (SCAN 2.1)
CMDQ somatoform disorder subscale (Christensen 2005)	n=701	Schedules for Clinical Assessment in Neuropsychiatry (SCAN)

AUC: area under the curve; Brief PHQ-r: revised brief Patient Health Questionnaire; CMDQ: Common Mental Disorders Questionnaire; DSM-IV: Diagnostic and Statistical Manual of Mental Disorders 4th edition; MINI Plus: Mini-International Neuropsychiatric Interview-Plus; PHQ-15: the Patient Health Questionnaire 15; PSC-51: Physical Symptom Checklist; ROC: Receiver Operating Characteristic; SCID-I: Structured Clinical Interview for DSM-IV Axis I disorders; SCID-R: Structured Clinical Interview for DSM disorders; SCL-SOM: Symptom Check-List Somatization subscale; Sens: sensitivity; Spec: specificity; κ: kappa; * See explanatory table 1.

Validation results	Methodological quality	Evidence rating*
κ = 0.65, sens = 0.65, spec = 0.96	Fair	+++
κ = 0.476, sens = 61.9%, spec = 92.5%	Excellent	
κ = 0.85, sens = 1.0, spec = 0.95. pos. predictive value = 0.78, neg. predictive value = 1.0	Excellent	
Optimal sum of sens and spec = 3. Accuracy = fair, AUC = 0.76	Excellent	
AUC = 0.76. Optimal cutoff of =10. (Point biserial r = 0.41 (p<0.01))	Good	
AUC = 0.80 (0,76 to 0,83). No optimal cut-off emerged from ROC curves. Cut-off 5: sens = 0,78, spec = 0,64	Fair	+
SCL-SOM cut-off score 5 - any somatoform disorder: sens = 0.65, spec = 0,63. Whiteley-7 cut-off score 2 - any somatoform disorder: sens = 0.74, spec = 0.56. AUC for SCL-SOM and Whiteley-7 ranged from 0.67 to 0.78 with highest values for somatization disorder	Fair	-

Table 9: Overall evidence rating

	PHQ(-15)	4DSQ somatization subscale	SCL-90-R somatization subscale	SEPS
Internal consistency	**+++**	**+++**	**?**	**?**
Test-retest reliability	**++**	**+**		
Measurement error		**?**		
Structural validity	**+++**	**+++**		**- - -**
Hypotheses testing	**++**	**++**	**++**	**+**
Cross-cultural validity		**Translation: ?** **Validation: +++**		
Criterion validity	**+++**			
Responsiveness		**+**		

+++/- - -: strong level of positive/negative evidence; ++/- -: moderate level of positive/negative evidence; +/-: limited level of positive/negative evidence;
?: unknown due to poor quality of the studies

	PSC-51	CMDQ somatoform disorder subscale	GMSCS	(R)-SOMS-2	BDS checklist
			+++	?	+++
				?	
			- - -		+++
	+			?	
	+	-			

DISCUSSION

Summary of evidence

We identified 24 articles on various measurement properties of 9 self-report questionnaires measuring somatization in the primary care setting. The PHQ-15 and the 4DSQ somatization subscale were described the most, in 8 articles each. The SCL-90-R somatization subscale was described in two articles and the remaining questionnaires were described in only one article per questionnaire, which weakens the level of evidence for the results.

The PHQ-15 and the 4DSQ somatization subscale show the strongest evidence for positive results on a broad array of measurement properties. The quality of the measurement properties of the PHQ-15 and 4DSQ somatization subscale outweighs each other, so a preference can be based on practical considerations. The two questionnaires have nearly the same number of items, however the PHQ-15 enquires about symptoms in the previous four weeks, whereas the 4DSQ has a recall period of one week. Having to recall symptoms over a longer period of time could cause more recall bias. However, reporting symptoms from the previous week only, could possibly leave out important information about symptoms that were present previously, but by chance were less prominent in the past week. Or, on the contrary, transient symptoms that may not point to somatization could be detected and seem more prominent than possibly necessary when using the 4DSQ. Also, the PHQ-15 includes two items enquiring about symptoms linked to menstruation and sexual intercourse. The 4DSQ does not include items on these topics. The choice for one of the two instruments could therefore depend on the patient population. For instance, when using a questionnaire with female patients within their reproductive age range, the PHQ-15 could provide useful information. With older patients, for instance women after having reached menopause, the 4DSQ could possibly be more suitable. A health care provider interested in screening for the DSM-IV somatoform disorder may opt for the PHQ-15, as this instrument has been compared against this diagnosis. On the other hand, the 4DSQ somatization subscale may be more suitable for Polish, French and Dutch speaking patients due to its validation within these population groups.

A promising questionnaire is the BDS checklist. Although its measurement properties were only investigated in one study, strong evidence was found for good internal consistency and structural validity. However, more research is needed to investigate the quality of the remaining measurement properties.

Based on the studies included in this review the SEPS, GMSCS and the CMDQ somatoform disorder subscale seem less suitable for measuring somatization in primary care, due to poor structural validity of the first two instruments, and poor criterion validity of the latter, which was found in studies of excellent and fair quality. However, more research of these questionnaires is still needed to be able to draw more solid conclusions about their quality.

The remaining questionnaires, the SCL-90-R somatization subscale, PSC-51 and the (R)-SOMS-2 have been described in a small number of articles where only two measurement properties were assessed. The studies on the (R)-SOMS-2 were of poor quality, therefore, no conclusions were drawn about it in this review. Limited to moderate evidence supported findings of several good measurement properties of the SCL-90-R somatization scale and the PSC-51. Studies on the SCL-90-R somatization subscale in other populations (general population and various secondary care patients) show acceptable to good measurement properties (16, 58–60). So once again, more research on these questionnaires in primary care is needed. A point of consideration is that the PSC-51 and the (R)-SOMS-2 consist of 51 and 46 (or 29 in the short version) items, respectively, and are therefore time-consuming. This may be less suitable for a primary care setting due to the time constraints faced by GPs and additionally being more burdensome for patients.

Embedding in existing literature

One previous review provided an overview of diagnostic measurement instruments for somatoform disorders (18). However, this review only focused on the assessment of somatoform disorders, which are at the most severe spectrum of symptoms of somatization and therefore their findings only cover part of the broad range of symptom severity that is encountered in primary care. Also, although measurement properties of the measurement instruments were mentioned, no structured, thorough evaluation of measurement properties was made.

Another article provided an overview of self-report questionnaires for common somatic symptoms for use in large-scale epidemiological studies in any type of population, so not specifically for primary care (17). The authors recommend the use of the PHQ-15 and the SCL-90 somatization subscale. However, their aim was to determine which questionnaire was most suitable for research purposes in large-scale epidemiological studies. The authors focussed on applicability in largescale studies by examining low burden to participants, suitability for investigators with no specific expertise in the assessment of somatization, and relevance for future studies in terms of symptom types

and language. In addition, two important measurement properties, i.e. measurement error and responsiveness were not taken into account in their review. In our review we also found positive results for the PHQ-15 in primary care, but less so for the SCL-90-R somatization subscale.

Strengths and limitations

The most important strength of this study is that we used the COSMIN taxonomy for deciding which measurement properties were assessed, and that we took the methodological quality of each individual measurement property into account when interpreting the results of the studies, and drawing conclusions on the quality of the included questionnaires. The COSMIN methodology provided a structured tool for assessing all questionnaires in a consistent way.

We slightly modified the COSMIN standard, by omitting the missing data items from the 4-point rating scale. This modification had consequences for the evidence rating of the questionnaires: the results were of stronger quality due to the modification and stronger conclusions could be drawn because of this. This was especially the case for internal consistency, structural validity and criterion validity. Results were stronger after the modification because in order to determine the overall score of the methodological quality of a certain measurement property, a COSMIN box with standards for the corresponding measurement property is completed (19). Each box consists of a number of items covering the standards of methodological quality. Each item is rated from poor to excellent. The 'worst score counts' method is applied in order to determine the overall score of the methodological quality. So the entire box is scored 'poor' if only one item in the box is scored as 'poor'. Therefore, if the item on missing data was scored as poor, the entire box was scored as 'poor' as well, leading to weaker overall methodological quality results, even if no other item in the box has been scored 'poor'. Omitting the items on missing data, therefore led to results of stronger quality.

A point of consideration is our definition of the term 'somatization', in which we only focussed on the presence and severity of multiple physical symptoms and their burden for the patient. The widely used definition by Lipowski (2) also incorporates cognitions and behavior of the patients with regard to their multiple physical symptoms. However, as virtually no questionnaire enquires about all those aspects simultaneously, it was impossible to include questionnaires measuring somatization according to Lipowski's definition. Therefore, we chose to focus on the measurement of experienced multiple physical symptoms and their severity which can be used as a proxy for somatization. However, to cover the entire original definition of somatization, future studies could take

the psychological and behavioural aspects of somatization into account as well. This could for instance be done by combining one of the somatic symptom questionnaires such as the PHQ-15 or the somatization subscale of the 4DSQ with another questionnaire that assesses the psychological and behavioural aspects of somatization, such as the recently developed and validated Somatic Symptom Disorder- B Criteria Scale (SSD-12) (61, 62), which measures the psychological features of the DSM-5 somatic symptom disorder.

Another point of consideration is that there is no true gold standard for somatization. A relatively objective measure that approaches a gold standard is a diagnosis of a somatoform disorder according to the DSM-IV criteria. We therefore referred to these criteria as the gold standard in this review as well. However, it must be noted that these diagnoses only cover the extreme end of the spectrum of somatization. As there is in fact no gold standard for somatization, the comparison with the DSM-IV criteria could be considered hypotheses testing rather than criterion validity.

A third limitation is that we did not include grey literature such as dissertations and conference abstracts. This choice may have contributed to selection bias. Also, due to practical reasons we did not use indirect evidence from studies in which the questionnaires were actually used. Finally, we excluded full-text articles that were written in a language other than English or Dutch. As we chose to include Dutch as well as English articles, this could possibly have contributed to the fact that more studies were found about Dutch questionnaires such as the 4DSQ. Therefore, this questionnaire may have been overrepresented and this might lead to a possible bias in our conclusive recommendation about this questionnaire. Furthermore, we did not take the different questionnaires' year of origin into account. Therefore, it has not yet been possible to investigate more recently developed questionnaires such as the BDS checklist in different settings or languages, as has been the case with the earlier developed questionnaires.

A final point of consideration is that we chose to limit our search to articles on the quality of questionnaires studied in primary care. Questionnaires studied in other populations, such as the general population, community samples, students, secondary care, have therefore been excluded, as measurement properties of questionnaires may be different in different populations and settings. However, it is possible that some questionnaires could be useful, but have not yet been studied in primary care, such as the Somatic Symptom Index (SSI) (63), the SOMS-7 (64) or the Somatic Symptom Scale-8 (SSS-8) (65). The latter questionnaire for instance performs similarly to the PHQ-15 in secondary care and has less items. Also, more studies are available on the SCL-90-R somatization

subscale in the general population, secondary care and psychiatric populations (16, 58–60, 66). Validation of these questionnaires in primary care may yield interesting information.

CONCLUSIONS

Based on our overview, we recommend using the PHQ-15 or 4DSQ for measuring somatization in primary care. The choice of a preferred questionnaire can differ depending on the measurement properties that have priority to the user of the questionnaire. Health care providers interested in the closest approximation to a somatoform disorder may favour the PHQ-15, as it had the best results for criterion validity, whereas health care providers seeking the best questionnaire for Polish, French or Dutch-speaking patients may choose the 4DSQ somatization subscale instead. If information on menstruation or sexual intercourse complaints is relevant, the PHQ-15 can be used, but if these items are less important, the 4DSQ may be more suitable. Other questionnaires, such as the BDS checklist, SCL-90-R somatization subscale and PSC-51 could benefit from further study primary care. However, the BDS checklist and the PSC-51 consist of a larger number of items than the PHQ-15 and 4DSQ somatization subscale and are therefore more time-consuming. Finally, measurement properties such as measurement error, content validity, cross-cultural validity and responsiveness should be studied further in all questionnaires using sound research methods.

REFERENCES

1. Terluin B, van Marwijk HW, Ader HJ, de Vet HC, Penninx BW, Hermens ML, et al. The Four-Dimensional Symptom Questionnaire (4DSQ): a validation study of a multidimensional self-report questionnaire to assess distress, depression, anxiety and somatization. BMC Psychiatry. 2006;6:34.
2. Lipowski ZJ. Somatization: the concept and its clinical application. Am J Psychiatry. 1988;145:1358-68.
3. Arnold IA, de Waal MW, Eekhof JA, van Hemert AM. Somatoform disorder in primary care: course and the need for cognitive-behavioral treatment. Psychosomatics. 2006;47(6):498-503.
4. de Waal MWM, Arnold IA, Eekhof JAH, Van Hemert AM. Somatoform disorders in general practice: prevalence, functional impairment and comorbidity with anxiety and depressive disorders. Br J Psychiatry. 2004;184(6):470-476.
5. APA. Diagnostic and Statistical Manual of Mental Disorders, 4th ed., (DSM-IV) 1994, American Psychiatric Association; Washington, DC.
6. Diagnostic and Statistical Manual of Mental Disorders, 5th ed. (DSM-5) 2013, American Psychiatric Publishing; Arlington, VA.
7. Jackson JL, Kroenke K. Prevalence, impact, and prognosis of multisomatoform disorder in primary care: a 5-year follow-up study. Psychosom Med. 2008;70(4):430-434.
8. Rosendal M, Olde Hartman TC, Aamland A, van der Horst H, Lucassen P, Budtz-Lilly A, et al. "Medically unexplained" symptoms and symptom disorders in primary care: prognosis-based recognition and classification. BMC Fam Pract. 2017;18(1):18.
9. Murray AM, Toussaint A, Althaus A, Lowe B. The challenge of diagnosing non-specific, functional, and somatoform disorders: a systematic review of barriers to diagnosis in primary care. J Psychosom Res. 2016;80:1-10.
10. Crombez G, Beirens K, Van Damme S, Eccleston C, Fontaine J. The unbearable lightness of somatisation: a systematic review of the concept of somatisation in empirical studies of pain. Pain. 2009;145(1-2):31-35.
11. olde Hartman TC, Borghuis MS, Lucassen PL, van de Laar FA, Speckens AE, van Weel C. Medically unexplained symptoms, somatisation disorder and hypochondriasis: course and prognosis. A systematic review. J Psychosom Res. 2009;66(5):363-377.
12. Tomenson B, Essau C, Jacobi F, Ladwig KH, Leiknes KA, Lieb R, et al. Total somatic symptom score as a predictor of health outcome in somatic symptom disorders. Br J Psychiatry. 2013;203(5):373-380.
13. Lowe B, Spitzer RL, Williams JB, Mussell M, Schellberg D, Kroenke K. Depression, anxiety and somatization in primary care: syndrome overlap and functional impairment. Gen Hosp Psychiatry. 2008;30(3):191-199.
14. Bierke S, Haner M, Petersen W. Influence of somatization and depressive symptoms on the course of pain within the first year after uncomplicated total knee replacement: a prospective study. Int Orthop. 2016;40(7):1353-1360.
15. Haggarty JM, O'Connor BP, Mozzon JB, Bailey SK. Shared mental healthcare and somatization: changes in patient symptoms and disability. Prim Health Care Res Dev. 2016;17(3):277-286.
16. van Driel TJW, Hilderink PH, Hanssen DJC, de Boer P, Rosmalen JGM, Oude Voshaar RC. Assessment of somatization and medically unexplained symptoms in later life. Assessment. 2017:DOI: 10.1177/1073191117721740.
17. Zijlema WL, Stolk RP, Lowe B, Rief W, BioShaRe, White PD, et al. How to assess common somatic symptoms in large-scale studies: a systematic review of questionnaires. J Psychosom Res. 2013;74(6):459-468.
18. Hiller W, Janca A. Assessment of somatoform disorders: a review of strategies and instruments. Acta Neuropsychiatr. 2003;15(4):167-179.
19. Mokkink LB, Terwee CB, Patrick DL, Alonso J, Stratford PW, Knol DL, et al. The COSMIN checklist for assessing the methodological quality of studies

on measurement properties of health status measurement instruments: an international Delphi study. Qual Life Res. 2010;19(4):539-549.
20. Mokkink LB, Terwee CB, Patrick DL, Alonso J, Stratford PW, Knol DL, et al. The COSMIN study reached international consensus on taxonomy, terminology, and definitions of measurement properties for health-related patient-reported outcomes. J Clin Epidemiol. 2010;63(7):737-745.
21. Moher D, Liberati A, Tetzlaff J, Altman DG, Group P. Preferred reporting items for systematic reviews and meta-analyses: the PRISMA statement. J Clin Epidemiol. 2009;62(10):1006-1012.
22. Terwee CB, Jansma EP, Riphagen LL, de Vet HC. Development of a methodological PubMed search filter for finding studies on measurement properties of measurement instruments. Qual Life Res. 2009;18(8):1115-1123.
23. Terwee CB, Mokkink LB, Knol DL, Ostelo RW, Bouter LM, de Vet HC. Rating the methodological quality in systematic reviews of studies on measurement properties: a scoring system for the COSMIN checklist. Qual Life Res. 2012;21(4):651-657.
24. WHO. SCAN. Schedules for Clinical Assessment in Neuropsychiatry, version 2.1. Geneva: World Health Organization, Division of Mental Health; 1998.
25. First MB, Spitzer RL, M. G, Williams JBW. Structured Clinical Interview for DSM-IV-TR Axis I Disorders, Research Version, Patient Edition. New York: Biometrics Research, New York State Psychiatric Institute; 2010.
26. Lecrubier Y, Sheehan DV, Weiller E, Amorim P, Bonora I, Harnett Sheehan K, et al. The Mini International Neuropsychiatric Interview (MINI). A short diagnostic structured interview: reliability and validity according to the CIDI. Eur Psychiatry. 1997;12:224-231.
27. Prinsen CA, Vohra S, Rose MR, Boers M, Tugwell P, Clarke M, et al. How to select outcome measurement instruments for outcomes included in a "Core Outcome Set" - a practical guideline. Trials. 2016;17(1):449.
28. van Tulder M, Furlan A, Bombardier C, Bouter LM. Updated method guidelines for systematic reviews in the Cochrane collaboration back review group. Spine (Phila Pa 1976). 2003;28(12):1290-1299.
29. Korber S, Frieser D, Steinbrecher N, Hiller W. Classification characteristics of the Patient Health Questionnaire-15 for screening somatoform disorders in a primary care setting. J Psychosom Res. 2011;71(3):142-147.
30. Kroenke K, Spitzer RL, Williams JB. The PHQ-15: Validity of a new measure for evaluating the severity of somatic symptoms. Psychosom Med. 2002;64:258-266.
31. Interian A, Allen LA, Gara MA, Escobar JI, Diaz-Martinez AM. Somatic complaints in primary care: further examining the validity of the Patient Health Questionnaire (PHQ-15). Psychosomatics. 2006;47(5):392-398.
32. Witthoft M, Hiller W, Loch N, Jasper F. The latent structure of medically unexplained symptoms and its relation to functional somatic syndromes. Int J Behav Med. 2013;20(2):172-183.
33. van Ravesteijn H, Wittkampf K, Lucassen P, van de Lisdonk E, van den Hoogen H, van Weert H, et al. Detecting somatoform disorders in primary care with the PHQ-15. Ann Fam Med. 2009;7(3):232-238.
34. Becker S, Al Zaid K, Al Faris E. Screening for somatization and depression in Saudi Arabia: a validation study of the PHQ in primary care. Int J Psychiatry Med. 2002;32(3):271-283.
35. Muramatsu K, Miyaoka H, Kamijima K, Muramatsu Y, Yoshida M, Otsubo T, et al. The Patient Health Questionnaire, Japanese version: validity according to the MINI-International Neuropsychiatric Interview-Plus. Psychol Rep. 2007;101:952-960.
36. Corapcioglu A, Ozer GU. Adaptation of revised Brief PHQ (Brief-PHQ-r) for diagnosis of depression, panic disorder and somatoform disorder in primary healthcare settings. Int J Psychiatry Clin Pract. 2004;8(1):11-18.
37. Tebbe BB, Terluin B, Koelewijn JM. Assessing psychological health in midwifery practice: a validation study of the Four-Dimensional Symptom Questionnaire (4DSQ), a Dutch primary care instrument.

Midwifery. 2013;29(6):608-615.
38. Terluin B, Smits N, Miedema B. The English version of the four-dimensional symptom questionnaire (4DSQ) measures the same as the original Dutch questionnaire: a validation study. Eur J Gen Pract. 2014;20(4):320-326.
39. Czachowski S, Terluin B, Izdebski A, Izdebski P. Evaluating the cross-cultural validity of the Polish version of the Four-Dimensional Symptom Questionnaire (4DSQ) using differential item functioning (DIF) analysis. Fam Pract. 2012;29(5):609-615.
40. Chambe J, Le Reste JY, Maisonneuve H, Sanselme AE, Oho-Mpondo J, Nabbe P, et al. Evaluating the validity of the French version of the Four-Dimensional Symptom Questionnaire with differential item functioning analysis. Fam Pract. 2015;32(4):474-479.
41. Terluin B. Factoranalyse van drie klachtenlijsten [Factor analysis of three questionnaires]. Tijdschrift voor Psychosomatische Fysiotherapie. 1999;4:25-31.
42. Terluin B. Wat meet de Vierdimensionale Klachtenlijst (4DKL) in vergelijking met enkele bekende klachtenlijsten? [What does the Four-Dimensional Symptom Questionnaire (4DSQ) measure compared to several known questionnaires?]. Tijdschrift voor Gezondheidswetenschappen. 1998;76:435-441.
43. Terluin B. De Vierdimensionale Klachtenlijst (4DKL). Een vragenlijst voor het meten van distress, depressie, angst en somatisatie. [The Four-Dimensional Symptom Questionnaire (4DSQ). A questionnaire for measuring distress, depression, anxiety and somatization]. Huisarts Wet. 1996;39(12):538-547.
44. Schmitz N, Hartkamp N, Kiuse J, Franke GH, Reister G, Tress W. The Symptom-Check-List-90-R (SCL-90-R): AsGerman validation study. Qual Life Res. 2000;9:185-193.
45. Katerndahl DA, Amodei N, Larme AC, Palmer R. Psychometric assessment of measures of psychological symptoms, functional status, life events, and context for low income Hispanic patients in a primary care setting. Psychol Rep. 2002;91:1121-1128.
46. Tyrer H, Ali L, Cooper F, Seivewright P, Bassett P, Tyrer P. The Schedule for Evaluating Persistent Symptoms (SEPS): a new method of recording medically unexplained symptoms. Int J Soc Psychiatry. 2013;59(3):281-287.
47. de Waal MW, Arnold IA, Spinhoven P, Eekhof JA, Assendelft WJ, van Hemert AM. The role of comorbidity in the detection of psychiatric disorders with checklists for mental and physical symptoms in primary care. Soc Psychiatry Psychiatr Epidemiol. 2009;44(1):78-85.
48. Christensen KS, Fink P, Toft T, Frostholm L, Ornbol E, Olesen F. A brief case-finding questionnaire for common mental disorders: the CMDQ. Fam Pract. 2005;22(4):448-457.
49. Beirens K, Fontaine JR. Development of the Ghent Multidimensional Somatic Complaints Scale. Assessment. 2010;17(1):70-80.
50. Fabiao C, Silva MC, Barbosa A, Fleming M, Rief W. Assessing medically unexplained symptoms: evaluation of a shortened version of the SOMS for use in primary care. BMC Psychiatry. 2010;10:34.
51. Budtz-Lilly A, Fink P, Ornbol E, Vestergaard M, Moth G, Christensen KS, et al. A new questionnaire to identify bodily distress in primary care: the 'BDS checklist'. J Psychosom Res. 2015;78(6):536-545.
52. APA. Diagnostic and statistical manual of mental disorders, 3d edition. Washington, DC: American Psychiatric Association; 1980.
53. Derogatis LR. SCL-90-R. Administration, Scoring and Procedures. MANUAL-II. Towson, MD: Clinical Psychometric Research; 1983.
54. Fink P, Ewald H, Jensen J, Sorensen L, Engberg M, Holm M, et al. Screening for somatization and hypochondriasis in primary care and neurological in-patients: a seven-item scale for hypochondriasis and somatization. J Psychosom Res. 1999;46(3):261-273.
55. Rief W, Hessel A, Braehler E. Somatization symptoms and hypochondriacal features in the general population. Psychosom Med. 2001;63:595-602.

56. Rief W, Hiller W. Toward empirically based criteria for the cllassification of somatoform disorders. J Psychosom Res. 1999;46(6):507-518.
57. de Vet HC, Terwee CB, Mokkink LB, Knol DL. Measurement in medicine, a practical guide. Cambridge, United Kingdom: Cambridge University Press; 2015.
58. Smits IAM, Timmerman ME, Barelds DPH, Meijer RR. The Dutch Symptom Checklist-90-Revised. Is the use of the subscales justified? Eur J Psychol Assess. 2014;31(4):263–271.
59. Sereda Y, Dembitskyi S. Validity assessment of the symptom checklist SCL-90-R and shortened versions for the general population in Ukraine. BMC Psychiatry. 2016;16:300.
60. Tomioka M, Shimura M, Hidaka M, Kubo C. The reliability and validity of a Japanese version of Symptom Checklist 90 Revised. Biopsychosoc Med. 2008;2:19.
61. Toussaint A, Lowe B, Brahler E, Jordan P. The Somatic Symptom Disorder - B Criteria Scale (SSD-12): factorial structure, validity and population-based norms. J Psychosom Res. 2017;97:9-17.
62. Toussaint A, Murray AM, Voigt K, Herzog A, Gierk B, Kroenke K, et al. Development and validation of the Somatic Symptom Disorder-B Criteria Scale (SSD-12). Psychosom Med. 2016;78(1):5-12.
63. Escobar JI, Rubio-Stipec MS, Canino G, Karno M. Somatic Symptom Index (SSI): a new and abridged somatization construct. J Nerv Ment Dis. 1989;177(3):140-146.
64. Rief W, Hiller W. A new approach to the assessment of the treatment effects of somatoform disorders. Psychosomatics. 2003;44(6):492-498.
65. Gierk B, Kohlmann S, Toussaint A, Wahl I, Brunahl CA, Murray AM, et al. Assessing somatic symptom burden: a psychometric comparison of the patient health questionnaire-15 (PHQ-15) and the somatic symptom scale-8 (SSS-8). J Psychosom Res. 2015;78(4):352-355.
66. Hart DL, Werneke MW, George SZ, Deutscher D. Single-item screens identified patients with elevated levels of depressive and somatization symptoms in outpatient physical therapy. Qual Life Res. 2012;21(2):257-268.

Appendix A. Search strategies

Search strategy for Pubmed/Medline

1. Construct

(Somatization[tw] OR somatisation[tw] OR somatoform[tw] OR hysteri*[tw] OR briquet[tw] OR polysymptom*[tw] OR multisomatoform[tw] OR somatizer*[tw] OR somatic symptom*[tw] OR MUPS[tw] OR medical unexplained[tw] OR medically unexplained[tw] OR unexplained medical*[tw] OR unexplained symptom*[tw] OR unexplained syndrom*[tw] OR frequent attend*[tw] OR "multiple physical symptom"[tw] OR "multiple physical symptoms"[tw] OR "multiple symptom diagnosed"[tw] OR "multiple symptom diagnoses"[tw] OR "multiple symptom diagnosis"[tw] OR "multiple symptom diagnostics"[tw] OR neurastheni*[tw])

AND

2. Population

("family"[all fields] OR physician*[all fields] OR practice*[tw] OR "primary care"[all fields] OR "Primary Health Care"[mh] OR primary[tw] OR general pract*[tiab] OR gp[tiab] OR gps[tiab])

AND

3. Measurement instruments

(HR-PRO[tiab] OR HRPRO[tiab] OR HRQL[tiab] OR HRQoL[tiab] OR QL[tiab] OR QoL[tiab] OR quality of life[tw] OR life quality[tw] OR health index*[tiab] OR health indices[tiab] OR health profile*[tiab] OR health status[tw] OR ((patient[tiab] OR self[tiab] OR child[tiab] OR parent[tiab] OR carer[tiab] OR proxy[tiab]) AND ((report[tiab] OR reported[tiab] OR reporting[tiab]) OR (rated[tiab] OR rating[tiab] OR ratings[tiab]) OR based[tiab] OR (assessed[tiab] OR assessment[tiab] OR assessments[tiab]))) OR ((disability[tiab] OR function[tiab] OR functional[tiab] OR functions[tiab] OR subjective[tiab] OR utility[tiab] OR utilities[tiab] OR wellbeing[tiab] OR well being[tiab]) AND (index[tiab] OR indices[tiab] OR instrument[tiab] OR instruments[tiab] OR measure[tiab] OR measures[tiab] OR questionnaire[tiab] OR questionnaires[tiab] OR profile[tiab] OR profiles[tiab] OR scale[tiab] OR scales[tiab] OR score[tiab] OR scores[tiab] OR status[tiab] OR survey[tiab] OR surveys[tiab])))

AND

4. Measurement properties

(instrumentation[sh] OR methods[sh] OR "Validation Studies"[pt] OR "Comparative Study"[pt] OR "psychometrics"[MeSH] OR psychometr*[tiab] OR clinimetr*[tw] OR clinometr*[tw] OR "outcome assessment (health care)"[MeSH] OR "outcome assessment"[tiab] OR "outcome measure*"[tw] OR "observer variation"[MeSH] OR "observer variation"[tiab] OR "Health Status Indicators"[Mesh] OR "reproducibility of results"[MeSH] OR reproducib*[tiab] OR "discriminant analysis"[MeSH] OR reliab*[tiab] OR unreliab*[tiab] OR valid*[tiab] OR "coefficient of variation"[tiab] OR coefficient[tiab] OR homogeneity[tiab] OR homogeneous[tiab] OR "internal consistency"[tiab] OR (cronbach*[tiab] AND (alpha[tiab] OR alphas[tiab])) OR (item[tiab] AND (correlation*[tiab] OR selection*[tiab] OR reduction*[tiab])) OR agreement[tw] OR precision[tw] OR imprecision[tw] OR "precise values"[tw] OR test-retest[tiab] OR (test[tiab] AND retest[tiab]) OR (reliab*[tiab] AND (test[tiab] OR retest[tiab])) OR stability[tiab] OR interrater[tiab] OR inter-rater[tiab] OR intrarater[tiab] OR intra-rater[tiab] OR intertester[tiab] OR inter-tester[tiab] OR intratester[tiab] OR intra-tester[tiab] OR interobserver[tiab] OR inter-observer[tiab] OR intraobserver[tiab] OR intra-observer[tiab] OR intertechnician[tiab] OR inter-technician[tiab] OR intratechnician[tiab] OR intra-technician[tiab] OR interexaminer[tiab] OR inter-examiner[tiab] OR intraexaminer[tiab] OR intra-examiner[tiab] OR interassay[tiab] OR inter-assay[tiab] OR intraassay[tiab] OR intra-assay[tiab] OR interindividual[tiab] OR inter-individual[tiab] OR intraindividual[tiab] OR intra-individual[tiab] OR interparticipant[tiab] OR inter-participant[tiab] OR intraparticipant[tiab] OR intra-participant[tiab] OR kappa[tiab] OR kappa's[tiab] OR kappas[tiab] OR repeatab*[tw] OR ((replicab*[tw] OR repeated[tw]) AND (measure[tw] OR measures[tw] OR findings[tw] OR result[tw] OR results[tw] OR test[tw] OR tests[tw])) OR generaliza*[tiab] OR generalisa*[tiab] OR concordance[tiab] OR (intraclass[tiab] AND correlation*[tiab]) OR discriminative[tiab] OR "known group"[tiab] OR "factor analysis"[tiab] OR "factor analyses"[tiab] OR "factor structure"[tiab] OR "factor structures"[tiab] OR dimension*[tiab] OR subscale*[tiab] OR (multitrait[tiab] AND scaling[tiab] AND (analysis[tiab] OR analyses[tiab])) OR "item discriminant"[tiab] OR "interscale correlation*"[tiab] OR error[tiab] OR errors[tiab] OR "individual variability"[tiab] OR "interval variability"[tiab] OR "rate variability"[tiab] OR (variability[tiab] AND (analysis[tiab] OR values[tiab])) OR (uncertainty[tiab] AND (measurement[tiab] OR measuring[tiab])) OR "standard error of measurement"[tiab] OR sensitiv*[tiab] OR responsive*[tiab] OR (limit[tiab] AND detection[tiab]) OR "minimal detectable concentration"[tiab] OR interpretab*[tiab] OR ((minimal[tiab] OR minimally[tiab] OR clinical[tiab] OR clinically[tiab]) AND (important[tiab] OR significant[tiab] OR detectable[tiab]) AND (change[tiab] OR difference[tiab])) OR

(small*[tiab] AND (real[tiab] OR detectable[tiab]) AND (change[tiab] OR difference[tiab])) OR "meaningful change"[tiab] OR "ceiling effect"[tiab] OR "floor effect"[tiab] OR "Item response model"[tiab] OR IRT[tiab] OR Rasch[tiab] OR "Differential item functioning"[tiab] OR DIF[tiab] OR "computer adaptive testing"[tiab] OR "item bank"[tiab] OR "cross-cultural equivalence"[tiab])

NOT

("addresses"[Publication Type] OR "biography"[Publication Type] OR "case reports"[Publication Type] OR "comment"[Publication Type] OR "directory"[Publication Type] OR "editorial"[Publication Type] OR "festschrift"[Publication Type] OR "interview"[Publication Type] OR "lectures"[Publication Type] OR "legal cases"[Publication Type] OR "legislation"[Publication Type] OR "letter"[Publication Type] OR "news"[Publication Type] OR "newspaper article"[Publication Type] OR "patient education handout"[Publication Type] OR "popular works"[Publication Type] OR "congresses"[Publication Type] OR "consensus development conference"[Publication Type] OR "consensus development conference, nih"[Publication Type] OR "practice guideline"[Publication Type]) NOT ("animals"[MeSH Terms] NOT "humans"[MeSH Terms])

Search strategy for EMBASE

1. Construct

(somatization:de,ab,ti OR somatisation:de,ab,ti OR somatoform:de,ab,ti OR hysteri*:de,ab,ti OR briquet:de,ab,ti OR polysymptom*:de,ab,ti OR multisomatoform:de,ab,ti OR somatizer*:de,ab,ti OR (somatic NEXT/3 symptom*):de,ab,ti OR mups:de,ab,ti OR 'medical unexplained':de,ab,ti OR 'medically unexplained':de,ab,ti OR (unexplained NEXT/1 medical*):de,ab,ti OR (unexplained NEXT/3 symptom*):de,ab,ti OR (unexplained NEXT/3 syndrom*):de,ab,ti OR (frequent NEXT/1 attend*):de,ab,ti OR ('multiple physical' NEXT/3 symptom*):de,ab,ti OR ('multiple symptom' NEXT/3 diagnos*):de,ab,ti OR neurastheni*:de,ab,ti)

AND

2. Population

(family OR physician* OR practice*:de,it,lnk,ab,ti OR 'primary care' OR 'Primary Health Care'/exp OR primary:de,it,lnk,ab,ti OR (general NEXT/1 pract*) OR gp:ab,ti OR gps:ab,ti)

AND

3. Measurement properties

'intermethod comparison'/exp OR 'data collection method'/exp OR 'validation study'/exp OR 'feasibility study'/exp OR 'pilot study'/exp OR 'psychometry'/exp OR 'reproducibility'/exp OR reproducib*:ab,ti OR 'audit':ab,ti OR psychometr*:ab,ti OR clinimetr*:ab,ti OR clinometr*:ab,ti OR 'observer variation'/exp OR 'observer variation':ab,ti OR 'discriminant analysis'/exp OR 'validity'/exp OR reliab*:ab,ti OR valid*:ab,ti OR 'coefficient':ab,ti OR 'internal consistency':ab,ti OR (cronbach*:ab,ti AND ('alpha':ab,ti OR 'alphas':ab,ti)) OR 'item correlation':ab,ti OR 'item correlations':ab,ti OR 'item selection':ab,ti OR 'item selections':ab,ti OR 'item reduction':ab,ti OR 'item reductions':ab,ti OR 'agreement':ab,ti OR 'precision':ab,ti OR 'imprecision':ab,ti OR 'precise values':ab,ti OR 'test-retest':ab,ti OR ('test':ab,ti AND 'retest':ab,ti) OR (reliab*:ab,ti AND ('test':ab,ti OR 'retest':ab,ti)) OR 'stability':ab,ti OR 'interrater':ab,ti OR 'inter-rater':ab,ti OR 'intrarater':ab,ti OR 'intra-rater':ab,ti OR 'intertester':ab,ti OR 'inter-tester':ab,ti OR 'intratester':ab,ti OR 'intratester':ab,ti OR 'interobeserver':ab,ti OR 'inter-observer':ab,ti OR 'intraobserver':ab,ti OR 'intraobserver':ab,ti OR 'intertechnician':ab,ti OR 'inter-technician':ab,ti OR 'intratechnician':ab,ti OR 'intratechnician':ab,ti OR 'interexaminer':ab,ti OR 'inter-examiner':ab,ti OR 'intraexaminer':ab,ti OR 'intraexaminer':ab,ti OR 'interassay':ab,ti OR 'inter-assay':ab,ti OR 'intraassay':ab,ti OR 'intra-assay':ab,ti OR 'interindividual':ab,ti OR 'inter-individual':ab,ti OR 'intraindividual':ab,ti OR 'intra-individual':ab,ti OR 'interparticipant':ab,ti OR 'inter-participant':ab,ti OR 'intraparticipant':ab,ti OR 'intraparticipant':ab,ti OR 'kappa':ab,ti OR 'kappas':ab,ti OR 'coefficient of variation':ab,ti OR repeatab*:ab,ti OR (replicab*:ab,ti OR 'repeated':ab,ti AND ('measure':ab,ti OR 'measures':ab,ti OR 'findings':ab,ti OR 'result':ab,ti OR 'results':ab,ti OR 'test':ab,ti OR 'tests':ab,ti)) OR generaliza*:ab,ti OR generalisa*:ab,ti OR 'concordance':ab,ti OR ('intraclass':ab,ti AND correlation*:ab,ti) OR 'discriminative':ab,ti OR 'known group':ab,ti OR 'factor analysis':ab,ti OR 'factor analyses':ab,ti OR 'factor structure':ab,ti OR 'factor structures':ab,ti OR 'dimensionality':ab,ti OR subscale*:ab,ti OR 'multitrait scaling analysis':ab,ti OR 'multitrait scaling analyses':ab,ti OR 'item discriminant':ab,ti OR 'interscale correlation':ab,ti OR 'interscale correlations':ab,ti OR ('error':ab,ti OR 'errors':ab,ti AND (measure*:ab,ti OR correlat*:ab,ti OR evaluat*:ab,ti OR 'accuracy':ab,ti OR 'accurate':ab,ti OR 'precision':ab,ti OR 'mean':ab,ti)) OR 'individual variability':ab,ti OR 'interval variability':ab,ti OR 'rate variability':ab,ti OR 'variability analysis':ab,ti OR ('uncertainty':ab,ti AND ('measurement':ab,ti OR 'measuring':ab,ti)) OR 'standard error of measurement':ab,ti OR sensitiv*:ab,ti OR responsive*:ab,ti OR ('limit':ab,ti AND 'detection':ab,ti) OR 'minimal detectable concentration':ab,ti OR interpretab*:ab,ti OR (small*:ab,ti AND ('real':ab,ti OR 'detectable':ab,ti) AND ('change':ab,ti OR

'difference':ab,ti)) OR 'meaningful change':ab,ti OR 'minimal important change':ab,ti OR 'minimal important difference':ab,ti OR 'minimally important change':ab,ti OR 'minimally important difference':ab,ti OR 'minimal detectable change':ab,ti OR 'minimal detectable difference':ab,ti OR 'minimally detectable change':ab,ti OR 'minimally detectable difference':ab,ti OR 'minimal real change':ab,ti OR 'minimal real difference':ab,ti OR 'minimally real change':ab,ti OR 'minimally real difference':ab,ti OR 'ceiling effect':ab,ti OR 'floor effect':ab,ti OR 'item response model':ab,ti OR 'irt':ab,ti OR 'rasch':ab,ti OR 'differential item functioning':ab,ti OR 'dif':ab,ti OR 'computer adaptive testing':ab,ti OR 'item bank':ab,ti OR 'cross-cultural equivalence':ab,ti

Search strategy for PSYCINFO

1. Construct

TI ((somatization OR somatisation OR somatoform OR hysteri* OR briquet OR polysymptom* OR multisomatoform OR somatizer* OR (somatic W2 symptom*) OR MUPS OR "medical unexplained" OR "medically unexplained" OR "unexplained medical*" OR (unexplained W2 symptom*) OR (unexplained W2 syndrom*) OR ("frequent attend*") OR (multiple W2 "physical symptom*") OR (multiple W2 "symptom diagnos*") OR neurastheni*)) OR AB ((somatization OR somatisation OR somatoform OR hysteri* OR briquet OR polysymptom* OR multisomatoform OR somatizer* OR (somatic W2 symptom*) OR MUPS OR "medical unexplained" OR "medically unexplained" OR "unexplained medical*" OR (unexplained W2 symptom*) OR (unexplained W2 syndrom*) OR ("frequent attend*") OR (multiple W2 "physical symptom*") OR (multiple W2 "symptom diagnos*") OR neurastheni*)) OR SU ((somatization OR somatisation OR somatoform OR hysteri* OR briquet OR polysymptom* OR multisomatoform OR somatizer* OR (somatic W2 symptom*) OR MUPS OR "medical unexplained" OR "medically unexplained" OR "unexplained medical*" OR (unexplained W2 symptom*) OR (unexplained W2 syndrom*) OR ("frequent attend*") OR (multiple W2 "physical symptom*") OR (multiple W2 "symptom diagnos*") OR neurastheni*))

AND

2. Population

(DE "Primary Health Care") OR TI ((family OR physician* OR primary care OR practice* OR primary OR general pract* OR gp OR gps)) OR AB ((family OR physician* OR primary care OR practice* OR primary OR general pract* OR gp OR gps)) OR SU ((family OR physician* OR primary care))

AND

3. Measurement properties

(DE “psychometrics”) or (TI psychometr* or AB psychometr*) or (TI clinimetr* or AB clinimetr*) or (TI clinometr* OR AB clinometr*) or “outcome assessment” or (TI outcome assessment or AB outcome assessment) or (TI outcome measure* or AB outcome measure*) or “health status indicators” or “reproducibility of results” or “discriminant analysis” or (DE “test validity”) or ((TI reproducib* or AB reproducib*) or (DE “test reliability”) or (TI reliab* or AB reliab*) or (TI unreliab* or AB unreliab*)) or ((TI valid* or AB valid*) or (TI coefficient or AB coefficient) or (TI homogeneity or AB homogeneity)) or (TI homogeneous or AB homogeneous) or (TI “coefficient of variation” or AB “coefficient of variation”) or (TI “internal consistency” or AB “internal consistency”) or “internal consistency” or “reliability” or “measurement error” or (DE “consistency (measurement)”) or (DE “error of measurement”) or “content validity” or “hypothesis testing” or “structural validity” or “cross-cultural validity” or “criterion-related validity” or “responsiveness” or “interpretability” or (TI reliab* or AB reliab*) and ((TI test or AB test) OR (TI retest or AB retest)) or (TI stability or AB stability) or (TI interrater or AB interrater) or (TI inter-rater or AB inter-rater) or (TI intrarater or AB intrarater) or (TI intra-rater or AB intrarater) or (TI intertester or AB intertester) or (TI inter-tester or AB inter-tester) or (TI intratester or AB intratester) or (TI intra-tester or AB intra-tester) or (TI interobserver or AB interobserver) or (TI inter-observer or AB inter-observer) or (TI intraobserver or AB intraobserver) or (TI intra-observer or AB intra-observer) or (TI intertechnician or AB intertechnician) or (TI inter-technician or AB inter-technician) or (TI intratechnician or AB intratechnician) or (TI intra-technician or AB intra-technician) or (TI interexaminer or AB interexaminer) or (TI inter-examiner or AB inter-examiner) or (TI intraexaminer or AB intraexaminer) OR (TI intra-examiner or AB intra-examiner) or (TI intra-examiner or AB intraexaminer) or (TI interassay or AB interassay) or (TI inter-assay or AB inter-assay) or (TI intraassay or AB intraassay) or (TI intra-assay or AB intra-assay) or (TI interindividual or AB interindividual) or (TI inter-individual or AB inter-individual) OR (TI intraindividual or AB intraindividual) or (TI intra-individual or AB intra-individual) or (TI interparticipant or AB interparticipant) or (TI inter-participant or AB inter-participant) or (TI intraparticipant or AB intraparticipant) or (TI intra-participant or AB intra-participant) or (TI kappa or AB kappa) or (TI kappa’s or AB kappa’s) or (TI kappas or AB kappas) or (TI repeatab* or AB repeatab*) or (TI responsive* or AB responsive*) or (TI interpretab* or AB interpretab*)

Search strategy for CINAHL

1. Construct

TI ((somatization OR somatisation OR somatoform OR hysteri* OR briquet OR polysymptom* OR multisomatoform OR somatizer* OR (somatic W2 symptom*) OR MUPS OR "medical unexplained" OR "medically unexplained" OR "unexplained medical*" OR (unexplained W2 symptom*) OR (unexplained W2 syndrom*) OR ("frequent attend*") OR (multiple W2 "physical symptom*") OR (multiple W2 "symptom diagnos*") OR neurastheni*)) OR AB ((somatization OR somatisation OR somatoform OR hysteri* OR briquet OR polysymptom* OR multisomatoform OR somatizer* OR (somatic W2 symptom*) OR MUPS OR "medical unexplained" OR "medically unexplained" OR "unexplained medical*" OR (unexplained W2 symptom*) OR (unexplained W2 syndrom*) OR ("frequent attend*") OR (multiple W2 "physical symptom*") OR (multiple W2 "symptom diagnos*") OR neurastheni*)) OR SU ((somatization OR somatisation OR somatoform OR hysteri* OR briquet OR polysymptom* OR multisomatoform OR somatizer* OR (somatic W2 symptom*) OR MUPS OR "medical unexplained" OR "medically unexplained" OR "unexplained medical*" OR (unexplained W2 symptom*) OR (unexplained W2 syndrom*) OR ("frequent attend*") OR (multiple W2 "physical symptom*") OR (multiple W2 "symptom diagnos*") OR neurastheni*))

AND

2. Population

(MH "Primary Health Care") OR TI ((family OR physician* OR primary care OR practice* OR primary OR general pract* OR gp OR gps)) OR AB ((family OR physician* OR primary care OR practice* OR primary OR general pract* OR gp OR gps)) OR SU ((family OR physician* OR primary care))

AND

3. Measurement properties

(MH "Psychometrics") or (TI psychometr* or AB psychometr*) or (TI clinimetr* or AB clinimetr*) or (TI clinometr* OR AB clinometr*) or (MH "Outcome Assessment") or (TI outcome assessment or AB outcome assessment) or (TI outcome measure* or AB outcome measure*) or (MH "Health Status Indicators") or (MH "Reproducibility of Results") or (MH "Discriminant Analysis") or ((TI reproducib* or AB reproducib*) or (TI reliab* or AB reliab*) or (TI unreliab* or AB unreliab*)) or ((TI valid* or AB valid*) or (TI coefficient or AB coefficient) or (TI homogeneity or AB homogeneity)) or (TI

homogeneous or AB homogeneous) or (TI “coefficient of variation” or AB “coefficient of variation”) or (TI “internal consistency” or AB “internal consistency”) or (MH “Internal Consistency+”) or (MH “Reliability+”) or (MH “Measurement Error+”) or (MH “Content Validity+”) or “hypothesis testing” or “structural validity” or “cross-cultural validity” or (MH “Criterion-Related Validity+”) or “responsiveness” or “interpretability” or (TI reliab* or AB reliab*) and ((TI test or AB test) OR (TI retest or AB retest)) or (TI stability or AB stability) or (TI interrater or AB interrater) or (TI inter-rater or AB inter-rater) or (TI intrarater or AB intrarater) or (TI intra-rater or AB intrarater) or (TI intertester or AB intertester) or (TI inter-tester or AB inter-tester) or (TI intratester or AB intratester) or (TI intra-tester or AB intra-tester) or (TI interobserver or AB interobserver) or (TI inter-observer or AB inter-observer) or (TI intraobserver or AB intraobserver) or (TI intra-observer or AB intra-observer) or (TI intertechnician or AB intertechnician) or (TI inter-technician or AB inter-technician) or (TI intratechnician or AB intratechnician) or (TI intra-technician or AB intra-technician) or (TI interexaminer or AB interexaminer) or (TI inter-examiner or AB inter-examiner) or (TI intraexaminer or AB intraexaminer) OR (TI intra-examiner or AB intra-examiner) or (TI intra-examiner or AB intraexaminer) or (TI interassay or AB interassay) or (TI inter-assay or AB inter-assay) or (TI intraassay or AB intraassay) or (TI intra-assay or AB intra-assay) or (TI interindividual or AB interindividual) or (TI inter-individual or AB inter-individual) OR (TI intraindividual or AB intraindividual) or (TI intra-individual or AB intra-individual) or (TI interparticipant or AB interparticipant) or (TI inter-participant or AB inter-participant) or (TI intraparticipant or AB intraparticipant) or (TI intra-participant or AB intra-participant) or (TI kappa or AB kappa) or (TI kappa’s or AB kappa’s) or (TI kappas or AB kappas) or (TI repeatab* or AB repeatab*) or (TI responsive* or AB responsive*) or (TI interpretab* or AB interpretab*)

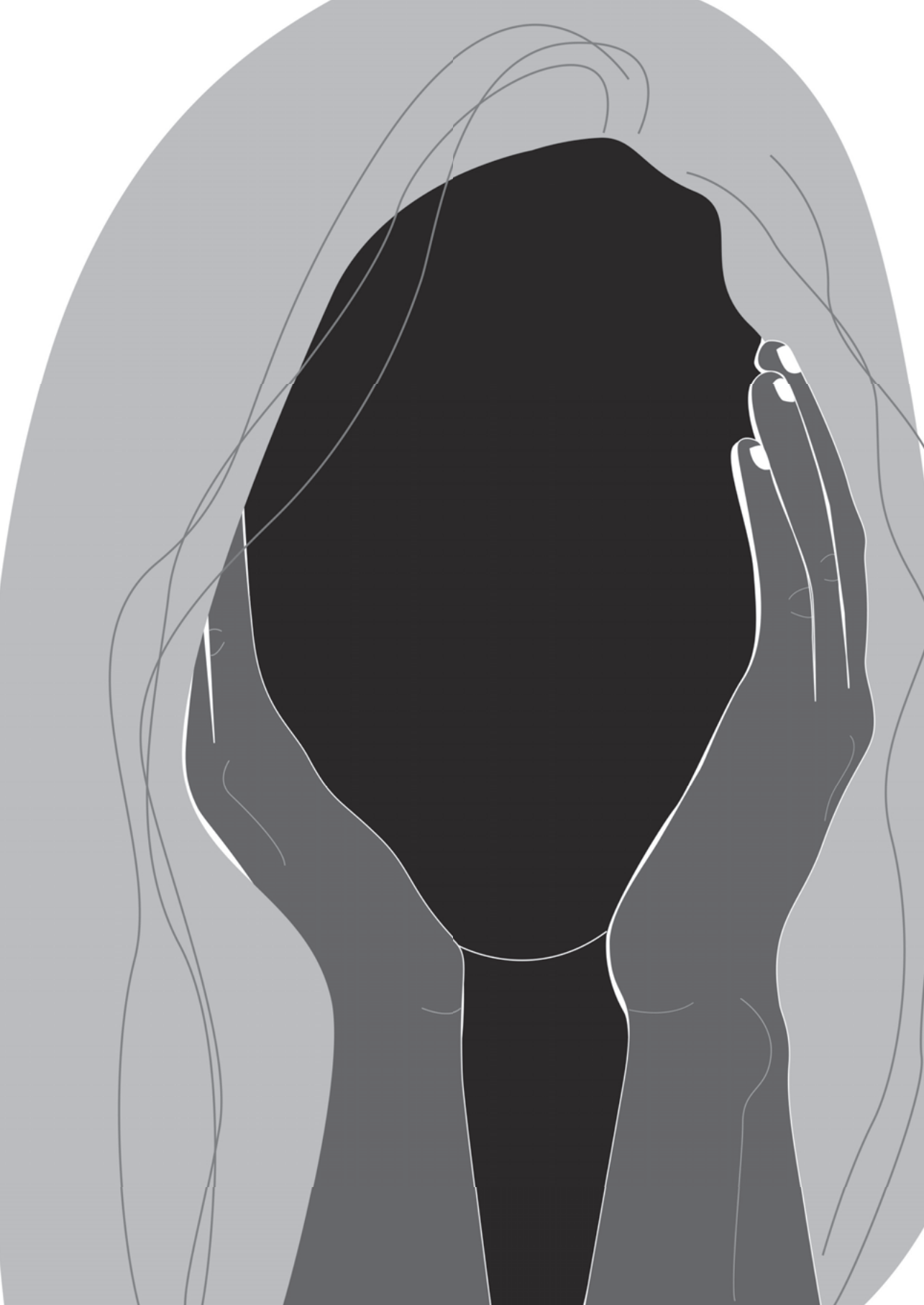
Appendix B. Criteria for good measurement properties

Measurement property	Rating*	Criteria
Internal consistency	+	At least limited evidence for unidimensionality or positive structural validity AND Cronbach's alpha(s) ≥0.70 and ≤0.95
	?	Not all information for '+' reported OR conflicting evidence for unidimensionality or negative structural validity
	-	Criteria for '+' not met
Reliability	+	ICC or weighted Kappa ≥ 0.70 OR Pearson's r ≥ 0.80
	?	ICC, weighted Kappa, or Pearson's r not reported
	-	Criteria for '+' not met
Measurement error	+	SDC or LoA < MIC
	?	MIC not defined
	-	Criteria for '+' not met
Content validity	+	All items refer to relevant aspects of the construct to be measured AND are relevant for the target population AND are relevant for the context of use AND together comprehensively reflect the construct to be measured
	?	Not all information for '+' reported
	-	Criteria for '+' not met
Structural validity	+	**CTT**: Unidimensionality: EFA: First factor accounts for at least 20% of the variability AND ratio of the variance explained by the first to the second factor > 4 OR Bi-factor model: Standardized loadings on a common factor >0.30 AND correlation between individual scores under a bi-factor and unidimensional model >0.90 Structural validity: CFI or TLI or comparable measure >0.95 AND RMSEA <0.06 OR SRMR <0.08 **Rasch/IRT**: At least limited evidence for unidimensionality or positive structural validity AND no evidence for violation of local independence: Rasch: standardized item-person fit residuals between -2.5 and 2.5; OR IRT: residual correlations among the items after controlling for the dominant factor < 0.20 OR Q3's < 0.37 AND no evidence for violation of monotonicity: adequate looking graphs OR item scalability >0.30 AND adequate model fit: Rasch: infit and outfit mean squares ≥ 0.5 and ≤ 1.5 OR Z-standardized values > -2 and <2; OR IRT: G2 >0.01 Optional additional evidence: Adequate targeting; Rasch: adequate person-item threshold distribution; IRT: adequate threshold range

Measurement property	Rating*	Criteria
		No important DIF for relevant subject characteristics (such as age, gender, education), McFadden's R2 < 0.02
	?	CTT: Not all information for '+' reported IRT: Model fit not reported
	-	Criteria for '+' not met
Hypothesis testing	+	At least 75% of the results are in accordance with the hypotheses
	?	No correlations with instrument(s) measuring related construct(s) AND no differences between relevant groups reported
	-	Criteria for '+' not met
Cross-cultural validity	+	No important differences found between language versions in multiple group factor analysis or DIF analysis
	?	Multiple group factor analysis AND DIF analysis not performed
	-	One or more criteria for '+' not met
Criterion validity	+	Convincing arguments that gold standard is "gold" AND correlation with gold standard ≥ 0.70 OR AUC ≥ 0.70 OR SE and SP ≥0.60
	?	Not all information for '+' reported
	-	Criteria for '+' not met
Responsiveness	+	At least 75% of the results are in accordance with the hypotheses
	?	No correlations with changes in instrument(s) measuring related construct(s) AND no differences between changes in relevant groups reported
	-	Criteria for '+' not met

AUC: area under the curve, CFI: comparative fit index; CTT: classical test theory, DIF: differential item functioning; EFA: exploratory factor analysis, ICC: intraclass correlation coefficient, IRT: item response theory, LoA: limits of agreement, MIC: minimal important change, RMSEA: root mean square error of approximation, SE: sensitivity, SP: specificity, SEM: standard error of measurement, SDC: smallest detectable change, SRMR: standardized root mean residuals, TLI: Tucker-Lewis index

* + = positive rating; ? = indeterminate rating; - = negative rating



Chapter 3

The CIPRUS study, a nurse-led psychological treatment for patients with undifferentiated somatoform disorder in primary care: study protocol for a randomised controlled trial

Kate Sitnikova
Stephanie S Leone
Lyonne NL Zonneveld
Harm WJ van Marwijk
Judith E Bosmans
Johannes C van der Wouden
Henriëtte E van der Horst

Published in: Trials 2017;18:206.

ABSTRACT

Background: Up to a third of patients presenting medically unexplained physical symptoms in primary care may have a somatoform disorder, of which undifferentiated somatoform disorder (USD) is the most common type. Psychological interventions can reduce symptoms associated with USD and improve functioning. Previous research has either been conducted in secondary care or interventions have been provided by general practitioners (GPs) or psychologists in primary care. As efficiency and costs-effectiveness are imperative in primary care, it is important to investigate whether nurse-led interventions are effective as well. The aim of this study is to examine the effectiveness and cost-effectiveness of a short cognitive behavioural therapy (CBT)-based treatment for patients with USD provided by mental health nurse practitioners (MHNP), compared to usual care.

Methods: In a cluster randomised controlled trial, 212 adult patients with USD will be assigned to the intervention or care as usual. The intervention group will be offered a short, individual CBT-based treatment by the MHNP in addition to usual GP care. The main goal of the intervention is that patients become less impaired by their physical symptoms and cope with symptoms in a more effective way. In six sessions patients will receive problem-solving treatment. The primary outcome is improvement in physical functioning, measured by the physical component summary score of the RAND-36. Secondary outcomes include health-related quality of life measured by the separate subscales of the RAND-36, somatization (PHQ-15) and symptoms of depression and anxiety (HADS). Problem-solving skills, health anxiety, illness perceptions, coping, mastery and working alliance will be assessed as potential mediators. Assessments will be done at 0, 2, 4, 8 and 12 months. An economic evaluation will be conducted from a societal perspective with quality of life as the primary outcome measure assessed by the EQ-5D-5L. Health care, patient and lost productivity costs will be assessed with the Tic-P.

Discussion: We expect that the intervention will improve physical functioning and is cost-effective compared to usual care. If so, more patients might successfully be treated in general practice, decreasing the number of referrals to specialist care.

BACKGROUND

Physical symptoms that cannot be sufficiently explained by organic pathology are commonly presented by patients in all health care settings (1, 2). In primary care up to one third of patients present these medically unexplained physical symptoms (MUPS) to the GP (3). Most of these symptoms are self-limiting, but some persist and cluster. Up to a third of patients presenting with such symptoms in primary care can be diagnosed with a somatoform disorder (3–5). The most common type of somatoform disorder is the undifferentiated somatoform disorder (USD), including patients who suffer from at least one impairing unexplained physical symptom lasting longer than 6 months (6). USD is associated with considerable functional impairment and reduced quality of life, which in turn results in a high illness burden. Anxiety and depression are comorbid in at least 13.7% of the cases (7, 8) and may aggravate symptoms and functional limitations (5).

Additionally, USD is associated with high health care costs due to frequent and excessive health care use, repeated diagnostic procedures and high lost productivity costs (9, 10). In 2010 the total costs of all somatoform disorders in Europe amounted to €21 billion, which was considered to be a conservative estimate (11, 12).

Previous research shows that only half of patients with USD seeks help from a mental health care provider (13). However, several reviews on treatment of MUPS and somatoform disorders show that cognitive behavioural therapy (CBT) is effective, with moderate effect sizes (14, 15). Moreover, CBT interventions are effective in treating anxiety and depression (16) that often co-occur with USD.

A Cochrane review on non-pharmacological interventions for MUPS and somatoform disorders identified eight studies that recruited patients from primary care (14). However, in only two of these studies treatment was actually performed in the primary care setting (17). In one study (18), psychologists from secondary care performed a CBT-based group training within general practices. This training proved to be effective in increasing physical and emotional functioning and quality of life. Another study also offered a group intervention in the general practice, but the treatment was given by specifically trained GPs and psychosomatic specialists (17). This intervention was effective in improving mental but not physical functioning. However, the group formats may not appeal to all patients and take considerable time.

Less robust evidence shows that psychotropic medication, such as antidepressants, may also have some effect on the symptoms but these can induce dependence and may have

side-effects (19). Also, a recent pilot study on a brief, multimodal psychosomatic therapy, combining physical and psychological components and delivered by physiotherapists showed improvement in perceived symptom severity, somatization and hyperventilation, but larger trials are needed to draw further conclusions (20).

Meanwhile, patients with USD frequently visit their GP (12). However, to support and treat these patients can be a challenging task (21, 22). GPs may feel powerless because they cannot find a somatic cause for the physical symptoms with which these patients typically present, and the patients themselves may fear serious disease (23). In some cases, this may lead to unnecessary referrals to medical specialists for diagnostic and therapeutic purposes (24, 25). Although GPs recognise the need of discussing psychological issues with these patients, this is often not feasible due to time constraints or GPs may feel ill-equipped to do so themselves (21). Knowledge about treatment of USD in primary care and its cost-effectiveness in comparison with usual care is lacking.

Currently, the contribution of the mental health nurse practitioner (MHNP) within general practices in the Netherlands is increasing. A part of mental health care, in the form of short psychological treatment or coaching sessions, is taken over from the GP by the MHNP. MHNPs work within the general practice, and are trained to provide short-term psychological treatment. They are seemingly in a good position to offer psychological help to patients with USD in a more accessible way, provided such tasks are clear and there is evidence that such extra care is effective. However, they do not yet have a standard evidence-based protocol for treating these patients. We will, therefore, adapt an existing and effective secondary care protocol for primary care. To the best of our knowledge no previous research on the effectiveness of individual treatment conducted by primary care health care workers, such as MHNPs, has been executed yet.

The aim of this study is to examine the effectiveness and cost-effectiveness of a short CBT-based psychological treatment for patients with USD provided by MHNPs, in comparison with usual care.

METHODS/DESIGN

This protocol was developed in accordance with the Standard Protocol Items: Recommendations for Interventional Trials (SPIRIT) Statement. We will conduct a parallel-group, multicentre, cluster randomised controlled trial in 39 primary care centres. Cluster randomisation will take place at the MHNP and general practice level in order to avoid contamination between the intervention and control condition. Clusters will be

composed by matching MHNPs to all participating general practices where the MHNP works and to all other MHNPs who also work in these general practices. The entire cluster consisting of one or more MHNPs and general practices will then be randomly assigned to the intervention or to care as usual (CAU) group, prior to the inclusion of patients. An independent statistician, not involved in the selection of the practices and MHNPs, will carry out the randomisation. In order to balance the size of the intervention and CAU groups, randomisation will be stratified according to cluster size (small: less than 5000 patients, and large: 5000 patients or more). Assessments will take place at baseline (T0), during the intervention period at approximately 2 months (T1), directly after the intervention period at 4 months (T2), at 8 months (T3) and at 12 months (T4) after baseline.

This study is registered at the Dutch Trial Registry (NTR4686) and was approved by the VU Medical Center Ethics Committee. It will be conducted according to the principles of the Declaration of Helsinki (version 2013). Important protocol modifications will be communicated with the Dutch Trial Registry and the VU Medical Center Ethics Committee.

Participants

Participants will be recruited from various general practices and care groups situated in different geographical locations in The Netherlands. Patients will be eligible for the study if they are 18 years or older and meet the criteria for USD according to the Diagnostic and Statistical Manual of Mental disorders IV (DSM-IV) (6).

Patients will be excluded from participation in the study if they have a medical or psychological disorder explaining their symptoms, a severe psychiatric disorder (e.g. psychotic disorders), are currently receiving psychological treatment for USD or have poor language skills or handicaps that prevent them from understanding the intervention. Patients can withdraw from the study at any time for any reason without any consequences.

Inclusion procedure

The researchers and GPs will select adult patients (aged 18 years or older) from the GP's electronic database, who consulted the GP with one or more symptoms from the Robbins list (26) (Table 1) at least twice in the previous 3 months. The presented symptom does not necessarily have to be the same for each visit. The Robbins list consists of 23 physical symptoms that are associated with functional somatic syndromes. The symptoms on this list are likely to be medically unexplained if they lack an accompanying

'diagnostic' or 'disease' International Classification of Primary Care (ICPC) code (i.e. ICPC code >70). Following this step, the participating GPs will check the selected patients for exclusion criteria in order to, for example, avoid inclusion of patients with actual somatic pathology.

Table 1. Symptoms from the Robbins list (26)

1. Back pain
2. Joint pain
3. Extremity pain
4. Headaches
5. Weakness
6. Fatigue
7. Sleep disturbance
8. Difficulty concentrating
9. Loss of appetite
10. Weight change
11. Restlessness
12. Thoughts slower
13. Chest pain
14. Shortness of breath
15. Palpitations
16. Dizziness
17. Lump in throat
18. Numbness
19. Nausea
20. Loose bowels
21. Gas or bloating
22. Constipation
23. Abdominal pain

Patients identified as potentially eligible will then receive brief information about the study and the Patient Health Questionnaire somatization scale (PHQ-15) (27) from their GP. Patients who are interested in participation in the study and who have a PHQ-15 score of 5 (low symptom severity) or higher will receive extensive study information. We chose the cut-off point for low symptom severity in order to make sure that patients with disabling somatic complaints are not wrongly excluded at this point. Patients will then be invited to participate in a clinical interview (Structured Clinical Interview for DSM-IV Axis I Disorders (SCID-I)) to assess DSM-IV criteria for USD. Trained members of the

research team will administer the interview by telephone. Patients meeting the DSM-IV criteria for USD (the presence of one or more medically unexplained physical symptoms that last at least 6 months and significantly impair functioning/quality of life) will receive an Informed Consent Form. After completing the Informed Consent Form and sending it back to the researcher they will be included in the intervention or CAU group based on the allocation of the MHNP and general practice to which they belong. An overview of the inclusion procedure is provided in Figure 1.

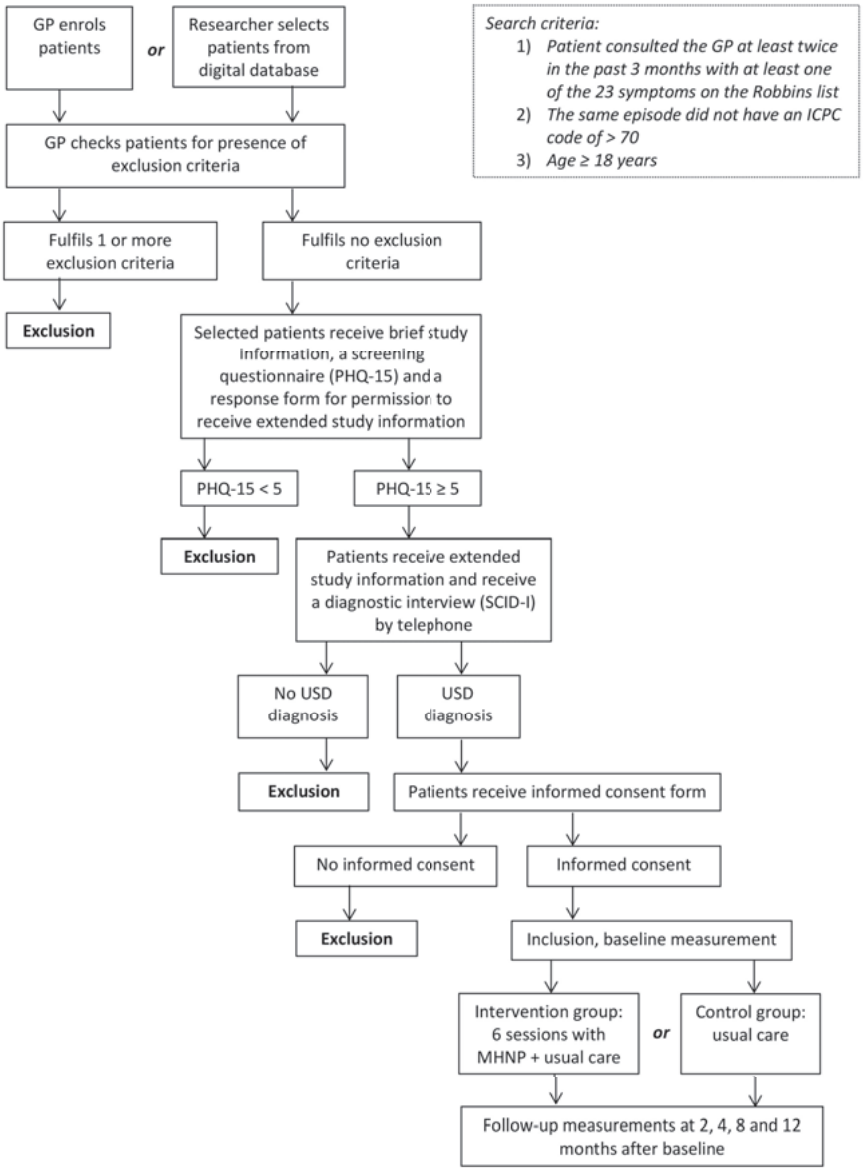


Figure 1. Flow chart of the study.
GP general practitioner, *ICPC* International Classification of Primary Care, *PHQ-15* the Patient Health Questionnaire somatization scale, *SCID-I* Structured Clinical Interview for DSM-IV for Axis I disorders, *USD* undifferentiated somatoform disorder, *MHNP* mental health nurse practitioner.

The intervention

The intervention consists of extra care in addition to usual care. We realise that, specifically for this group, acknowledging and receiving help partly depends on the willingness to accept psychological assistance. Therefore, we hope that providing the intervention in their own general practice by their own MHNP will lower possible barriers to receiving psychological help.

The intervention consists of six individual sessions of 30 min each. The aim of the treatment is to improve physical functioning by helping patients cope with the consequences of their physical symptoms and with everyday problems in general. The rationale for the intervention is based on the consequences model for somatoform complaints (28, 29).

The consequences model for somatoform complaints has been found to be effective in previous Dutch randomised intervention studies (18, 30) and it is used in secondary mental health care. The model focuses on the consequences or problems that arise due to somatoform complaints rather than on the causes of somatoform complaints which are by definition unknown. In our study, we use the consequences model modified by Zonneveld et al. (31). The modified consequences model assumes that physical symptoms lead to various consequences in the daily life of patients which are in fact survival strategies in reaction to the physical symptoms. Although these survival strategies must have been beneficial at the beginning (otherwise they would not have been developed) they can aggravate symptoms and become harmful or devastating in the long run. In the model, patients can improve (physical) functional and quality of life by enhancing the more beneficial survival strategies in the long run instead of the harmful ones. Identified consequences or problems are then tackled using cognitive behavioural problem-solving techniques which will be learned and applied by patients in a stepwise manner according to the steps outlined in problem-solving therapy (PST). The goal is to help patients develop more helpful survival strategies in the long run. PST has proven to be effective for depression in primary care and is suitable for being carried out by trained GPs or nurses (32, 33). Although PST has been investigated less thoroughly in patients with USD, several studies show promising results (18, 34, 35). Patients become more skilled in developing (helpful) survival strategies or solutions by working through the following seven steps: 1) identifying and specifying a problematic situation, 2) setting a clear goal for resolving this problematic situation, 3) formulating as many survival strategies or solutions as possible to reach this goal, 4) making a cost-benefit analysis for each strategy or solution, 5) selecting the most favourable strategy or solution, 6) specifying the necessary steps to implement the strategy or solution and 7)

implementing the strategy or solution and evaluating its results. An intervention protocol was developed in which the intervention is described in detail. More information about the intervention can be requested from the first author.

Training of mental health nurse practitioners

Before administering the intervention, all MHNPs in the intervention group will receive two group-training sessions of 3 to 3.5 h each, depending on the size of the group, during a 2-week period (one session a week). A registered clinical psychologist with broad clinical expertise in treating patients with somatoform complaints will lead the training sessions. The aim is to train the MHNPs in understanding and applying the intervention according to the intervention protocol. In the first training session the theoretical background of MUPS, USD and the modified consequences model will be given. In the second training session, the steps of PST will be introduced, explained, modelled by the clinical psychologist and finally practised by the MHNPs by applying the treatment to a co-trainee. In between the two training sessions, MHNPs will be asked to go through the PST steps to solve a minor problem of their own as homework. MHNPs will also receive a copy of the treatment protocol that will serve as a guideline for the intervention.

Feasibility testing and ongoing supervision

In order to determine whether the intervention is feasible in practice, two MHNPs will be asked to pilot test the protocol with a patient. Also, two patients will be asked about their opinion on the treatment protocol. Their feedback will be collected and the protocol will be adjusted if needed. During the intervention period a supervision session with the clinical psychologist (by telephone or face-to-face) will be offered. MHNPs can also contact the researcher during the entire intervention period for any questions regarding the treatment and the study.

Furthermore, five randomly selected MHNPs will be asked to record all sessions with all of their patients. The researcher (KS) will listen to the recordings and identify topics in the obstacles encountered by the MHNPs. The researcher will communicate this with the clinical psychologist, so that these topics can be addressed during supervision.

Care as usual (CAU) group

Patients in the CAU group will not be offered any additional specific intervention other than the care they would usually receive from the GP and/or MHNP, carried out according to the guideline for medically unexplained symptoms by the Dutch College of General Practitioners (NHG) (36).

Outcome measures

An overview of all outcome measures and the time points of assessments can be found in Table 2.

Table 2. Overview of assessment moments and outcome measurements

	Instrument	Baseline (T0)	2 months (T1)	4 months (T2)	8 months (T3)	12 months (T4)
Primary outcomes						
Physical functioning	PCS of RAND-36	x	x	x		x
Direct and indirect costs	Tic-P	x		x	x	x
Quality of life	EQ-5D-5L	x	x	x	x	x
Secondary outcomes						
HRQoL	RAND-36 subscales	x	x	x		x
Anxiety and depression	HADS	x	x	x		x
Number/ severity of symptoms	PHQ-15	x	x	x		x
Potential mediators						
Health anxiety	Whitely Index	x	x	x		
Illness perceptions	Brief IPQ	x	x	x		
Cognitions and coping	CBRQ	x	x	x		
Social problem-solving skills	SPSI-R:S	x	x	x		
Mastery	Pearlin Mastery Scale	x	x	x		
Working alliance	WAI-SF*		x	x		

PCS physical component summary score, *RAND-36* RAND-36-item Health Survey, *Tic-P* Trimbos and iMTA questionnaire on Costs associated with Psychiatric illness, *EQ-5D-5L* EuroQol 5D – 5 Level version, *HRQoL* health-related quality of life, *HADS* Hospital Anxiety and Depression Scale, *PHQ-15* Patient Health Questionnaire somatization scale, *IPQ* Illness Perception Questionnaire, *CBRQ* Cognitive and Behavioural Responses Questionnaire, *SPSI-R:S* Social Problem-Solving Inventory Revised: Short form, *WAI-SF* Working Alliance Inventory-Short form.

* only administered in the intervention group

Primary outcomes

The primary clinical outcome of this study is the improvement in physical functioning during the total follow-up period measured by the physical component summary score (PCS) of the RAND-36-item Health Survey (RAND-36). The RAND-36 is a widely used, valid and reliable health-related quality of life (HRQoL) instrument comprising 36 items that assess eight health domains: physical functioning, role limitations caused by physical health problems, role limitations caused by emotional problems, social functioning, emotional well-being, energy/fatigue, pain and general health perceptions (37–39). Physical and mental health summary scores (PCS and MCS, respectively) can be derived from the eight scales. The raw scores are transformed into scores ranging from 0-100 with a higher score indicating better functioning.

The primary outcome for the economic evaluation is quality of life as measured by the EuroQol 5D – 5 level version (EQ-5D-5L) (40). The EQ-5D-5L is a standardised instrument including five health domains (mobility, self-care, usual activities, pain/discomfort and anxiety/depression). It comprises five items, one single item referring to each domain, that are assessed on a 5-point scale: 0 = 'no problems with', 1 = 'slight problems with', 2 = 'moderate problems with', 3 = 'severe problems with' and 4 = 'unable to'. The health state indicated by patients on the EQ-5D-5L will be converted to a utility score using the Dutch EQ-5D-5L tariff (41). The EQ-5D-5L utility scores at different time points will be used to calculate quality-adjusted life years (QALYs) using the area under the curve method. Changes between health states at different time points are considered to be linear. Also the patient is asked to rate their general health on a 0-100 scale (40).

Societal costs will be assessed with the Trimbos and iMTA questionnaire on Costs associated with Psychiatric illness (Tic-P) (42). The Tic-P is an instrument that assesses self-reported health care utilisation, medication use, informal care, absenteeism from paid and unpaid work, and presenteeism. The costs of the intervention will be estimated using a bottom-up approach. For the valuation of health care utilisation and informal care, standard prices published in the Dutch costing guidelines will be used (42). Medication use will be valued using prices of the Royal Dutch Society for Pharmacy (Z-index). The friction cost approach will be used to estimate absenteeism from paid work based on sex specific mean wages in the Dutch population.

Secondary outcomes

Secondary outcome measures are the eight separate subscales scales of the RAND-36 (physical functioning, role limitations caused by physical health problems, role limitations caused by emotional problems, social functioning, emotional well-being,

energy/fatigue, pain and general health perceptions) (39), severity of somatization (Patient Health Questionnaire 15-item somatic symptom scale (PHQ-15) (27)) and depressive and anxiety symptoms (Hospital Anxiety and Depression Scale (HADS) (43)). PHQ-15 is a somatic symptom scale derived from the Patient Health Questionnaire (PHQ). It comprises 15 items, each of which can be scored from 0 = 'not bothered at all' to 2 = 'bothered a lot', which results in a total score ranging from 0 to 30. Higher scores indicate higher somatic symptom severity. Scores of 5, 10 and 15 represent cut-off points for low, medium, and high somatic symptom severity, respectively (27).

The HADS is a 14-item instrument assessing symptoms of anxiety (seven items) and depression (seven items). There are four answer categories that are given scores of 0-3, resulting in a total score ranging from 0 to 21 for each scale. Higher scores on the HADS indicate more severe symptoms of anxiety and depression. A score of 0-7 indicates no depressive or anxiety disorder. A score of 8-10 indicates a possible depressive or anxiety disorder, whereas a score of 11-21 indicates a probable depressive or anxiety disorder (43).

Patient and illness characteristics

In order to examine which patients are most likely to benefit from the intervention, we will assess various potential effect modifiers. The choice of these factors was based in part on the Dutch multidisciplinary guideline for MUPS and somatoform disorders which advocates the use of three patient profiles: mild, moderate and severe symptom profile (36). Severity depends on factors such as duration, severity and number of symptoms and comorbidity. The potential effect modifiers selected in this study are:

- Demographic factors: age, gender and education (self-report)
- Illness duration (self-report)
- Severity of somatization (somatization scale of the PHQ-15 (27))
- Physical comorbidity (self-report)
- Psychiatric comorbidity: anxiety and depression (HADS (43))

Potential mediators

The intervention is expected to have positive effects on physical functioning and quality of life through developing problem-solving skills and increasing adequate coping. Moreover, health anxiety and dysfunctional somatic causal attributions are thought to be important aggravating factors of somatoform disorders which are reflected in the new DSM-5 criteria for somatic symptom disorder (formerly somatoform disorders) (44). Therefore, the following factors will be assessed as potential mediators:

- Problem-solving skills (Social Problem-Solving Inventory (45))
- Health anxiety (Whiteley Index (46))
- Illness perceptions (brief IPQ (47))
- Coping and beliefs about symptoms (CBRQ (48))
- Mastery (Pearlin Mastery Scale (49))
- Working alliance (WAI-SF (50))

The Social Problem-solving Inventory Revised: Short Form (SPSI-R:S) measures an individual's problem-solving strengths and weaknesses. It assesses five dimensions: positive problem orientation, rational problem solving, negative problem orientation, impulsive/careless style and avoidance style. It consists of 25 items that can each be scored on a 5-point Likert scale (0 = 'not at all true for me' to 4 = 'extremely true for me'). Higher scores indicate greater effective social problem-solving skills (45).

The Whiteley Index is a short instrument measuring health anxiety and is often used in investigating symptoms of hypochondria. It consists of 14 items with two answer categories 'yes' and 'no'. A higher score indicates greater health anxiety (46).

The brief Illness Perception Questionnaire (IPQ) is based on the widely used IPQ-R. It is designed in order to rapidly assess the cognitive and emotional representations of illness. It comprises nine items in total. Five items assess cognitive illness representations: consequences, timeline, personal control, treatment control and identity. Two items assess emotional representations: concern and emotions. One item assesses illness comprehensibility. The last item is an open-ended question assessing causal representation (47).

The Cognitive and Behavioural Responses Questionnaire (CBRQ) measures patients' cognitive and behavioural responses to their illness. This tool has been developed to measure specific cognitions and coping styles in patients with physical symptoms. It consists of 41 items that can be rated on a 5-point Likert scale ranging from 'strongly disagree' to 'strongly agree'. These items add up to four cognitive subscales: catastrophising, damaging beliefs, embarrassment avoidance and symptom focusing; and two behavioural subscales: all-or-nothing behaviour and avoidance/resting behaviour (48).

The Pearlin Mastery Scale measures the level of perceived control, or mastery, over situations in one's life. It comprises seven items that are scored on a 5-point Likert scale, ranging from 'completely disagree' to 'completely agree' (49).

The Working Alliance Inventory-Short Form (WAI-SF) is a shortened version of the Working Alliance Inventory (WAI). It is used to measure the therapeutic alliance in an ongoing client-therapist interaction. It comprises 12 items that are scored on a 5-point Likert scale, ranging from 'never or rarely' to 'very often' (50).

Factors influencing implementation

Non-response and patient satisfaction

Patients who do not want to participate in the study will be asked for reasons for non-participation. Participating patients will receive questions about satisfaction with the content and relevance of the treatment at the 4 months follow-up. The questions will cover topics such as: patients' expectations and needs prior to treatment; whether the treatment lived up to their expectations; whether patients considered themselves to be the target audience (suffering from USD); patients' reasons for participating; whether the treatment helped in the short and the long term; satisfaction with the duration of the treatment; and whether they would recommend the treatment to someone else in the future. The above questions will be assessed using a 5-point Likert-scale ranging from 1 = 'very satisfied' to 5 = 'very unsatisfied'. Also, reasons for drop-out will be assessed.

Evaluation MHNPs

The MHNPs in the intervention group will receive a questionnaire to evaluate their opinion about the content and relevance of the treatment. The questions will cover topics such as whether MHNPs felt the treatment was effective for the patients; whether the intervention protocol was useful to the MHNPs and whether they followed the protocol as they were asked to do. Potential reasons for non-compliance will be explored. The questions on the opinions of the MHNPs will also be assessed on a 5-point Likert-scale. Additionally, MHNPs will be interviewed by telephone or in person to gain insight into factors that they deem relevant for successful implementation of the intervention in the future and what possible barriers they identify. The interviews will be recorded and transcribed for analysis.

Handling and storage of data and documents

Data will be collected and stored digitally using the web database NetQuestionnaires. This database treats data with strict confidence and does not disclose data to any third parties without permission of the user. NetQuestionnaires also takes safety measures for collecting, storing and processing data to prevent unauthorized access. In case people prefer a paper version of the questionnaires, this will be sent by regular mail. The completed paper questionnaires will be stored in a locked closet at the department of General Practice and Elderly Care Medicine. When working with data, subjects will be

assigned with a code. The code list will be safeguarded by the principal investigator. Only the principal investigator and research assistants will be able to access the source data. Data will be kept for 15 years.

Power calculation/sample size calculation

We based the sample size calculation on an expected increase in the primary clinical outcome PCS of the RAND-36 during the total follow-up period (4 and 12 months after baseline). Based on previous findings, we aim to detect a clinically relevant effect size of 0.4. This effect size was previously shown to be feasible in a similar population (18). The clinically relevant difference on the PCS ranges from 3 to 5 points (51). Based on an estimate of the standard deviation of the PCS score of 10 (52), our effect size corresponds to a 4-point difference between the intervention group and the CAU group. We assume a two-sided significance level of 5% and a power of 80%. The ratio of the number of subjects in the compared groups is 1:1 and there are two measurements of follow-up for our primary outcome. Although we assume that the results will be similar on both repeated measurements, we opted for a more conservative correlation coefficient of 0.5. Using linear mixed models with these values leads to a sample size of n=74 patients per condition.

However, since our study is a cluster randomised controlled trial, we applied an additional correction for the 'design effect' (53) with an expected average cluster size of four. As previous research found that 90% of intraclass correlation coefficients (ICCs) in primary care research are smaller than 0.055 (54), we chose an ICC of 0.05. After applying the correction, the sample size resulted in n=85 patients (1.15 x 74) per condition. Taking a potential dropout rate of 20% into account, we aim to include 106 (85/0.8) patients in each condition.

Statistical analyses

Primary outcomes

Differences in the change scores between the intervention group and the CAU group on the PCS of the RAND-36 will be analysed with linear mixed models according to the intention-to-treat principle as outlined in the Consolidated Standards of Interventional Trials (CONSORT) Statement with extension to cluster randomised trials (55). This analysis technique allows for the clustering of patients within MHNPs and for dependence of observations within individuals over time.

Secondary outcomes

Differences in change scores between the intervention group and the CAU group on the eight separate scales of RAND-36, HADS and PHQ-15 (secondary outcome measures) will also be analysed with linear mixed models according to the intention-to-treat principle.

Economic evaluation

Cost-effectiveness analysis (CEA): Both a cost-effectiveness and cost-utility analysis will be performed from a societal perspective. The time period of the economic evaluation will be 12 months; therefore, discounting is not necessary. Sensitivity analyses will be performed to assess the robustness of the results using different assumptions regarding costs and effects.

Societal costs will be related to the following effect measures in the economic evaluation:

1) Physical functioning as measured by the RAND-36
2) Quality-adjusted life years (QALYs) based on the Dutch tariff for the EuroQol (EQ-5D-5L) (41)
3) Severity of somatization (PHQ-15) and mental health (HADS)

The analysis will be done according to the intention-to-treat principle. Missing cost and effect data will be imputed using a multiple imputation technique. Incremental cost-effectiveness ratios (ICERs) will be calculated by dividing the difference in mean total costs between the groups, by the difference in mean effects between the groups. Bootstrapping with 5000 replications will be used to estimate 95% confidence intervals around cost differences and the uncertainty surrounding the ICERs. Rubin's rules will be used to pool the results from the different multiply imputed datasets. Uncertainty surrounding the ICERs will graphically be presented on cost-effectiveness planes. Cost-effectiveness acceptability curves showing the probability that the intervention was cost-effective in comparison with usual care for a range of different ceiling ratios will also be estimated (56). We will adjust for confounders and effect modifiers, when necessary.

Budget impact analysis (BIA): In BIA, the cost-effectiveness of the short-term psychological intervention and usual care will be extrapolated using a simple Markov model over a period of 5 years based on the estimates obtained from the proposed study. Societal, government and insurer perspectives will be considered. Different scenarios will be evaluated including the following: 1) the intervention is not implemented, i.e. all patients receive usual care, 2) the intervention is offered to the whole patient population,

3) the intervention is implemented over a period of 4 years (25% of the patient population per year) and 4) the intervention is only offered to specific subgroups of the potential patient population. These subgroups will be defined based on the results of the study (e.g. subgroups that particularly benefitted from the intervention).

The total number of patients eligible for the intervention will be estimated based on Dutch incidence and prevalence rates of USD. Resource utilisation will be calculated by multiplying the number of eligible patients with the resource utilisation rates obtained from the cost-effectiveness analysis. Different prices will be used to value resource use depending on the perspective of the analysis: Dutch standard costs for the societal perspective, actual Dutch Healthcare Authority (NZA) tariffs for the government perspective, and average NZA tariffs for the insurer perspective. Both resource use and annual costs will be presented over a 5-year period for all perspectives. Aggregated and disaggregated (e.g. GP care, secondary care, and productivity losses) total costs per year will be presented for the different perspectives and scenarios.

Analysis of patients most likely to benefit from the intervention: if power allows, interaction terms between potential effect modifiers (e.g. severity, comorbidity, gender) and group allocation will be tested in mixed regression models. In case of insufficient power, only exploratory analyses will be done.

Mediation analyses: mediation analyses will be conducted to determine whether the intervention affected physical functioning through changes in problem-solving skills, health anxiety, illness perceptions, coping, mastery and/or working alliance. The Krull and MacKinnon method (57) will be used for this purpose.

Factors influencing implementation: data will be obtained on participation rate, satisfaction with the intervention and characteristics of non-responders. Data collected from non-responders will be limited to demographic characteristics, such as age and gender and reasons for not participating in the study. Impeding and facilitating factors for implementation as observed by MHNPs will be ascertained by conducting interviews with the latter. The interviews will be transcribed verbatim. The transcripts will be analysed by coding the texts using Atlas.ti and themes will be identified and described.

DISCUSSION

To date, this is the first study to evaluate the effectiveness and cost-effectiveness of a short CBT-based treatment provided by MHNPs for patients with undifferentiated

somatoform disorder in primary care versus usual care. The aim of the treatment is to improve physical functioning by increasing problem-solving skills to cope with consequences of the physical symptoms and with everyday problems in general. We assume that higher physical functioning and quality of life will result in lower health care-related and work-related costs.

A strength of this study is that the intervention is provided within the patients' own general practices. Given the prevalence of USD and the large societal costs that accompany this disorder, there is an urgent need to provide easily accessible and affordable treatment for patients with USD. A previous study showed that providing psychological help in general practice was effective in improving quality of life (18). However, the intervention in this study was conducted in a group setting by psychologists from secondary care who offered treatment to patients in general practice. This might not always be feasible and time-efficient. MHNPs, working in general practices and trained in providing (short) psychological treatment, are in a much more convenient position to help patients.

Furthermore, receiving psychological treatment in the patients' general practice may create a safe and low-threshold environment for patients with USD, especially those who would otherwise not seek psychological help in the mental health care setting. Also, by focussing on the consequences and not on the causes of physical symptoms, the possible struggle about the cause of the symptoms is avoided. Regardless of the cause, patients suffer from consequences of USD. This approach may create more acceptance from the patients.

By testing our intervention directly within general practices and with MHNPs who are already employed there, this study has a high clinical relevance. If successful, the intervention is likely to be easily implemented in daily practice as the number of MHNPs employed in the general practices is growing. This is especially relevant in The Netherlands, as more emphasis is being placed on provision of mental health care services in general practice as a result of organisational changes in health care services.

Additionally, the current intervention combines a cognitive-behavioural theoretical framework with PST intervention techniques for somatoform disorders. To date, PST has been widely investigated in depression (32, 58, 59) but rarely in somatoform disorders. One preliminary study containing 11 subjects who received PST found that PST was acceptable to patients and reduced symptoms, hypochondriacal preoccupation and psychiatric morbidity (35). Another study investigated PST in 162 patients and found a

positive impact on symptoms, functioning and costs, but in this study PST was combined with other CBT techniques (18).

A possible limitation to our study is that we use the diagnosis undifferentiated somatoform disorder (USD) as defined by the DSM-IV. The fifth edition of the DSM (DSM-5) has a new classification for the previous category somatoform disorders. Previously, the category of somatoform disorders included somatization disorder, undifferentiated somatoform disorder, conversion disorder, pain disorder, hypochondriasis, body dysmorphic disorder and somatoform disorder not otherwise specified. All of these somatoform disorders except hypochondriasis, body dysmorphic disorder and conversion disorder have now been categorised under somatic symptom disorder (SSD). Furthermore, for the classification of SSD the somatic symptoms do not have to be medically unexplained. Our study was designed and funded before DSM-5 was introduced. SSD is a new DSM-5 classification and issues such as its usefulness and accuracy are currently a topic of debate in the field in addition to proposals for modifications of the criteria for SSD (60–62). Moreover, since its introduction, the DSM-5 has not yet been widely used in research and practice, and no appropriate diagnostic instrument equivalent to the SCID or the Mini International Neuropsychiatric Interview (MINI) was available at the time of recruitment of patients. Therefore, for practical reasons it was impossible to diagnose patients reliably and validly according to the DSM-5. Recently the Health Preoccupation Diagnostic Interview (HPDI), a new structured diagnostic interview for the classification of somatic symptom disorder and illness anxiety disorder, has been developed by Axelsson and colleagues (63). However, this diagnostic interview has not yet been validated.

Another limitation is that, due to the nature of the intervention, it is not possible to blind patients, health care providers and researchers to treatment allocation.

A final point of consideration concerns the primary outcome measure, the physical component summary score (PCS) of the RAND-36. The PCS is a general physical functioning scale, comprising subscales measuring physical aspects of health. Because the PCS is a summary score, the total score may be somewhat insensitive to change, as potential changes on the separate subscales may not lead to a difference in the total score. This may make it more difficult to detect an effect on the PCS. However, several previous studies have successfully used the PCS scale as a primary or secondary outcome and were able to detect change (18, 64–67). Despite potential shortcomings, we also chose the PCS as the primary outcome measure, because a more specific measure, such as the subscale 'physical pain' of the RAND-36, may focus on only a part

of physical functioning, whereas we aim to investigate physical functioning in a more generic manner. After all, the aim of the intervention is not to reduce the symptoms, but to improve physical functioning as a whole. We will also separately investigate the changes on the separate subscales as a secondary outcome in order to see whether more specific changes take place.

Overall, if this study shows that the treatment is effective and cost-effective, the treatment could result in great benefits in primary care by making psychological treatments more available to patients with USD and providing broader treatment possibilities for primary care professionals. Quality of life of patients with USD may be improved and health care costs may be reduced.

ETHICAL APPROVAL AND CONSENT TO PARTICIPATE

This study was approved by the VU Medical Centre Ethics Committee on 9 July 2014, reference number 2014.305. It will be conducted according to the principles of the Declaration of Helsinki (version 2013). Informed consent will be obtained from all participants.

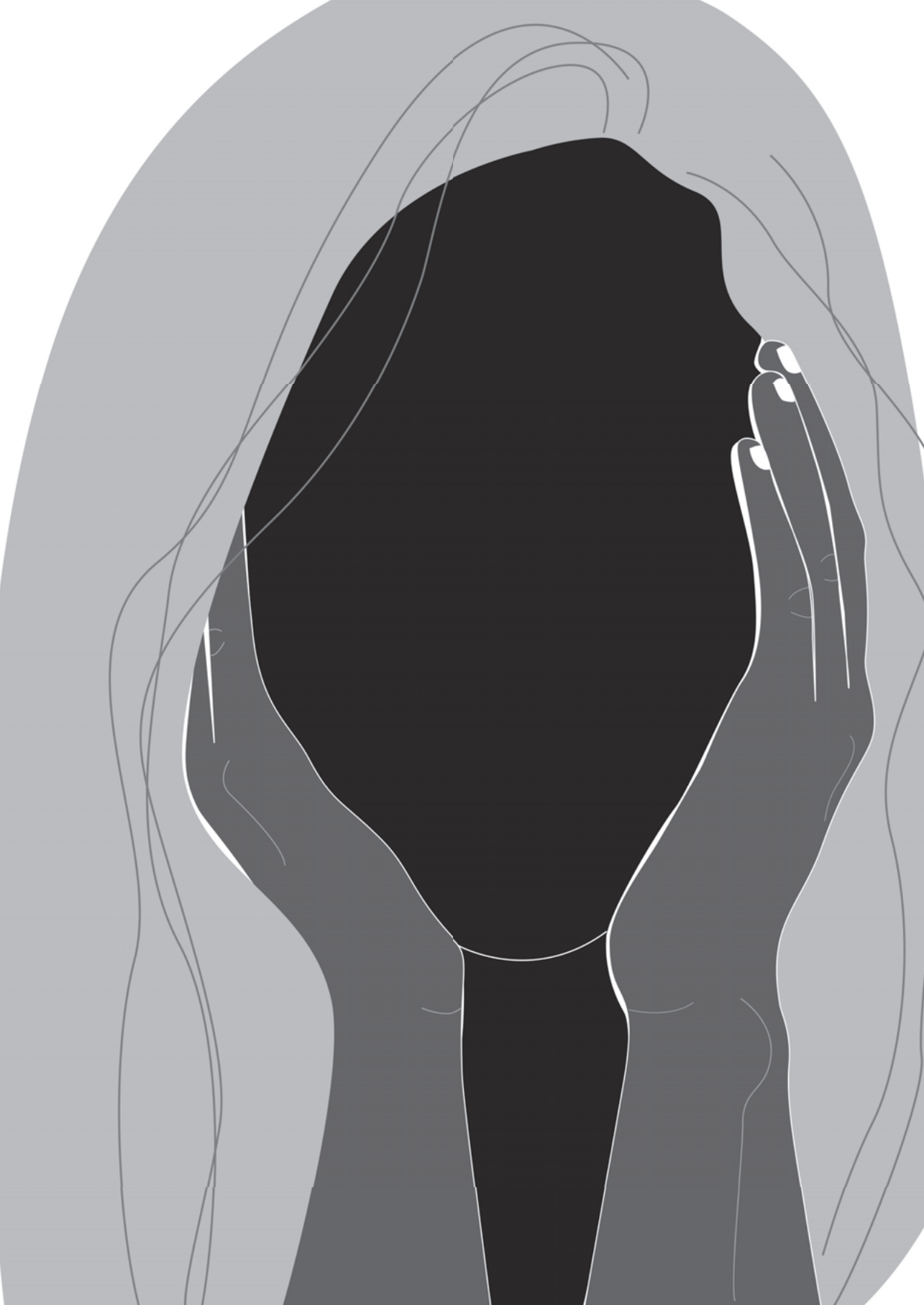
REFERENCES

1. Henningsen P, Zipfel S, Herzog W. Management of functional somatic syndromes. Lancet. 2007;369(9565):946–55.
2. Burton C. Beyond somatisation: a review of the understanding and treatment of medically unexplained physical symptoms (MUPS). Br J Gen Pract. 2003;53(488):231–9.
3. Steinbrecher N, Koerber S, Frieser D, Hiller W. The prevalence of medically unexplained symptoms in primary care. Psychosomatics. 2011;52:263–71.
4. Haller H, Cramer H, Lauche R, Dobos G. Somatoform disorders and medically unexplained symptoms in primary care. Dtsch Arztebl Int. 2015;112:279–87.
5. De Waal M, Arnold I, Eekhof J, van Hemert A. Somatoform disorders in general practice: Prevalence, functional impairment and comorbidity with anxiety and depressive disorders. Br J Psychiatry. 2004;184(6):470–6.
6. American Psychiatric Association. Diagnostic and Statistical Manual of Mental Disorders. 4th ed. Washington, DC: Arlington, VA; 2000.
7. Mergl R, Seidscheck I, Allgaier A, Möller H, Hegerl U, Henkel V. Depressive, anxiety, and somatoform disorders in primary care: prevalence and recognition. Depress Anxiety. 2007;24(3):185–95.
8. Janssens KAM, Zijlema WL, Joustra ML, Rosmalen JGM. Mood and anxiety disorders in chronic fatigue syndrome, fibromyalgia, and irritable bowel syndrome: results from the LifeLines cohort study. Psychosom Med. 2015;77:449–57.
9. Konnopka A, Schaefert R, Heinrich S, Kaufmann C, Luppa M, Herzog W, et al. Economics of medically unexplained symptoms: a systematic review of the literature. Psychother Psychosom. 2012;81(5):265–75.
10. Barsky AJ, Orav EJ, Bates DW. Somatization increases medical utilization and costs independent of psychiatric and medical comorbidity. Arch Gen Psychiatry. 2005;62(8):903–10.
11. Bermingham SL, Cohen A, Hague J, Parsonage M. The cost of somatisation among the working-age population in England for the year 2008-2009. Ment Health Fam Med. 2010;7(2):71–84.
12. Gustavsson A, Svensson M, Jacobi F, Allgulander C, Alonso J, Beghi E, et al. Cost of disorders of the brain in Europe 2010. Eur Neuropsychopharmacol. 2012;22(3):237–8.
13. Mack S, Jacobi F, Gerschler A, Strehle J, Höfler M, Busch MA, et al. Self-reported utilization of mental health services in the adult German population – evidence for unmet needs? Results of the DEGS1-Mental Health Module (DEGS1-MH). Int J Methods Psychiatr Res. 2014;23(3):289–303.
14. Van Dessel N, Den Boeft M, Van der Wouden J, Kleinstäuber M, Leone S, Terluin B, et al. Non-pharmacological interventions for somatoform disorders and medically-unexplained physical symptoms (MUPS) in adults. Cochrane Database Syst Rev. 2014;11:CD01114.
15. Kroenke K, Swindle R. Cognitive-behavioral therapy for somatization and symptom syndromes: a critical review of controlled clinical trials. Psychother Psychosom. 2000;69:205–15.
16. Watts SE, Turnell A, Kladnitski N, Newby JM, Andrews G. Treatment-as-usual (TAU) is anything but usual: a meta-analysis of CBT versus TAU for anxiety and depression. J Affect Disord. 2015;175:152–67.
17. Schaefert R, Kaufmann C, Wild B, Schellberg D, Boelter R, Faber R, et al. Specific collaborative group intervention for patients with medically unexplained symptoms in general practice: a cluster randomized controlled trial. Psychother Psychosom. 2013;82(2):106–19.
18. Zonneveld LNL, van Rood YR, Timman R, Kooiman CG, Van't Spijker A, Busschbach JJ V. Effective group training for patients with unexplained physical symptoms: a randomized controlled trial with a non-randomized one-year follow-up. PLoS One. 2012;7(8):e42629.
19. Kleinstäuber M, Witthöft M, Steffanowski A, van Marwijk H, Hiller W, Lambert MJ. Pharmacological interventions for

somatoform disorders in adults. Cochrane Database Syst Rev. 2014;7(11).
20. Wortman MSH, Lucassen PLBJ, van Ravesteijn HJ, Bor H, Assendelt P, Lucas C, et al. Brief multimodal psychosomatic therapy in patients with medically unexplained symptoms: feasibility and treatment effects. Fam Pract. 2016;33(4):346–53.
21. Wileman L, May C, Chew-Graham CA. Medically unexplained symptoms and the problem of power in the primary care consultation: a qualitative study. Fam Pract. 2002;19(2):178–82.
22. Reid S, Whooley D, Crayford T, Hotopf M. Medically unexplained symptoms-GPs' attitudes towards their cause and management. Fam Pract. 2001;18(5):519–23.
23. Peters S, Stanley IAN, Rose M, Salmon P. Patients with medically unexplained symptoms: sources of patients' authority and implications for demands on medical care. Soc Sci Med. 1998;46:559–65.
24. Grandes G, Montoya I, Arietaleanizbeaskoa M, Arce V, Sanchez A. The burden of mental disorders in primary care. Eur Psychiatry. 2011;26(7):428–35.
25. Cape J, Barker C, Buszewicz M, Pistrang N. General practitioner psychological management of common emotional problems (I): Definitions and literature review. Br J Gen Pract. 2000;50(453):313–8.
26. Robbins J, Kirmayer L, Hemami S. Latent variable models of functional somatic distress. J Nerv Ment Dis. 1997;185(10):606–15.
27. Kroenke K, Spitzer RL, Williams JB. The PHQ-15: validity of a new measure for evaluating the severity of somatic symptoms. Psychosom Med. 2002;64(2):258–66.
28. Zonneveld LNL, van 't Spijker A, Passchier J, van Busschbach JJ, Duivenvoorden HJ. The effectiveness of a training for patients with unexplained physical symptoms: protocol of a cognitive behavioral group training and randomized controlled trial. BMC Public Health. 2009;9:251.
29. Van Rood Y, van Ravesteijn H, de Roos C, Spinhoven P, Speckens A. [Protocol for diagnosis and treatment of patients with medically unexplained physical symptoms: The consequences model] (in Dutch). Keijsers, van Minnen & Hoogduin (eds): [Protocolized treatments for adults with psychological complaints 2] (in Dutch). Amsterdam: Boom; 2011. p. 15–47.
30. Speckens A, van Hemert A, Spinhoven P, Hawton KE, Bolk JH, Rooijmans HG. Cognitive behavioural therapy for medically unexplained physical symptoms: a randomised controlled trial. Br Med J. 1995;311(7016):1328–32.
31. Zonneveld LNL, Duivenvoorden HJ, Passchier J, Spijker A Van. Tailoring a cognitive behavioural model for unexplained physical symptoms to patient's perspective: a bottom-up approach. Clin Psychol Psychother. 2010;17(6):528–35.
32. Mynors-Wallis LM, Gath DH, Day A, Baker F. Randomised controlled trial of problem solving treatment, antidepressant medication, and combined treatment for major depression in primary care. Br Med J. 2000;320:26–30.
33. Mynors-Wallis L, Gath D, Lloyd-Thomas A, Tomlinson D, Gath D, Tomlinson D. Randomised controlled trial comparing problem solving treatment with amitriptyline and placebo for major depression in primary care. BMJ. 1995;310(6977):441–5.
34. Hassink-Franke LJA, van Weel-Baumgarten EM, Wierda E, Engelen M, Beek M, Bor H, et al. Effectiveness of problem-solving treatment by general practice registrars for patients with emotional symptoms. J Prim Health Care. 2011;3(3):181–9.
35. Wilkinson P, Mynors-Wallis L. Problem-solving therapy in the treatment of unexplained physical symptoms in primary care: a preliminary study. J Psychosom Res. 1994;38(6):591–8.
36. Olde Hartman T, Blankenstein A, Molenaar A, Bentz van den Berg D, van der Horst H, Arnold I, et al. [NHG guideline on medically unexplained symptoms (MUS)] (in Dutch). Huisarts Wet. 2013;56(5):222–30.
37. Hays RD, Morales LS. The RAND-36 measure of health-related quality of life. Ann Med. 2001;33:350–7.
38. Van der Zee K, Sanderman R, Heyink

J, de Haes H. Psychometric qualities of the RAND 36-item health survey 1.0: A multidimensional measure of general health status. Int J Behav Med. 1996;3(2):104–22.
39. Van der Zee K, Sanderman R. [Measuring general health with the RAND-36, a guideline] (in Dutch). 2nd ed. Groningen; 2012.
40. Herdman M, Gudex C, Lloyd A, Janssen M, Kind P, Parkin D, et al. Development and preliminary testing of the new five-level version of EQ-5D (EQ-5D-5L). Qual Life Res. 2011;20(10):1727–36.
41. Versteegh MM, Vermeulen KM, Evers SMAA, Wit GA De, Prenger R, Stolk EA. Dutch tariff for the five-level version of EQ-5D. Value Heal. 2016;19(4):343–52.
42. Hakkaart-van Roijen L, van Straten A, Donker M, Tiemens B. Manual Trimbos/ IMTA questionnaire for costs associated with psychiatric illness (TIC-P). Rotterdam: Institute for Medical Technology Assessment, Erasmus Univeristy; 2002.
43. Zigmond A, Snaith R. The hospital anxiety and depression scale. Acta Psychiatr Scand. 1983;67(6):361–70.
44. American Psychiatric Association. Diagnostic and statistical manual of mental disorders. 5th ed. Arlington, VA: American Psychiatric Publishing; 2013.
45. Dreer LE, Berry J, Rivera P, Elliott TR, Miller D, Little TD. Efficient assessment of social problem-solving abilities in medical and rehabilitation settings: a Rasch analysis of the Social Problem-Solving Inventory-Revised. 2009;65(7):653–69.
46. Pilowsky I. Dimensions of hypochondriasis. Br J Psychiatry. 1967;113(494):89–93.
47. Broadbent E, Petrie KJ, Main J, Weinman J. The brief illness perception questionnaire. J Psychosom Res. 2006;60(6):631–7.
48. Moss-Morris RE, Chalder T. Illness representations: where to from here? 16th Conference of the European Health Psychology Society, Kos. 2003.
49. Pearlin L, Schooler C. The structure of coping. J Health Soc Behav. 1978;19(1):2–21.
50. Stinckens N, Elliott R, Leijssen M. Bridging the gap between therapy research and practice in a person-centered/experiential therapy training program: the Leuven systematic case study research protocol. Pers Exp Psychother. 2009;8:143–62.
51. Samsa G, Edelman D, Rothman M, Willieams G, Lipscomb J, Matchar D. Determining clinically important differences in health status measures: a general approach with illustration to the Health Utilities Index Mark II. Pharmacoeconomics. 1999;15:141–55.
52. Farivar SS, Cunningham WE, Hays RD. Correlated physical and mental health summary scores for the SF-36 and SF-12 Health Survey, V.I. Health Qual Life Outcomes. 2007;5:54.
53. Campbell MK, Fayers PM, Grimshaw JM. Determinants of the intracluster correlation coefficient in cluster randomized trials: the case of implementation research. Clin Trials. 2005;2:99–107.
54. Adams G, Gulliford MC, Ukoumunne OC, Eldridge S, Chinn S, Campbell MJ. Patterns of intra-cluster correlation from primary care research to inform study design and analysis. J Clin Epidemiol. 2004;57(8):785–94.
55. Campbell M, Piaggio G, Elbourne D, Altman D. Consort 2010 statement: extension to cluster randomised trials. BMJ. 2012;345:e5661.
56. Fenwick E, O'Brien BJ, Briggs A. Cost-effectiveness acceptability curves - facts, fallacies and frequently asked questions. Health Econ. 2004;13(5):405–15.
57. Krull J, MacKinnon D. Multilevel mediation modeling in group-based intervention studies. Eval Rev. 2016;23(4):418–44.
58. Bell AC, D'Zurilla TJ. Problem-solving therapy for depression: a meta-analysis. Clin Psychol Rev. 2009;29(4):348–53.
59. Cuijpers P, van Straten A, Warmerdam L. Problem solving therapies for depression: a meta-analysis. Eur Psychiatry. 2007;22(1):9–15.
60. Frances A. The new somatic symptom disorder in DSM-5 risks mislabeling many people as mentally ill. BMJ. 2013;346:f1580.
61. Rief W, Martin A. How to use the new DSM-5 somatic symptom disorder diagnosis in research and practice: a Critical evaluation and a proposal for modifications. Annu Rev Clin Psychol. 2014;10:339–67.

62. Katz J, Rosenbloom B, Fashler S. Chronic pain, psychopathology, and DSM-5 somatic symptom disorder. Can J Psychiatry. 2015;60(4):160–7.
63. Axelsson E, Andersson E, Ljótsson B, Wallhed Finn D, Hedman E. The health preoccupation diagnostic interview: inter-rater reliability of a structured interview for diagnostic assessment of DSM-5 somatic symptom disorder and illness anxiety disorder. Cogn Behav Ther. 2016;45:259–69.
64. Katsamanis M, Lehrer PM, Escobar JI, Gara M a, Kotay A, Liu R. Psychophysiologic treatment for patients with medically unexplained symptoms: a randomized controlled trial. Psychosomatics. 2011;52(3):218–29.
65. Gili M, Magallón R, López-Navarro E, Roca M, Moreno S, Bauzá N, et al. Health related quality of life changes in somatising patients after individual versus group cognitive behavioural therapy: a randomized clinical trial. J Psychosom Res. 2014;76(2):89–93.
66. Sattel H, Lahmann C, Gündel H, Guthrie E, Kruse J, Noll-Hussong M, et al. Brief psychodynamic interpersonal psychotherapy for patients with multisomatoform disorder: randomised controlled trial. Br J Psychiatry. 2012;200(1):60–7.
67. Schröder A, Rehfeld E, Ornbøl E, Sharpe M, Licht RW, Fink P. Cognitive-behavioural group treatment for a range of functional somatic syndromes: randomised trial. Br J Psychiatry. 2012;200(6):499–507.



Chapter 4

Management of patients with persistent medically unexplained symptoms: a descriptive study

Kate Sitnikova
Renske Pret-Oskam
Sandra MA Dijkstra-Kersten
Stephanie S Leone
Harm WJ van Marwijk
Henriëtte E van der Horst
Johannes C van der Wouden

Published in: BMC Family Practice 2018;19:88.

ABSTRACT

Background: In 2013 the Dutch guideline for management of medically unexplained symptoms (MUS) was published. The aim of this study is to assess medical care for patients with persistent MUS as recorded in their electronic medical records, to investigate if this is in line with the national guideline for persistent MUS and whether there are changes in care over time.

Methods: We conducted an observational study of adult primary care patients with MUS. Routinely recorded health care data were extracted from electronic medical records of patients participating in an ongoing randomised controlled trial in 30 general practices in the Netherlands. Data on general practitioners' (GPs') management strategies during MUS consultations were collected in a 5-year period for each patient prior. Management strategies were categorised according to the options offered in the Dutch guideline. Changes in management over time were analysed.

Results: Data were collected from 1035 MUS consultations (77 patients). Beside history-taking, the most frequently used diagnostic strategies were physical examination (24.5%) and additional investigations by the GP (11.1%). Frequently used therapeutic strategies were prescribing medication (24.6%) and providing explanations (11.2%). As MUS symptoms persisted, GPs adjusted medication, discussed progress and scheduled follow-up appointments more frequently. The least frequently used strategies were exploration of all complaint dimensions (i.e. somatic, cognitive, emotional, behavioural and social) (3.5%) and referral to a psychologist (0.5%) or psychiatrist (0.1%).

Conclusions: Management of Dutch GPs is partly in line with the Dutch guideline. Medication was possibly prescribed more frequently than recommended, whereas exploration of all complaint dimensions, shared problem definition and referral to mental health care were used less.

BACKGROUND

Medically unexplained symptoms (MUS), i.e. physical symptoms that cannot entirely be accounted for by a known somatic disease, are extremely common in primary care (1, 2). Although most such symptoms are self-limiting, in some cases they persist and impair patients' functioning (3). In the latter case, persisting MUS may meet diagnostic criteria for (undifferentiated) somatoform disorder of the psychiatric classification system DSM-IV (4). Since the introduction of DSM-5, somatoform disorders have been replaced by somatic symptom disorders (5). The main criteria for somatic symptom disorder no longer require the nature of physical symptoms to be unexplained, but focus on maladaptive cognitions, emotions and/or behaviour with respect to the physical symptom(s).

The prevalence of persistent MUS, such as those classified as somatoform disorders, is 3–10% in general practice (6, 7, 8). Persistent MUS are disabling and are associated with high rates of comorbid mental health disorders (6, 9, 10). There are high direct and indirect health care costs due to increased health care use and productivity loss due to sickness absence (11).

Previous research shows that GPs may view MUS patients as challenging, as it can be difficult for the GP to exclude the possibility of a serious illness and at the same time satisfy patients' concerns about their health (12). GPs may develop a sense of uncertainty in their professional knowledge (12, 13, 14) and patients may be left feeling that their symptoms are not being taken seriously (13).

To aid GPs in the management of patients with MUS, the Dutch College of General Practitioners published a guideline in 2013 (15). Previous guidelines for MUS have also been published in Germany (16) and England (17). The diagnostic recommendations in the Dutch guideline include ample exploration of all dimensions of complaints (i.e. somatic, cognitive, emotional, behavioural and social dimensions) and a thorough physical examination. The GP should be cautious with additional investigations and diagnostic referrals and should evaluate the severity of the symptoms or a change in symptoms over time. The therapeutic recommendations describe a stepped-care process in three steps, in which the GP starts with the mildest possible treatment and intensifies treatment when there are no adequate results.

It is unclear what current management for persisting MUS entails and whether this is in line with the Dutch guideline. Although GPs' perceptions about giving explanations to

patients with persistent MUS have previously been investigated in a Dutch focus group study, the actual management strategies were not described (18).

The aim of this descriptive study is to gain more insight into the management of adult patients with persistent MUS that meet criteria for an undifferentiated somatoform disorder in Dutch general practice and its potential change in time, as recorded in the patients' medical records. We also aim to investigate to what extent this care is in line with the national guideline published by the Dutch College of General Practitioners.

METHOD

Study design and patient selection

We analysed the longitudinal electronic medical record data of persons participating in an ongoing randomised controlled trial (RCT) called the CIPRUS study. The CIPRUS study aims to establish the effectiveness of treatment of undifferentiated somatoform disorder by a mental health nurse practitioner (MHNP) within general practice, versus usual care. The design of the CIPRUS study has been described elsewhere in more detail (19). Potential participants were identified by running a search of the electronic medical records for patients who had consulted their GP at least twice in the previous 3 months with one or more complaints from the Robbins list (20). The Robbins list consists of 23 physical symptoms that are associated with functional somatic syndromes. GPs then checked the selected patients to verify that these patients indeed had MUS according to them, and excluded patients who fulfilled one or more exclusion criteria. Exclusion criteria were: presence of a medical or psychological disorder explaining the symptoms, presence of a severe psychiatric disorder, currently receiving psychological treatment for MUS, having poor language skills or handicap that prevented patients from understanding the intervention. After patients were verified as having MUS by their GP, they were interviewed using a structured clinical interview (SCID-I) in order to determine whether they fulfilled the DSM-IV criteria for undifferentiated somatoform disorder (USD) (21). Those fulfilling the criteria for USD were included in the current study. All participating patients gave written informed consent to extract data from their electronic medical records. In the current study, we used data from patients participating in the usual care group of the CIPRUS study. We used data from the group of MUS patients receiving usual care because we wanted to know which care patients received. Because we had already identified these patients as having MUS we used the data from this group. Obviously we could not use data from the intervention group, as the intervention consisted of a number of scheduled meetings with the MHNP within the general practice, which would also be recorded in the electronic medical records. This

data would, therefore, not only reflect usual care, but also care provided due to being part of the intervention group of our trial.

Data extraction

Data were manually extracted by 3 researchers from electronic medical records of all participating patients between 21 November 2016 until 31 August 2017 in 30 participating general practices. Data were extracted for all MUS consultations in the 5-year time period for each patient prior to the search date. Data were collected from fields for prescription, medical tests and referrals, and from the GPs' free text notes. Extracted data consisted of the date of the consultation, International Classification of Primary Care (ICPC) code corresponding to the consultation, and the management strategy of the GP, i.e. GPs' own notes in the electronic medical records on what was carried out during the consultation and what they were planning or had arranged to do as a next step. For persons who were younger than 18 years of age during the 5-year time period, data were only collected from age 18 onward. MUS consultations were defined as consultations in which the GPs used ICPC codes that corresponded with the symptoms from the Robbins list (Table 1) (20). Because there are no suitable corresponding ICPC codes for the symptoms 'restlessness' and 'thoughts slower', these two items from the Robbins list were omitted. Data from consultations that were not coded with ICPC codes corresponding to the Robbins list, but where the GP had noted 'MUS' or 'somatisation' were also collected.

Table 1. Robbins list and corresponding ICPC codes

Symptoms from the Robbins list	Corresponding ICPC codes	Number of consultations (%)	
Back pain	L01, L02, L03	179	(17.3)
Joint pain	L20	28	(2.7)
Extremity pain	L18*	114	(11.0)
Headaches	N01, N02	69	(6.7)
Weakness/fatigue	A04	132	(12.8)
Chronic fatigue syndrome	A04.01	65	(6.3)
Sleep disturbance	P06*	106	(10.2)
Difficulty concentrating	P20	2	(0.2)
Loss of appetite	T03	0	(0.0)
Weight change	T07, T08	9	(0.9)
Restlessness	N/A	N/A	
Thoughts slower	N/A	N/A	
Chest pain	L04	38	(3.7)
Shortness of breath	R02	7	(0.7)

Table 1. Continued

Symptoms from the Robbins list	Corresponding ICPC codes	Number of consultations (%)	
Palpitations	K04	24	(2.3)
Dizziness	N17*	29	(2.8)
Lump in throat	R21*	28	(2.7)
Numbness	N06*	5	(0.5)
Nausea	D09	16	(1.5)
Loose bowels	D11	20	(1.9)
Gas or bloating	D08	0	(0.0)
Constipation	D12	36	(3.5)
Abdominal pain	D01	57	(5.5)
Other (not part of the Robbins list)	A97, D93, P75, P78	71	(6.9)

* including subcodes

N/A: not applicable

Data categorisation

After collection, the extracted data on management were categorised by one researcher (KS) according to the options for diagnosis and treatment in the current Dutch GP guideline (15). The categories from the guideline used for classifying diagnostic strategies were exploration of all complaint dimensions, physical examination and additional diagnostic testing within and outside general practice (diagnostic referral). The categories used for classifying treatment strategies were shared problem definition, education and explanation, advice, treatment with medication, setting up a time contingent plan, scheduling follow-up appointments, referral to other primary care providers, and referral to secondary care (15). Within primary care a patient can be referred to other care providers such as a (psychosomatic) physiotherapist or exercise therapist, mental health nurse practitioner, primary care social psychiatric nurse or primary care psychologist (e.g. trained in cognitive behavioural therapy) (15).

Data analysis

Data were analysed using SPSS version 22 for Windows. We used descriptive statistics to describe the study population and the management strategies. In order to determine whether there were any trends of providing various management strategies over time, we used cross-tabs and the chi-square test for trend.

RESULTS

Figure 1 presents a flow chart of patients included in this study. The control group of the CIPRUS study consisted of 96 patients in total. Seventeen patients dropped out of the study, and did not give permission to collect data from their medical records. Therefore, data were collected from 79 patients. For two patients, no information on MUS consultations was found in their electronic medical records. Therefore, data from 77 patients were available. The GPs registered a total of 1035 MUS consultations for these patients, of which 13.6% took place before 2013, the year in which the Dutch GP guideline was published.

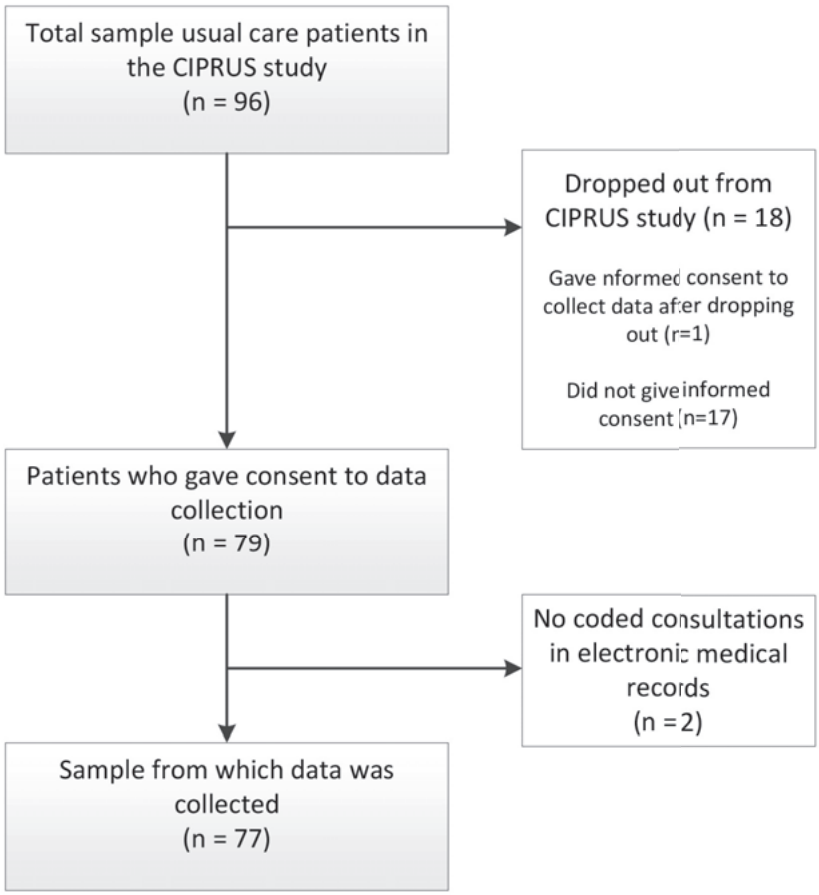


Figure 1. Flow chart of MUS patients
CIPRUS study = Cognitive-behavioural Intervention for PRimary care patients with Undifferentiated Somatoform disorder

Characteristics of patients with persistent MUS

Of the 77 patients, 80.5% were female. The mean age was 50 (SD: 17,1, range: 19–89). Over the 5 year period, the mean number of MUS consultations was 13 (SD: 17, range: 1–130), resulting in an average of 2,6 consultations a year. The symptoms patients presented with are provided in Table 1. The most frequently recorded symptoms were back pain (17,3%), weakness or fatigue (12,8%), extremity pain (11%) and sleep disturbance (10,2%). No consultations had codes for loss of appetite and gas or bloating. Seventy-one consultations (6,9%) were coded with codes other than those that appear on the Robbins list, however the GP had referred to MUS in these consultations. The codes used in this category were hysteria/hypochondria (ICPC code P75), neurasthenia/stress (P78), spastic colon/irritable bowel syndrome (D93), and 'no disease' (A97).

Recorded management strategies

GPs varied in the way they recorded what was done during the consultations. This varied per GP as well as per patient and per consultation. Table 2 provides examples of data extracted from electronic medical records of three patients. There are 2 examples of brief records (≤10 words), 2 examples of medium-length records (11–30 words) and 2 examples of long records (≥31 words). The table also illustrates how these were categorized according to the GP guideline categories.

An overview of the strategies the GPs used in the 1035 consultations is provided in Table 3. The most common diagnostic strategies were physical examination (24.5% of consultations, range among GPs 7.0–66.7%) and additional investigations within the GP practice (14.6%, range 0–50%). Of the additional investigations, laboratory tests such as various blood, urine and feces tests were done most frequently (11.1%, range 0–37.5%). Symptom exploration was recorded in 3.5% of the consultations (range 0–20.0%) and found among 40% of the GPs. Having administered the recommended symptom checklist enquiring about distress, depression, anxiety and somatisation symptoms, the 4-Dimensional Symptom Questionnaire (4DSQ) (22), was only recorded once (0.1%).

The most common treatment strategies were treatment with medication (24.6%, range 0–62.5%), followed by discussing progress (16.2%, range 0–41.5%), scheduling follow-up appointments (11.8%, range 0 –33.3%), vitamin pills/injections (11.7%, range 0–36.8%, recorded by less than a quarter of the GPs, mainly in the same patients), providing education and explanation (11.2%, range 0–35.7%) and giving advice (10.8%, range 0–42.3%). Wait and see strategies were also recorded frequently (9.4%, range 0–40.0%). Medication requiring a prescription was prescribed most (at least 19.4% of all treatment strategies, range 0 –40.0%). NSAIDs were prescribed most frequently (3.6% of

Table 2. Examples of categorisation of data extracted from electronic medical records

Patient	Length of record	Information extracted from medical records:	Categorised as:
113	≤10 words	Had a talk. Gave explanation.	*Diagnostic*: None *Therapeutic*: - Education and explanation - Discussing progress - Other: "talk"
171	≤10 words	Physical examination, referral to neurologist	*Diagnostic*: - Physical examination *Therapeutic*: - Referral to secondary care (unclear for diagnostics or treatment): neurologist
115	11-30 words	Explained that I don't know whether a scan is indicated, but that due to the long duration of complaints we can ask for the orthopedist's opinion: referral	*Diagnostic*: - Diagnostic referral *Therapeutic*: - Education and explanation
123	11-30 words	Stop tramal, start fentanyl patch, and follow up appointment after 2 weeks, is allergic to diclofenac, developed a rash, fentanyl patch 12 mcg/hr 5 pieces	*Diagnostic*: None *Therapeutic*: - Medication adjustment: discontinuation - Prescribed medication: opioids - Follow-up appointment

Table 2. Continued

Patient	Length of record	Information extracted from medical records:	Categorised as:
158	≥31 words	Carried out physical examination. Exploration. Does not feel reassured despite good lab results and echo abdomen. Will go to exercise therapist and an optometrist for visual test. Will return in a month for an evaluation. If there is insufficient improvement, referral to a psychiatrist. In my opinion no indication of physical cause. Patient will also fill in a diary with symptoms (because complaints are very inconsistent). Explanation when to return sooner.	*Diagnostic*: - Exploration of symptoms - Physical examination - Discussing test results *Therapeutic*: - Education and explanation - Symptom diary - Discussing progress - Follow-up appointment
165	≥31 words	Gave explanation: No somatic problem, no reason to be extra vigilant with normal heartbeat. Talked about the option to talk to the behaviour specialist, is going to do this. Will go to physiotherapist to learn not to focus on his normal heartbeat. Wants to go there as well because wants to hear from a professional whether everything is OK during a workout, prefers not to start a long treatment program (psychosomatic physiotherapy?)	*Diagnostic*: None *Therapeutic*: - Education and explanation - Referral within primary care: other GP consulting another health professional: other

all treatment strategies, range 0–13.5%), followed by psychopharmacological medication (3.4% of all treatment strategies, range 0–20.0%, recorded by almost half of the GPs) and opioids (3.0% of all treatment strategies, range 0–11.5%, also recorded by almost half of the GPs).

Referrals to a psychologist (0.5%, range 0–7.7%, recorded by 17% of the GPs) or a psychiatrist (0.1%), formulation of a shared problem definition (0.4%, range 0–7.7%, recorded by 10% of the GPs) and setting up a time contingent plan (0.1%) were management strategies that were used the least often. When referrals to secondary care were documented, it was often unclear whether the referral was for diagnostic or treatment purposes. Therefore, a category 'referral to secondary care (unclear for diagnostics or treatment)' was added. Finally, of the 44 management strategies categorised as 'other', GPs coded 28 consultations (2.7%) with 'talk', 'listening ear' and 'encouragement'.

Table 3. Overview of management strategies

Management strategies	n of consultations [a]	(%) [b]	n of patients
Diagnostic strategies			
Exploration of symptoms	36	(3.5)	22
Physical examination	254	(24.5)	67
Additional investigations within GP practice	151	(14.6)	61
Laboratory tests	115	(11.1)	57
ECG	17	(1.6)	14
X-ray	22	(2.1)	16
Echography	7	(0.7)	7
Other	7	(0.7)	7
Diagnostic referral	34	(3.3)	27
Discussing test results	62	(6.0)	39
Therapeutic strategies			
Shared problem definition	4	(0.4)	4
Education and explanation	116	(11.2)	46
Advice	112	(10.8)	45
Lifestyle/dietary advice	53	(5.1)	31
Physical exercise advice	52	(5.0)	28
Other advice	17	(1.7)	13

Table 3. Continued

Management strategies	n of consultations [a]	(%) [b]	n of patients
Symptom diary	12	(1.2)	10
Shared plan for symptom management	48	(4.6)	28
Setting up a time contingent plan	1	(0.1)	1
Discussing/giving advice about medication	97	(9.4)	35
Medication	255	(24.6)	65
Over the counter medication (OTC)	69	(6.7)	30
Prescribed medication	201	(19.4)	62
NSAIDs	37	(3.6)	24
Opioids	31	(3.0)	19
Psychopharmacological medication	35	(3.4)	20
Sleeping medication	25	(2.4)	15
Antibiotics	7	(0.7)	6
Other	83	(8.0)	44
Unclear OTC or prescribed medication	6	(0.6)	5
Vitamin pills/injections	121	(11.7)	12
Medication adjustment	71	(6.9)	29
Dose increase	22	(2.1)	15
Dose reduction	18	(1.7)	13
Discontinuation	37	(3.6)	21
Refill prescription	28	(2.7)	11
Referral within primary care	47	(4.5)	31
Physiotherapist	25	(2.4)	18
Mental health nurse practitioner	14	(1.4)	12
Other	9	(0.9)	9
Physiotherapist appointment	27	(2.6)	20
Mental health nurse practitioner appointment	29	(2.8)	8
Referral to secondary care for treatment	26	(2.5)	18
Medical specialist	14	(1.4)	9
Rehabilitation	15	(1.4)	13

Table 3. Continued

Management strategies	n of consultations [a]	(%) [b]	n of patients
Referral to secondary care (unclear for diagnostics or treatment)	46	(4.4)	31
Rheumatologist	11	(1.1)	10
Neurologist	10	(1.0)	9
Gastroenterologist	7	0.7)	6
Internist	6	(0.6)	6
Psychiatrist	1	(0.1)	1
Other	12	(1.1)	12
Referral to a psychologist	5	(0.5)	5
GP consulting another health professional	46	(4.4)	22
Colleague GP	21	(2.0)	6
Secondary care medical specialist	12	(1.2)	9
Other	13	(1.3)	10
Discussing progress	168	(16.2)	52
Follow-up appointment	122	(11.8)	51
Contact if necessary	87	(8.4)	41
Wait and see	97	(9.4)	38
Other	44	(4.3)	26
Unspecified	41	(4.0)	1

[a] Does not add up to 1035 because GPs recorded more than one ICPC codes during one consultation

[b] Does not add up to 100% because GPs recorded more than one ICPC codes during one consultation

Management strategies across time

We conducted chi-square tests for trend for the largest categories of management strategies (n of consultations ≥75). Over the 5 year time period, there appeared to be significant trends in the course of 'giving advice' χ (1) = 5.73, p = 0.017, 'medication adjustment' χ (1) = 11.67, p = 0.001, 'discussing progress' χ (1) = 11.31, p = 0.001, 'scheduling follow-up appointments' χ (1) = 10.75, p = 0.001 and 'contact if necessary' χ (1) = 4.11, p = 0.043. In all the above categories, the proportion of consultations in which the management strategies concerned were provided, increased over time. For all of the above management strategies except 'contact if necessary' there was a small decrease in percentage of the management strategy used within consultations after the first year, after which the percentages increased again. For 'contact if necessary', the percentage

of the consultations increased steadily across time. There were no significant trends over time for the other management strategies.

DISCUSSION

Summary

The most frequent management strategies recorded by Dutch GPs included diagnostic procedures such as physical examinations and additional investigations, and therapeutic procedures such as prescribing medication, discussing progress and providing education, explanation and advice. Other strategies that focus more on listening to the patient and involving patients in their own diagnostic and therapeutic process, and decision making, such as 'exploration of all complaint dimensions', 'shared problem definition', and 'shared plan for symptom management' did not seem to be adopted as frequently. These latter management strategies are especially important for MUS patients (23). Patients were also rarely referred to a psychologist or psychiatrist.

As the symptoms lasted longer, GPs tended to adjust medication more frequently, discuss progress more often, schedule more follow-up appointments and encourage patients more to contact the practice if necessary.

When comparing these strategies to the recommendations in the Dutch guideline, we can conclude that GPs partly used management strategies recommended by the guideline but several essential strategies were missing. This may possibly reflect either the GPs' or the patients' reluctance to seek mental health care for complaints that are perceived to be primarily physical. However, another reason could be that GPs in our sample were not sufficiently familiar with the guideline yet, since the guideline was published during the data extraction period. Even if GPs were familiar with the guideline, it may have taken some time to get used to the new approach, and they may not have started applying strategies, such as exploration of all complaint dimensions, with patients whom they had already seen often before.

Our findings could, however, also point to underreporting of these, more 'conversation-like' management strategies. Recording behaviour varied widely across GPs, so it is impossible to know whether the strategy was not provided or not recorded. Due to time constraints, GPs may only put the more objective management strategies such as results of physical examinations, additional investigations and medication prescriptions in the medical records.

Comparison with existing literature

Several studies investigated management of MUS in other countries. A Norwegian study found that the majority of Norwegian GPs offered supportive counselling (64%), followed by prescribing medication (24%) and additional tests or referrals (20%) (8). In our sample, the rates for prescription of medication (24%) and additional testing or referral (18%) were similar.

An Italian study found that Italian GPs mostly provided reassurance and support, listened to the patient, prescribed medication, ordered further medical tests and provided information (12). In our study, prescribing medication, doing further testing and providing information were also among the most commonly used strategies, however offering reassurance and support and listening to the patient were recorded less frequently. Although the GPs in our study coded 2.7% of their consultations as 'having a talk', 'listening ear' and 'encouragement', 'listening to the patient' was not one of the categories that we used in our classification of management strategies. Also, not all GPs may record their listening behaviour as such in the medical records.

In the dental field adherence to clinical practice guidelines has been found to be up to 72% on average (24). GPs also do not fully adhere to clinical practice guidelines (25, 26). GPs report that they are aware of the guidelines, but find it difficult to implement them with all individual patients, as they may feel that the guideline is not always fitting. GPs may therefore prefer to provide personalized care and let the patients have the final say in their treatment (25, 27). This may also apply to the GPs in our study.

Strengths and limitations

To the best of our knowledge this is the first study to investigate care for persons with MUS in such detail. A strength of this study is that we used real-world data directly from electronic medical records. We were therefore able to collect detailed information about every MUS consultation. Furthermore, we did not rely on self-report instruments such as surveys or interviews taken from GPs. This possibly led to having gathered more 'objective' data, free from various kinds of bias such as recall bias. Another strength is that the choice of categories for classifying management strategies was based on the current Dutch guideline, which provided clear classification options beforehand. At the same time, it must be noted that the guideline is a best practice statement, which is based on meta-analyses of high-quality randomized controlled trials where possible, but is not always the case. As a part of the recorded consultations took place before the guideline was published, a longer period of time is needed to draw firmer conclusions regarding adherence to the guideline.

Another limitation of this study is that by using electronic medical records our data were completely dependent on the registering behaviour of the GPs, which varied in amount of detail and coding. If the GP did not record certain management strategies or symptoms in a consultation, these data were missing. Our data, therefore, do not necessarily reflect what was actually done during the consultation, rather what was done and recorded. A comparative study with recordings of patients with and without USD, and comprehensive recording of all types of management strategies by GPs or videotaped consultations would be helpful in gaining thorough insight in their management and subsequent recording (28, 29).

A final limitation is that it was not possible to decide which consultation was the first consultation in the course of one or more MUS episodes. Because of this, all consultations were analysed as if they are independent. However, this is usually not the case. The policy of the GP can depend on the findings and results from previous consultations.

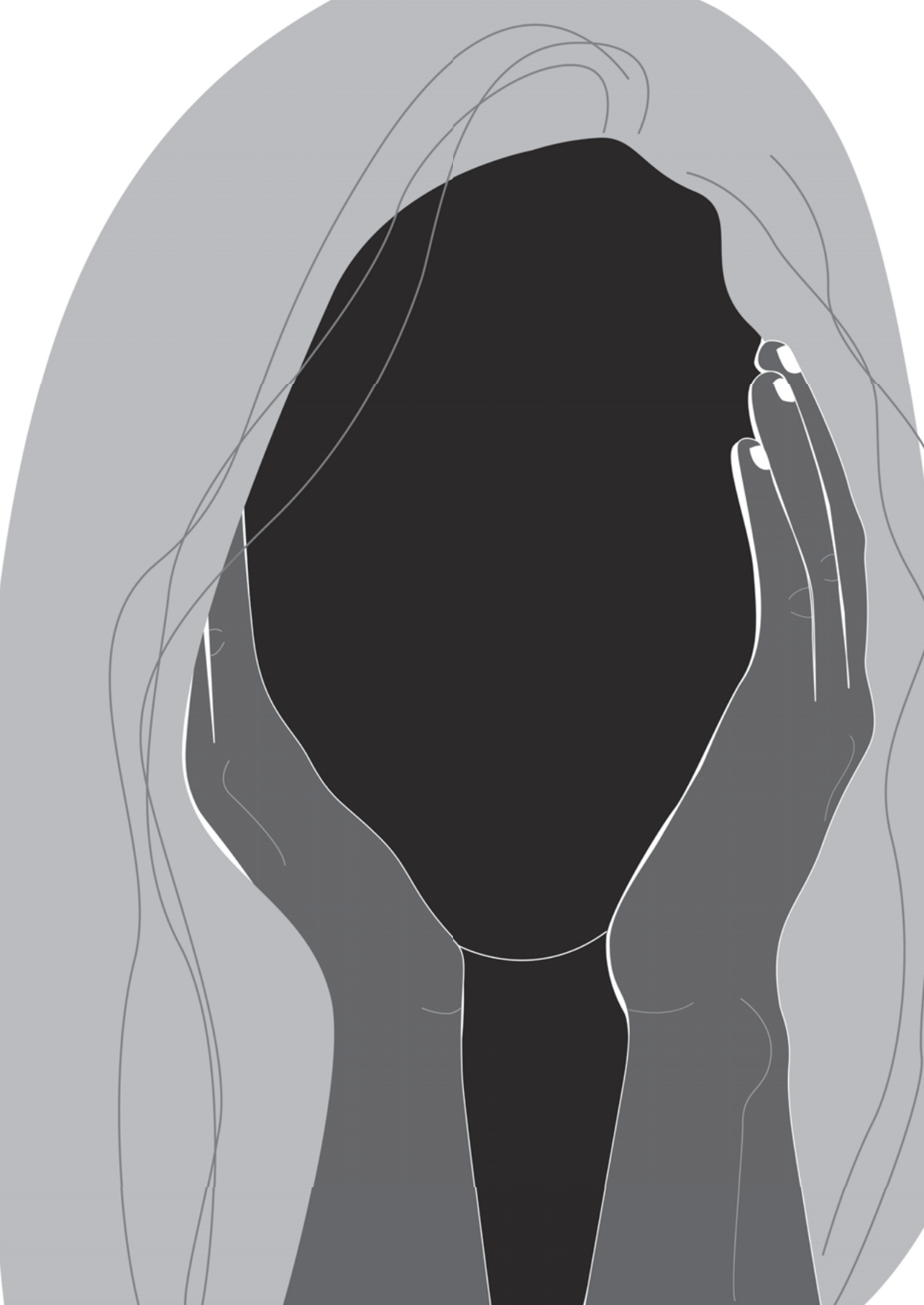
CONCLUSIONS

This is the first study that explores the primary care data of Dutch patients with MUS. GPs use standardised management strategies for persistent MUS, but seem to prescribe medication possibly more frequently and explore symptoms and refer to mental health care less frequently than desirable. Over time they seemed to adopt more monitoring and supportive management strategies for the same patient. When seeing patients with MUS, GPs should consider exploring cognitive, emotional, behavioural and social dimensions of MUS besides the somatic dimension, involving the patient more in the problem definition and treatment plan, referring to a mental health nurse practitioner within the practice, mental health care outside the practice and thorough recording in medical records.

REFERENCES

1. Haller H, Cramer H, Lauche R, Dobos G. Somatoform disorders and medically unexplained symptoms in primary care. Dtsch Arztebl Int 2015;112(16):279-287.
2. van der Weijden T, van Velsen M, Dinant G, van Hasselt C, Grol R. Unexplained complaints in general practice: prevalence, patients' expectations, and professionals' test-ordering behavior. Med Decis Making 2003;23:226-231.
3. Rask MT, Rosendal M, Fenger-Gron M, Bro F, Ornbol E, Fink P. Sick leave and work disability in primary care patients with recent-onset multiple medically unexplained symptoms and persistent somatoform disorders: a 10-year follow-up of the FIP study. Gen Hosp Psychiatry 2015;37(1):53-59.
4. APA: Diagnostic and statistical manual of mental disorders, 4th ed. (DSM-IV). Washington DC: American Psychiatric Association; 1994.
5. APA: Diagnostic and statistical manual of mental disorders, 5th ed. (DSM-5). Arlington, VA: American Psychiatric Publishing; 2013.
6. de Waal MWM, Arnold IA, Eekhof JAH, Van Hemert AM. Somatoform disorders in general practice: prevalence, functional impairment and comorbidity with anxiety and depressive disorders. Br J Psychiatry 2004;184(6):470-476.
7. Verhaak PF, Meijer SA, Visser AP, Wolters G. Persistent presentation of medically unexplained symptoms in general practice. Fam Pract 2006;23(4):414-420.
8. Aamland A, Malterud K, Werner EL. Patients with persistent medically unexplained physical symptoms: a descriptive study from Norwegian general practice. BMC Fam Pract 2014;15:107.
9. van Eck van der Sluijs J, Ten Have M, Rijnders C, van Marwijk H, de Graaf R, van der Feltz-Cornelis C. Medically unexplained and explained physical symptoms in the general population: association with prevalent and incident mental disorders. PLoS One 2015;10(4):e0123274.
10. van der Leeuw G, Gerrits MJ, Terluin B, Numans ME, van der Feltz-Cornelis CM, van der Horst HE, Penninx BW, van Marwijk HW. The association between somatization and disability in primary care patients. J Psychosom Res 2015;79(2):117-122.
11. Konnopka A, Schaefert R, Heinrich S, Kaufmann C, Luppa M, Herzog W, Konig HH. Economics of medically unexplained symptoms: a systematic review of the literature. Psychother Psychosom 2012;81(5):265-275.
12. Sirri L, Grandi S, Tossani E. Medically unexplained symptoms and general practitioners: a comprehensive survey about their attitudes, experiences and management strategies. Fam Pract 2017;34(2):201-205.
13. Johansen ML, Risor MB. What is the problem with medically unexplained symptoms for GPs? A meta-synthesis of qualitative studies. Patient Educ Couns 2017;100(4):647-654.
14. Howman M, Walters K, Rosenthal J, Ajjawi R, Buszewicz M. "You kind of want to fix it don't you?" Exploring general practice trainees' experiences of managing patients with medically unexplained symptoms. BMC Med Educ 2016;16:27.
15. Olde Hartman T, Blankenstein A, Molenaar A, Bentz van den Berg D, van der Horst H, Arnold I, Burgers J, Wiersma T, Woutersen-Koch H. [NHG guideline on medically unexplained symptoms (MUS)] (in Dutch). Huisarts Wet 2013;56(5):222-230.
16. Schaefert R, Hausteiner-Wiehle C, Hauser W, Ronel J, Herrmann M, Henningsen P. Non-specific, functional, and somatoform bodily complaints. Dtsch Arztebl Int 2012;109(47):803-813.
17. Medically unexplained symptoms / functional symptoms: positive practice guide: Improving Access to Psychological Therapies (IAPT), National Health Services (NHS); 2014.
18. olde Hartman TC, Hassink-Franke LJ, Lucassen PL, van Spaendonck KP, van Weel C. Explanation and relations. How do general practitioners deal with patients with persistent medically unexplained symptoms: a focus group study. BMC Fam Pract 2009;10(68).
19. Sitnikova K, Leone SS, Zonneveld LNL,

van Marwijk HWJ, Bosmans JE, van der Wouden JC, van der Horst HE. The CIPRUS study, a nurse-led psychological treatment for patients with undifferentiated somatoform disorder in primary care: study protocol for a randomised controlled trial. Trials 2017;18(1):206.
20. Robbins J, Kirmayer L, Hemami S. Latent variable models of functional somatic distress. J Nerv Ment Dis 1997;185(10):606-615.
21. First M, Spitzer R, Gibbon M, Williams J. Structured Clinical Interview for DSM-IV-TR Axis I Disorders, Research Version, Patient Edition (SCID-I/P). New York: Biometrics Research: New York State Psychiatric Institute; 2002.
22. Terluin B, van Marwijk HW, Ader HJ, de Vet HC, Penninx BW, Hermens ML, van Boeijen CA, van Balkom AJ, van der Klink JJ, Stalman WA. The Four-Dimensional Symptom Questionnaire (4DSQ): a validation study of a multidimensional self-report questionnaire to assess distress, depression, anxiety and somatization. BMC Psychiatry 2006;6:34.
23. den Boeft M, Claassen-van Dessel N, van der Wouden JC. How should we manage adults with persistent unexplained physical symptoms? Br Med J 2017;356:j268.
24. Clark R, Tonmukayakul U, Mangan Y, Smith M, Gussy M, Manton D, Bailey D, Calache H. Measuring adherence to evidence-based clinical practice guidelines. J Evid Based Dent Pract 2017;17(4):301-309.
25. van Peet PG, Drewes YM, Gussekloo J, de Ruijter W. GPs' perspectives on secondary cardiovascular prevention in older age: a focus group study in the Netherlands. Br J Gen Pract 2015;65(640):e739-747.
26. Franx G, Huyser J, Koetsenruijter J, van der Feltz-Cornelis CM, Verhaak PFM, Grol RPTM, Wensing M. Implementing guidelines for depression on antidepressant prescribing in general practice: a quasi-experimental evaluation. BMC Fam Pract 2014; 15(35).
27. Austad B, Hetlevik I, Mjolstad BP, Helvik AS. Applying clinical guidelines in general practice: a qualitative study of potential complications. BMC Fam Pract 2016;17:92.
28. olde Hartman TC, van Rijswijk E, van Dulmen S, van Weel-Baumgarten E, Lucassen PL, van Weel C. How patients and family physicians communicate about persistent medically unexplained symptoms. A qualitative study of video-recorded consultations. Patient Educ Couns 2013;90(3):354-360.
29. Jepson M, Salisbury C, Ridd MJ, Metcalfe C, Garside L, Barnes RK. The 'One in a Million' study: creating a database of UK primary care consultations. Br J Gen Pract 2017;67(658):e345-e351.



Chapter 5

Effectiveness of a cognitive behavioural intervention for patients with undifferentiated somatoform disorder: Results from the CIPRUS cluster randomized controlled trial in primary care

Kate Sitnikova
Stephanie S Leone
Harm WJ van Marwijk
Jos Twisk
Henriëtte E van der Horst
Johannes C van der Wouden

Published in: Journal of Psychosomatic Research 2019;127:109745.

ABSTRACT

Objective: To examine the effectiveness of a cognitive behavioural intervention delivered by mental health nurse practitioners (MHNPs) to patients with undifferentiated somatoform disorder (USD), compared to usual care.

Methods: We conducted a cluster randomized trial among primary care patients with USD comparing the intervention to usual care. The intervention consisted of six sessions with the MHNP. Primary outcome was physical functioning (RAND-36 physical component summary score). Secondary outcomes were the RAND-36 mental component summary score and the eight subscales; anxiety and depression (Hospital Anxiety and Depression Scale) and somatic symptom severity (Patient Health Questionnaire-15). Outcomes were assessed at baseline, 2, 4 and 12 months. We analysed data using linear mixed models by intention-to-treat, and investigated effect modifiers.

Results: Compared to usual care (*n*=87), the intervention group (*n*=111) showed an improvement in physical functioning (mean difference 2.24 [95% CI 0.51; 3.97]; *p*=.011), a decrease in limitations due to physical problems (mean difference 10.82 [95% CI 2.14; 19.49]; *p*.=0.015) and in pain (mean difference 5.08 [95% CI 0.58; 9.57]; *p*=.027), over 12 months. However effect sizes were small and less clinically relevant than expected. We found no differences for anxiety, depression and somatic symptom severity. Effects were larger and clinically relevant for patients with more recent symptoms and fewer physical diseases.

Conclusion: The cognitive behavioural intervention was effective in improving pain and physical functioning components of patients' health. It was particularly suitable for patients with symptoms that had been present for a limited number of years and with few comorbid physical diseases.

Trial registration: The trial is registered in the Dutch Trial Registry, www.trialregister.nl, under NTR4686.

INTRODUCTION

Medically unexplained physical symptoms (MUPS) are a diverse mixture of symptoms for which (currently) a medical explanation is lacking and which are extremely common in primary care (1, 2). In a small percentage of patients with MUPS a specific disorder may eventually prove to be the cause of the symptoms (3). Although most MUPS are self-limiting, symptoms may persist and cluster. In the latter case symptoms may fulfil the diagnostic criteria for the DSM-IV undifferentiated somatoform disorder (USD) (4). If physical symptoms are accompanied by disproportionate emotional, cognitive and behavioural reactions, they may fulfil DSM-5 diagnostic criteria for somatic symptom disorder (5). These psychiatric disorders are associated with a large burden of disease, poor quality of life, functional impairment, depression and anxiety (6, 7). Additionally, general practitioners (GPs) may experience some MUPS patients as 'difficult' or even 'heartsink' (8–10). They face a dilemma between pursuing medical investigations which will probably yield nothing important and may cause harm on the one hand, and refraining from further investigations with a very small chance of overlooking a (treatable) disease on the other hand (11, 12). Patients may feel uncertain, confused and distressed upon hearing that their symptoms do not currently fit with a diagnosable illness, and may regard that message as implying that their symptoms are feigned or "all in the mind" (13, 14). This may negatively impact their attitude towards mental health interventions for their symptoms.

Non-pharmacological interventions such as cognitive behavioural therapy (CBT) may improve functioning of patients with somatoform symptoms and reduce symptom severity (15). However, a German study found that only half of the patients with a somatoform disorder actually receive such mental health treatment (16), as CBT is commonly provided in secondary care or outside general practice. Patients and GPs may feel reluctant to turn to mental health services for physical symptoms. A CBT-based treatment provided in and suitable for primary care may therefore be a solution.

In the Dutch healthcare system, all citizens are registered with a general practice. The GP serves as gatekeeper to other healthcare providers. A recent reform of the Dutch mental healthcare system aimed to reduce the gap between general practice and mental health treatment (17). In 2014, the mental health nurse practitioner (MHNP) was introduced within general practice and currently nearly all surgeries in the Netherlands (87% in 2016) employ one (18). Dutch MHNPs have received higher vocational training in nursing or psychology and work under the supervision of the GP (17). They deliver short-term

interventions to patients with psychosocial problems, but their expertise in psychological techniques such as CBT can vary.

The Dutch guideline on MUPS for GPs recommends that the MHNP offers treatment for MUPS when symptoms are mild to moderate (19). A standardized, evidence-based treatment could be helpful for patients and feasible for MHNPs to deliver. However, such an intervention has never been evaluated in this patient group. The main aim of this study was to examine the effectiveness of a new short-term CBT-based intervention delivered by MHNPs to patients with USD, as compared to usual care.

METHOD

Trial design

We performed a multicentre, cluster randomized controlled trial with two parallel groups comparing a CBT-based intervention on top of usual care, to usual care alone. The study design is described in more detail elsewhere (20).

We chose a cluster design in order to prevent contamination between patients in the same general practice and to prevent MHNPs from having to carry out the intervention with some of the patients and not carry it out with others, which might lead to contamination.

Ethics

The study was conducted according to the declaration of Helsinki (version 2013) and was approved by the VU University Medical Center Ethics Committee (number 2014.305, 9 July 2014 (amendment 5 August 2016).

Participants

Eligibility criteria

Participants were recruited from general practices across the Netherlands, were aged 18 years and above and met the DSM-IV criteria for USD. Exclusion criteria were: having a medical or psychological disorder that explained the symptoms; having a severe psychiatric disorder (e.g. psychotic disorder); currently receiving psychological help for USD; having poor language skills or physical handicaps interfering with understanding the intervention or questionnaires.

Inclusion procedure

GPs selected patients aged 18 years and above from their electronic database, who had consulted the GP with one or more symptoms from “Robbins’ list” (21) at least

twice in the previous 3 months. Robbins' list consists of 23 physical symptoms that are associated with functional somatic syndromes and can be found in the trial protocol (20). GPs checked the list of selected patients for exclusion criteria. Patients identified as potentially eligible received concise information about the study and the Patient Health Questionnaire 15-item somatic symptom severity scale (PHQ-15) (22) by mail from their GP. Patients with a PHQ-15 score of at least 5 (low symptom severity) and who were interested in participation in the study received extensive information. Patients were then invited to participate in a clinical interview (Structure Clinical Interview for DSM-IV Axis I Disorders (SCID-I) (23)) to assess DSM-IV criteria for USD. Trained members of the research team administered these interviews by telephone. Patients meeting the criteria for USD received an informed consent form.

Intervention

The intervention consisted of six individual sessions of 30 min each with a MHNP. The intervention comprised a combination of two CBT-based techniques: a modified version of the 'consequences model' for somatoform disorders and Problem-Solving Treatment (PST). The consequences model is frequently used in treatment of somatoform disorders in Dutch secondary care (24, 25). It focusses on the consequences or problems that arise due to somatoform complaints rather than on their causes, which are by definition unknown. The model assumes that physical symptoms may lead to various consequences in the patient's daily life, which are in fact survival strategies in reaction to the physical symptoms. Although these survival strategies must have been beneficial initially (or they would not have been developed), they can aggravate symptoms and become harmful or devastating in the end. In the model, patients can improve their (physical) functioning and quality of life by strengthening the more beneficial, instead of the harmful, survival strategies in the end. Identified consequences or problems are then tackled using a CBT-based technique, which will be learned and applied by patients following the steps outlined in problem solving therapy (PST). PST is a practical treatment that is suitable for delivery by primary healthcare providers such as GPs and nurses (26, 27). The goal was to support patients in developing survival strategies that are more helpful in the long run.

Each session was described in detail in the intervention manual that all MHNPs received during their training. In session 1 the MHNP introduced and explained the treatment, the patient told the MHNP about their physical symptoms and consequences/problems in his/her life due to these symptoms. A consequence/problem was defined as something the patient would like to achieve but is unable to at the moment. In session 2 the MHNP explained the PST goals and steps. In each of the sessions 3 through 6, the patient

addressed a single consequence/problem, using the PST steps together with the MHNP. The goal was to stimulate people to practise improving their long-term problem solving skills, and to apply them to the consequences of their physical symptoms, but also to other problems in daily life. Patients applied the steps at home following written instructions. If not all the steps were covered during one session, they were addressed during the next session.

MHNPs

MHNPs followed two group training sessions lasting 3–3.5 h each. The training sessions were led by a clinical psychologist specialized in management of somatoform disorders. The training consisted of a theoretical part on USD, the consequences model and treatment rationale, and a practical part, in which MHNPs practiced PST. The clinical psychologist supervised the MHNPs during the study period.

Usual care

The usual care group did not receive any additional care, other than the usual care they received from their GP and any other healthcare providers they were referred to.

Outcome measures

Primary and secondary outcomes were assessed at baseline and 2, 4 and 12 months later. The measurements at 2 and 4 months corresponded to the intervention group being halfway through the intervention (3 sessions completed) and completing the intervention (6 sessions completed), respectively. Potential mediating variables were assessed at baseline and 2 and 4 months later. All outcome measures were assessed at the individual patient level.

Primary outcome

The primary outcome was the improvement in physical functioning during the total follow-up period, as measured by the physical component summary score (PCS) of the RAND-36 questionnaire. The PCS is one of the aggregated scores of the RAND-36, a validated questionnaire measuring health related quality of life. The raw scores were transformed into scores ranging from 0 to 100 with a higher score indicating better physical functioning (28). Item examples are “Does your health now limit you in lifting or carrying groceries? If so, how much?” and “How much bodily pain have you had during the past 4 weeks?”

Secondary outcomes

Secondary outcomes were the mental component summary score (MCS) and the eight subscales of the RAND-36 (physical functioning, role limitations due to physical health problems, role limitations due to emotional problems, social functioning, emotional well-being, energy/fatigue, bodily pain and general health perceptions). The scores of each separate subscale range from 0 to 100, higher scores indicating better health.

Depression and anxiety symptoms were measured by the Hospital Anxiety and Depression Scale (HADS) (29). This is an instrument with anxiety and depression subscales (HADS-A and HADS-D, respectively) with scores ranging from 0 to 21 for each scale. Higher scores indicate more severe symptoms. Somatization was measured with the PHQ-15 (22). This instrument has a score range of 0–30, with higher scores indicating higher somatic symptom severity.

Mediators

We took the following potential mediators into account: problem-solving skills (Social Problem-Solving Inventory) (30), health anxiety (Whitely Index) (31), cognitive and emotional representations of illness (brief version of the Illness Perception Questionnaire) (32), cognitive and behavioural responses to illness (Cognitive and Behavioural Responses Questionnaire) (33, 34), and level of perceived control (Pearlin Master Scale) (35). Data on these variables were collected at baseline, and at 2 and 4 months follow-up.

In the intervention group, we assessed the strength of the therapeutic alliance with the revised short-form Working Alliance Inventory (WAI-SR)) (36), at 2 months and at 4 months.

Process evaluation of the intervention

In order to better interpret our quantitative findings, we conducted a process evaluation by interviewing MHNPs from the intervention group. When they had completed most sessions, all 15 MHNPs were invited to participate in face-to-face interviews to evaluate their involvement in the trial; 13 accepted the invitation. The semi-structured interviews were based on a topic list with pre-identified themes.

At 4 months after baseline, patients in the intervention group were asked to answer 13 Likert items, evaluating their participation in the trial. The collected data were systematically analysed with both qualitative (interviews) and quantitative (questionnaire) methods.

Sample size

We aimed to detect a clinically relevant effect size of 0.4 sd on our primary outcome. We chose a two-sided significance level of 5% and a power of 80%. The allocation ratio was 1:1. We assumed a correlation coefficient of 0.5 for repeated measurements. Using linear mixed models with these values required a sample size of 74 patients per condition. We corrected for the cluster design with an expected average cluster size of 4 and assuming an ICC of 0.05 (37). Taking a potential dropout rate of 20% into account, we aimed to include 106 patients in each condition.

Randomization and blinding

At randomization, a cluster consisted of all participating general practices that employed one MHNP. An independent epidemiologist carried out concealed random allocation and assignment of clusters to the intervention group or control group by using a computer generated randomization list. She was not involved in the selection of general practices. In order to balance the size of the intervention and control groups, randomization was stratified according to cluster size (small: <5000 patients, and large: ≥5000 patients). Due to the nature of the intervention, it was not possible to blind researchers, GPs or patients to the allocation. MHNPs and GPs were informed about their allocation after signing a form that they agreed to participate. Patients were informed about their treatment allocation after they signed and returned the informed consent form.

Statistical analyses

We used descriptive statistics for baseline characteristics. The effect of the intervention on primary and secondary outcomes was analysed according to the intention-to-treat (ITT) principle (38). Linear mixed models analyses were used to take into account the dependence of repeated measurements in individual patients, without imputing missing data (39). Respondents were included if they had completed at least one follow-up measurement. For each outcome variable, we estimated the overall effect over time, and the effect per time point (2, 4 and 12 months after baseline). Time and the interaction between study group (intervention or control) and time were added to the models.

For each outcome measure, we performed a crude and adjusted analysis over the total follow-up period of 12 months. The crude analysis was only adjusted for the baseline value of the particular outcome. In the adjusted analysis, we evaluated whether the following variables were actual confounders: gender, age, level of education, duration of symptoms, somatic symptom severity (PHQ-15), anxiety symptoms (HADS-A), depressive symptoms (HADS-D), number of comorbid diseases, time intervals between completing baseline questionnaire and 2-month follow-up, and between completing the

2-month and 4-month follow-up questionnaires. We adjusted for the latter two because these intervals were different between the intervention group and the usual care group, for logistic reasons. Variables found to be actual confounders were added to the adjusted model.

For secondary outcomes, *p*-values should be interpreted cautiously due to multiple statistical comparisons, unless highly significant (e.g. $p < .01$).

To evaluate whether we should adjust for clustering within general practice, this variable was added as an additional level to the linear mixed model analysis. As this did not improve the model (likelihood ratio test: p=-0.90; ICC<0.01), clusters were not included in the final analyses.

All analyses were repeated applying the per protocol principle to the intervention group, as exploratory analyses. We defined three different per protocol populations: intervention patients who had 1) attended all 6 sessions with their MHNP (n=57); 2) attended all 6 sessions or less if their goals were achieved earlier (n=62); 3) attended at least 4 sessions (n=76). All control group patients were included in the per protocol analysis (n=87).

We carried out additional analyses by adding interaction terms to evaluate whether any of the pre-determined variables (age, gender, education level, symptom duration, somatic symptom severity (PHQ-15), physical comorbidity and anxiety and depressive symptoms (HADS)) were effect modifiers. These variables were chosen based on the Dutch multidisciplinary guideline for the management of MUPS and somatoform disorders, which identifies them as relevant factors in discerning patient profiles (40). Mediation analyses were carried out based on Krull & MacKinnon (41), using the Sobel-Goodman test to determine significance. The relationship between primary outcome and working alliance was assessed using Pearson's r.

Cohen's d for measuring effect size was calculated by dividing the regression coefficient of each outcome by its standard deviation. We used standard deviations for the total group at baseline.

Data were analysed using IBM SPSS Statistics version 22 and Stata version 14.

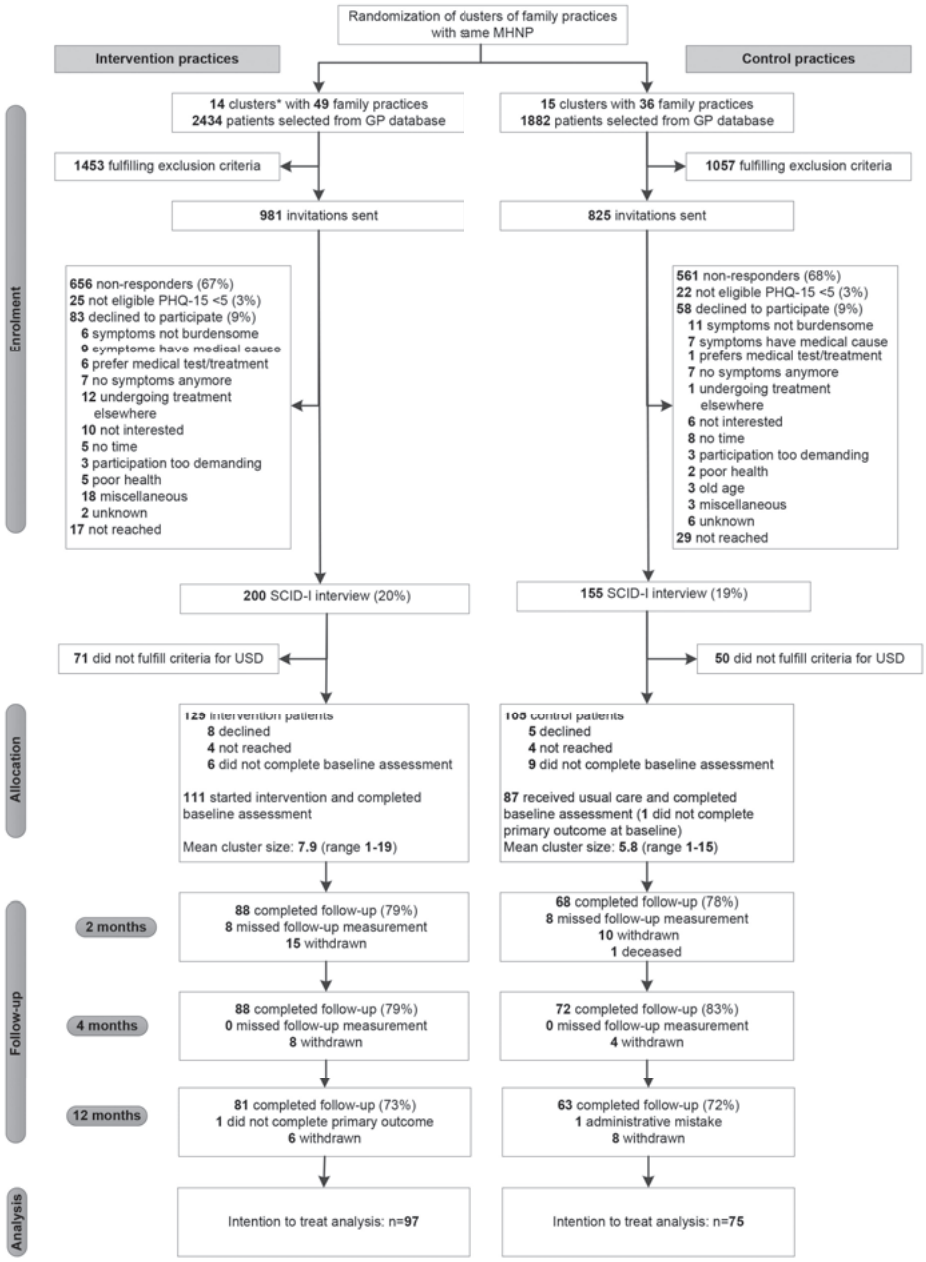


Figure 1. Flow of study participants

* For the intervention group, a cluster was composed by matching MHNPs to all participating general practices where the MHNP works and to all other MHNPs who also worked in these general practices.

GP General practitioner; *MHNP* mental health nurse practitioner; *PHQ-15* Patient Health Questionnaire 15-item somatic symptom severity scale; *SCID-I* Structured Clinical Interview for DSM-IV Axis I Disorders

RESULTS

Recruitment

Recruitment of patients took place between August 2015 and March 2017. Recruitment stopped when a total of 213 informed consent forms had been returned. Figure 1 provides an overview of the enrolment procedure.

Baseline characteristics

Socio-demographic and clinical baseline characteristics of the participants are provided in Table 1. The mean age of the total sample was 51.5 years (sd=16.3) and the majority of patients (74.5%) were female. There were more female patients (79.8%) in the control group than in the intervention group (70.3%) and the level of completed education was lower. The median duration of symptoms in the total sample was 5.7 years (IQR=2.7–15.7) and most commonly reported were musculoskeletal complaints (72.0%). Neurological symptoms were more common in the intervention group (35.1%) than in the control group (18.0%).

Table 1. Patients' baseline characteristics

Characteristics	Intervention group (n=111)	Control group (n=89) [a]	Total sample (n=200) [b]
Age, mean (sd)	53.00 (15.47)	49.69 (17.13)	51.53 (16.3)
Female	78 (70.3%)	71 (79.8%)	149 (74.5%)
Both parents born in the Netherlands	90 (81.1%)	72 (80.9%)	162 (81.0%)
Educational level			
Low	8 (7.3%)	7 (8.0%)	15 (7.6%)
Medium	58 (52.3%)	54 (62.1%)	112 (56.9%)
High education	44 (40.0%)	26 (29.8%)	70 (35.5%)
Work status [c]			
Employed	43 (38.7%)	36 (40.4.%)	79 (39.5%)
Unemployed	68 (61.3%)	53 (59.6%)	121 (60.5%)
Living situation			
Alone	28 (25.2%)	23 (25.8%)	51 (25.5%)
Not alone	83 (74.8%)	66 (74.2%)	149 (74.5%)
Symptom duration in years (self-report), median (IQR)	5.2 (2.8-15.5)	6.1 (2.7-16.1)	5.7 (2.7-15.7)
Most prominent symptoms [c]			
Musculoskeletal	78 (70.3%)	66 (74.2%)	144 (72.0%)

Table 1. Continued

Characteristics	Intervention group (n=111)	Control group (n=89) [a]	Total sample (n=200) [b]
General and unspecified	41 (36.9%)	36 (40.4%)	77 (38.5%)
Neurological	39 (35.1%)	16 (18.0%)	55 (27.5%)
Psychological	19 (17.1%)	18 (20.2%)	37 (18.5%)
Digestive	11 (9.9%)	7 (7.9%)	18 (9.0%)
Number of comorbid physical diseases, mean (sd)	3.16 (2.50)	3.34 (2.41)	3.24 (2.46)
Most reported comorbid physical diseases [c]			
Back problems	79 (71.2%)	63 (70.8%)	142 (71.0%)
Pulmonary	40 (36.0%)	28 (31.5%)	68 (34.0%)
Neurological	35 (31.5%)	31 (34.8%)	66 (33.0%)
Number of self-report comorbid psychiatric disorders, mean (sd)	0.69 (0.91)	0.71 (1.19)	0.70 (1.04)
Most reported comorbid psychiatric disorders [c]			
Distress/burn-out	27 (24.5%)	18 (20.9%)	45 (23.0%)
Depression	26 (23.4%)	17 (19.5%)	43 (21.7%)
Anxiety	19 (17.1%)	16 (17.4%)	34 (17.3%)
RAND-36			
PCS (primary outcome)	50.22 (9.89)	49.64 (9.81)	49.97 (9.83)
MCS	49.80 (9.97)	50.22 (10.93)	49.99 (10.38)
Physical functioning	62.78 (25.08)	59.25 (26.12)	61.23 (25.54)
Role functioning/physical	21.62 (32.59)	25.29 (32.84)	23.21 (32.67)
Role functioning/emotional	56.52 (44.13)	52.61 (46.01)	54.84 (44.87)
Social functioning	55.63 (27.73)	57.76 (27.84)	56.57 (27.73)
Bodily pain	46.24 (21.31)	45.79 (22.19)	46.04 (21.64)
Emotional well-being	59.72 (17.04)	60.43 (20.23)	60.03 (18.43)
Energy/fatigue	36.89 (16.90)	38.65 (15.95)	37.65 (16.48)
General health	43.81 (17.64)	41.43 (16.25)	42.78 (17.05)
Anxiety (HADS-A)	7.89 (3.80)	7.69 (4.52)	7.80 (4.11)
Depression (HADS-D)	7.00 (3.86)	7.52 (4.23)	7.22 (4.02)
Somatic symptom severity (PHQ-15)	13.63 (4.89)	13.47 (4.43)	13.56 (4.69)

Results are expressed as *n* (%) unless stated otherwise, and in *mean (sd)* for the RAND-36, HADS and PHQ-15.
Abbreviations: HADS: Hospital Anxiety and Depression Scale; MCS: Mental Component Summary Score; MUS: medically unexplained symptoms; PCS: Physical Component Summary Score; PHQ-15: Patient Health Questionnaire 15-item somatic symptom severity scale; sd: standard deviation
[a] 89 patients completed items on demographic characteristics, but 87 patients completed the primary outcome. Therefore, due to missing values, the available *n* ranged from 87-89.
[b] Due to missing values the available *n* ranged from 188-200.
[c] More than one answer option was permitted, so numbers do not necessarily add up to 100%

Numbers analysed

Data on the primary outcome on at least one follow-up assessment were available for 97/111 (87.4%) patients in the intervention group and 75/87 (86.2%) patients in the control group. Fig. 1 provides more details on withdrawals.

Primary outcome

Fig. 2 visually represents the course of the RAND-36 PCS for both groups. The ITT analysis showed a significant intervention effect over the 12 month period on patients' physical functioning (PCS score difference 2.24 [95% CI 0.51; 3.97]; *p*=.011; Cohen's d=0.23) (Table 2). Results per time point are presented in Appendix A. There was a statistically significant difference at 4 months after baseline (PCS score difference 2.93 [95% CI 0.77; 5.09]; *p*=.008; Cohen's d=0.30).

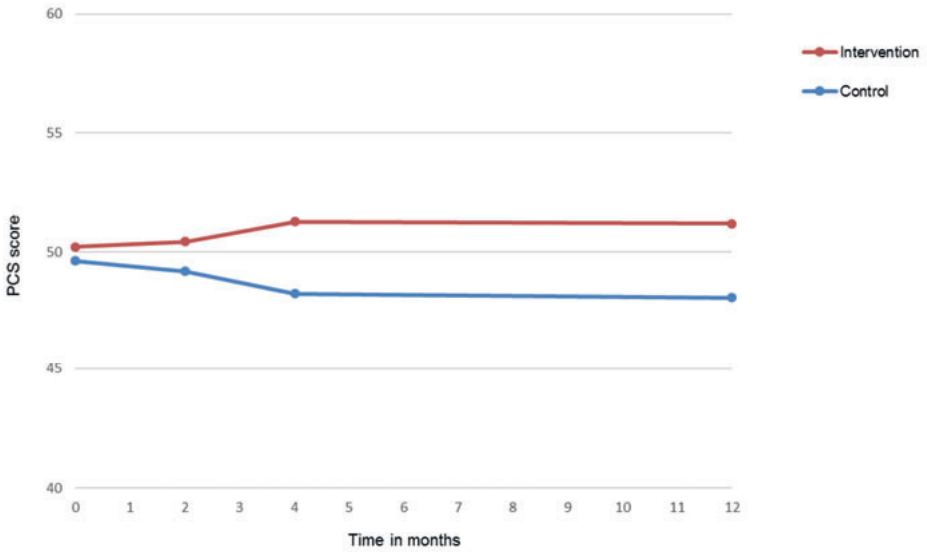


Figure 2. 12-month course of physical functioning as measured with the RAND-36 Physical Component Summary score (PCS)

Secondary outcomes

Significant intervention effects over 12 months were also found on patients' limitations in functioning due to physical health problems (RAND-36 role functioning/physical score difference 10.82 [95% CI 2.14; 19.49]; p.=0.015; Cohen's d=0.33) and bodily pain (RAND-36 bodily pain score difference 5.08 [95% CI 0.58; 9.57]; *p*=.027; Cohen's d=0.23) (Table 2).

Table 2. Results of the mixed models Intention-To-Treat analyses

	Crude analyses		Adjusted analyses [a]	
	B (95% CI)	**p-value**	**B (95% CI)**	**p-value**
Primary outcome RAND-36 Physical component summary score (PCS)	1.80 (0.19 to 3.42)	0.029*	2.24 (0.51 to 3.97)	0.011*
Secondary outcomes RAND-36				
Mental component summary score (MCS)	-0.55 (-2.47 to 1.37)	0.57	-0.35 (-2.22 to 1.52)	0.71
Physical functioning	1.47 (-2.08 to 5.02)	0.42	2.33 (-1.40 to 6.06)	0.21
Role functioning/ physical	7.17 (-1.16 to 15.50)	0.091	10.82 (2.14 to 19.49)	0.015*
Role functioning/ emotional	-2.14 (-11.63 to 7.36)	0.66	1.41 (-8.29 to 11.10)	0.78
Social functioning	2.65 (-2.85 to 8.14)	0.35	2.66 (-3.09 to 8.41)	0.37
Bodily pain	3.98 (-0.31 to 8.27)	0.069	5.08 (0.58 to 9.57)	0.027*
Emotional well-being	-0.77 (-3.88 to 2.35)	0.63	-0.13 (-3.28 to 3.02)	0.93
Energy/fatigue	2.56 (-0.64 to 5.75)	0.12	1.98 (-1.24 to 5.20)	0.23
General health	-0.28 (-3.90 to 3.34)	0.88	0.05 (-3.91 to 4.02)	0.98
Anxiety symptoms (HADS-A)	0.30 (-0.33 to 0.94)	0.35	0.33 (-0.29 to 0.94)	0.30
Depressive symptoms (HADS-D)	-0.06 (-0.66 to 0.55)	0.86	-0.23 (-0.89 to 0.43)	0.49
Somatic symptom severity (PHQ-15)	-0.51 (-1.43 to 0.40)	0.27	-0.69 (-1.64 to 0.24)	0.15

Abbreviations: 95%CI: 95% Confidence Interval; HADS-A: Hospital Anxiety and Depression Scale-Anxiety subscale; HADS-D: Hospital Anxiety and Depression Scale-Depression subscale; MCS: Mental Component Summary Score; PCS: Physical Component Summary Score; PHQ-15: Patient Health Questionnaire 15-item somatic symptom severity scale
a: adjusted for (if necessary): gender, age, level of education, duration of symptoms, somatic symptom severity (PHQ-15), anxiety symptoms (HADS-A), depressive symptoms (HADS-D), number of comorbid physical diseases, time interval baseline – 2-months follow-up, time interval baseline – 4-months follow-up
* p<0.05

When investigating the effects per time point (Appendix A), the largest and statistically significant differences were found at 4 months after baseline. Although there was no overall effect on the RAND-36 physical functioning subscale, there was a significant difference at 4 months (5.00 points [95% CI 0.28; 9.73]; *p*=.038; Cohen's d=0.20).
No significant intervention effects were found for the remaining domains of health related quality of life, anxiety, depression and somatic symptom severity.

Effect modification

Reported duration of symptoms at baseline significantly modified the effect of our intervention on the RAND-36 PCS (p=0.011), bodily pain (p=0.048) and general health subscales (p=0.006). Physical comorbidity significantly modified the intervention effect on the RAND-36 PCS (p=0.026) and general health subscale (p=0.031). No other variable modified any of the effects. In order to report the results separately, we split each effect-modifying variable on its median. Table 3 summarizes the results per group and Appendix B provides results per time point.

Generally, patients with a shorter duration of symptoms and fewer comorbid physical diseases showed improvement, as opposed to those with a longer duration of symptoms, who reported poorer general health after the intervention.

Table 3. Results of crude mixed models analyses per group for symptom duration and physical comorbidity

	Symptom duration below median (n=70-98)		**Symptom duration above median (n=72-98)**	
	Overall difference B (95% CI)	**p-value**	**Overall difference B (95% CI)**	**p-value**
Primary outcome				
RAND-36 PCS	3.83 (1.57 to 6.09)	0.001*	-0.18 (-2.40 to 2.03)	0.87
Secondary outcomes				
RAND-36				
Bodily pain	6.94 (1.05 to 12.84)	0.021*	0.68 (-5.41 to 6.78)	0.83
General health	5.74 (0.96 to 10.52)	0.019*	-5.98 (-11.09 to -0.86)	0.022*

	0-2 comorbid physical diseases (n=69-91)		**3 or more comorbid physical diseases (n=75-106)**	
	Overall difference B (95% CI)	**p-value**	**Overall difference B (95% CI)**	**p-value**
Primary outcome				
RAND-36 PCS	3.55 (1.13 to 5.97)	0.004*	-0.02 (-2.09 to 2.05)	0.99
Secondary outcome				
RAND-36 General health	2.38 (-2.60 to 7.36)	0.35	-3.09 (-8.16 to 1.98)	0.23

Abbreviations: 95%CI: 95% Confidence Interval; PCS: Physical Component Summary Score
* p<0.05

Mediation

None of the potential mediators actually mediated the effect on the primary outcome. With regard to the working alliance between patient and MHNP in the intervention

group, a weak, but significant, positive correlation was found between the RAND-36 PCS change score between baseline and 4 months and the WAI-SR bond scale (r=0.258, n=78, p=0.022) at 4 months after baseline.

Exploratory analyses

The results of the per protocol analyses are provided in Appendices C, D and E. For nearly all outcome variables the effect was similar to those in the ITT analyses.

Evaluation by MHNPs and patients

MHNPs were satisfied with the amount and content of training they received before delivering the intervention. Most found that 30 min was not enough for a single session. They reported that they generally adhered to the protocol but sometimes adjusted the length and pace of sessions by taking more time. MHNPs considered the CBT-based intervention to be a suitable technique for treating USD, that enhanced patients' problem-solving abilities and activated them in their daily life. MHNPs felt that most patients benefited from the intervention, as their functioning became less impaired, but thought that the intervention might not be effective for patients with comorbid physical and psychological disorders, psychosocial problems or a lower IQ.

Most MHNPs would use (elements of) the protocol again in the future. Those who would use the protocol again said the treatment manual improved their proficiency in a CBT-based intervention, provided them with structure during sessions and a more problem-solving mindset. For future use, MHNPs recommended personalizing the number and pace of sessions to the patient, and offering other treatment methods alongside the CBT-based method, such as Acceptance and Commitment Therapy, psychoeducation, and physical activation.

86 patients (77%) completed the patient evaluation questionnaire at 4 months after baseline. The (selected) results are provided in Table 4. The majority of patients (66%) rated the quality of the intervention as good, 11% as excellent, 14% as mediocre and 1 person (1%) as very poor. Half of the patients (51%) reported that the intervention helped them deal with their physical symptoms, 22% were neutral and 17% said it did not help. Most patients were fairly (42%), or extremely satisfied (25%) with the intervention and 20% were neutral. Only a few were (somewhat) unsatisfied (4%). More than half (54%) said they would certainly or probably recommend the intervention to a friend or family member with USD.[1]

[1] Figures in this paragraph do not add to 100% due to missing values.

Table 4. (Selected) results from the patient evaluation questionnaire*

1. Did the individual training help you deal better with your physical complaints?	*Yes, it helped me a lot*	*Yes, it helped me somewhat*	*Neutral*	*No, it did not help me*	*No, it aggravated my complaints*
	4.6%	46.0%	21.8%	17.2%	0%
2. What did you think of the quality of the individual training that you followed?	*Excellent*	*Good*	*Medium*	*Bad*	*Very bad*
	10.3%	65.5%	13.8%	0%	1.1%
3. Did the intervention meet your expectations?	*All my expectations were met*	*Most of my expectations were met*	*Some of my expectations were met*	*Only a few of my expectations were met*	*None of my expectations were met*
	8.0%	35.6%	21.8%	18.4%	5.7%
4. How satisfied are you in general with the intervention you received?	*Very satisfied*	*Somewhat satisfied*	Neutral	Somewhat unsatisfied	Very unsatisfied
	25.3%	41.4%	19.5%	3.4%	1.1%
5. Imagine that someone you know happens to have unexplained physical complaints, would you recommend this intervention?	*Yes, definitely*	*Yes, I think so*	*Maybe*	*No, I don't think so*	*No, definitely not*
	23.0%	31.0%	26.4%	11.5%	0%
6. Imagine that you encounter unexplained physical symptoms again in the future, would you follow this intervention again with your MHNP?	*Yes, definitely*	*Yes, I think so*	*Maybe*	*No, I don't think so*	*No, definitely not*
	17.2%	14.9%	28.7%	26.4%	3.4%

* Numbers do not add up to 100% due to missing values

DISCUSSION

Summary of findings

Our intervention improved physical components of patients' health. Physical functioning improved, bodily pain and limitations due to physical problems and pain decreased. This effect was more pronounced for patients with physical symptoms that had been present for a limited number of years and with few comorbid physical diseases. These patients also experienced improved general health.

Our study demonstrates that a relatively short and light intervention such as ours in primary care is suitable for patients with less persistent symptoms, but insufficient for those with more persistent symptoms. Our findings are supported by information from the interviews with the MHNPs, most of whom clearly distinguished between more severe and less severe patients, and reported that the latter benefited more from the intervention. Patients with symptoms that had lasted longer than the median duration deteriorated somewhat in their general health perceptions after our intervention. These patients assessed their health as poor and expected it to deteriorate further in the future. A possible explanation is that these patients, whose symptoms are already more difficult to treat due to their duration, became demoralized after receiving (possibly yet another) treatment that did not seem to help. In general, patients in the intervention group were satisfied with the intervention.

Although most patients reported that they were satisfied with the intervention and would recommend it to a friend with MUPS, a smaller percentage was less positive. Presumably these were the patients with more persistent symptoms who did not benefit from the intervention.

Surprisingly, none of the variables that we hypothesized to be potential mediators actually mediated the effect on patients' well-being and symptoms. Thus, our study was unable to shed light on the mechanism of change. We did find a positive, though low, correlation with therapeutic alliance, which corresponds to previous findings of a positive therapeutic relationship being partly responsible for the effects of a psychological intervention and improving quality of life (42, 43).

Embedding in existing literature

Overall we found statistically significant effects of our intervention, but effect sizes were small (d=0.22 for RAND-36 PCS, d=0.33 for role functioning/physical, and d=0.23 for bodily pain), and lower than we aimed for (0.4 sd for the RAND-36 PCS). Also, the

effect on the primary outcome was not clinically relevant (difference of 2.24 whereas a difference of 3–5 is considered clinically relevant (44)). However, effect sizes for the primary outcome were substantially higher in patients with a duration of symptoms shorter than the median (0.39) and with < 3 comorbid physical diseases (0.36). These are considered small, clinically relevant effect sizes. Our overall results are in line with previous findings from RCTs that investigated psychological interventions for patients with somatic complaints (15), where small to medium effect sizes are usually found on functional disability and quality of life. The effect sizes in our trial are also of similar magnitude to those demonstrated for interventions administered in primary care for other common mental disorders such as depression and anxiety (45, 46). For patients who do not respond to brief primary care based interventions, more intense interventions could be offered (47, 48).

In previous research the effectiveness of psychological interventions for patients with multiple MUPS was investigated when provided by various healthcare providers, such as psychotherapists and GPs (49). We investigated the effectiveness of an intervention carried out by MHNPs, a new role in Dutch primary care. Interventions by nurse practitioners seem to have beneficial effects on patient satisfaction and quality of life in primary care patients with somatic problems (50). On the flip side, a recent study in Dutch general practices found that having a MHNP in the surgery resulted in MHNPs offering additional long consultations to patients with mental health problems, but did not reduce

visits to the GP (17). Interventions delivered by MHNPs in general and for patients with somatoform complaints in particular must, therefore, be studied more extensively. Furthermore, incorporating other treatment methods such as physical exercise (49) or relaxation and mindfulness techniques (51) could be helpful for this patient group.

Strengths and limitations

This is the first study that examined the effects of an individual, CBT-based intervention by MHNPs for patients with USD versus usual care. We conducted the study in the actual setting of the general practice, making this treatment easier to implement. Another strength of this study is that we used qualitative data from process evaluation interviews with the MHNPs to deepen understanding of our results.

Although the desired number of patients signed an informed consent form (*n*=213), not all of them completed all of the measurements. The dropout rate also turned out higher than expected (27% rather than 20%). This might be attributed to our case-finding

(persons might be less motivated to change), a relatively long follow-up period (12 months) and the length of the questionnaires.

Furthermore, we used the diagnosis of USD according to the, now outdated, DSM-IV, as our trial was initiated in the transition period from DSM-IV to DSM-5, and a diagnostic interview for the DSM-5 was not available yet. The entire DSM-IV category 'somatoform disorders', to which USD belonged (4), has been replaced by 'somatic symptom disorder' (SSD) in the DSM-5 (5). A study comparing these diagnostic criteria found that patients with SSD always fulfil the criteria for USD, and have more severe symptoms and a lower quality of life (52). Therefore, as all SSD patients fulfil the criteria for USD, our findings could be generalizable to patients with SSD. However, considering that the latter are a more severe group and as our findings show that patients with a longer duration of symptoms do not benefit from our intervention, this needs to be verified in future studies.

We opted for cluster randomization, in order to keep the effect of the intervention as pure as possible, so that trained MHNPs would not have to switch between providing and not providing the intervention to similar patients. However, the choice to use cluster randomization also implied using larger clusters and a more complex definition of cluster because individual various MHNPs working in the same surgery also worked part-time in separate other surgeries.

A final point of consideration is that our trial was conducted in the Dutch healthcare setting, in which every citizen has access to general practice and virtually every general practice has an employed MHNP. Our results may be less generalizable to countries with different healthcare systems.

CONCLUSION

Our study demonstrated promising results for a nurse-led CBT-based intervention for patients with USD over usual primary care. The short-term and relatively light intervention appears effective for patients with a shorter symptom duration and with few other somatic diseases.

REFERENCES

1. Steinbrecher N, Koerber S, Frieser D, Hiller W. The prevalence of medically unexplained symptoms in primary care. Psychosomatics. 2011;52:263-71.
2. Haller H, Cramer H, Lauche R, Dobos G. Somatoform disorders and medically unexplained symptoms in primary care. Dtsch Arztebl Int. 2015;112(16):279-87.
3. Eikelboom EM, Tak LM, Roest AM, Rosmalen JGM. A systematic review and meta-analysis of the percentage of revised diagnoses in functional somatic symptoms. J Psychosom Res. 2016;88:60-7.
4. APA. Diagnostic and statistical manual of mental disorders, 4th ed. (DSM-IV). Washington DC: American Psychiatric Association; 1994.
5. APA. Diagnostic and statistical manual of mental disorders, 5th ed. (DSM-5). Arlington, VA: American Psychiatric Publishing; 2013.
6. de Waal MWM, Arnold IA, Eekhof JAH, Van Hemert AM. Somatoform disorders in general practice: prevalence, functional impairment and comorbidity with anxiety and depressive disorders. Br J Psychiatry. 2004;184(6):470-6.
7. Dijkstra-Kersten SM, Sitnikova K, van Marwijk HW, Gerrits MM, van der Wouden JC, Penninx BW, et al. Somatisation as a risk factor for incident depression and anxiety. J Psychosom Res. 2015;79(6):614-9.
8. Butler CC, Evans M. The 'heartsink' patient revisited. British Journal of General Practice. 1999;49(440):230-3.
9. Rosendal M, Fink P, Bro F, Olesen F. Somatization, heartsink patients, or functional somatic symptoms? Towards a clinical useful classification in primary health care. Scandinavian Journal of Primary Health Care. 2009;23(1):3-10.
10. Hahn SR. Physical symptoms and physician-experienced difficulty in the physician–patient relationship. Annals of Internal Medicine. 2001;134(9 pt 2):897-904.
11. Sirri L, Grandi S, Tossani E. Medically unexplained symptoms and general practitioners: a comprehensive survey about their attitudes, experiences and management strategies. Fam Pract. 2017;34(2):201-5.
12. Johansen ML, Risor MB. What is the problem with medically unexplained symptoms for GPs? A meta-synthesis of qualitative studies. Patient Educ Couns. 2017;100(4):647-54.
13. Nettleton S. 'I just want permission to be ill': towards a sociology of medically unexplained symptoms. Soc Sci Med. 2006;62(5):1167-78.
14. Brown RJ. Introduction to the special issue on medically unexplained symptoms: background and future directions. Clin Psychol Rev. 2007;27(7):769-80.
15. van Dessel N, den Boeft M, van der Wouden J, Kleinstäuber M, Leone S, Terluin B, et al. Non-pharmacological interventions for somatoform disorders and medically unexplained physical symptoms (MUPS) in adults. Cochrane Database Syst Rev. 2014;11, CD01114.
16. Mack S, Jacobi F, Gerschler A, Strehle J, Hofler M, Busch MA, et al. Self-reported utilization of mental health services in the adult German population--evidence for unmet needs? Results of the DEGS1-Mental Health Module (DEGS1-MH). Int J Methods Psychiatr Res. 2014;23(3):289-303.
17. Magnee T, de Beurs DP, de Bakker DH, Verhaak PF. Consultations in general practices with and without mental health nurses: an observational study from 2010 to 2014. BMJ Open. 2016;6(7):e011579.
18. Verhaak PFM, Nielen M, de Beurs D. Factsheet 1: [Demand for care - aid regarding psychological and social problems in the general practice 2011-2016] (in Dutch). Utrecht: Netherlands Institute of Health Services Research (NIVEL); 2017.
19. Olde Hartman T, Blankenstein A, Molenaar A, Bentz van den Berg D, van der Horst H, Arnold I, et al. [NHG guideline on medically unexplained symptoms (MUS)] (in Dutch). Huisarts Wet. 2013;56(5):222-30.
20. Sitnikova K, Leone SS, Zonneveld LNL, van Marwijk HWJ, Bosmans JE, van der Wouden JC, et al. The CIPRUS study, a

nurse-led psychological treatment for patients with undifferentiated somatoform disorder in primary care: study protocol for a randomised controlled trial. Trials. 2017;18(1):206.
21. Robbins J, Kirmayer L, Hemami S. Latent variable models of functional somatic distress. J Nerv Ment Dis. 1997;185(10):606-15.
22. Kroenke K, Spitzer RL, Williams JBW. The PHQ-15: validity of a new measure for evaluating the severity of somatic symptoms. Psychosom Med. 2002;64:258-66.
23. First M, Spitzer R, Gibbon M, Williams J. Structured clinical interview for DSM-IV-TR axis I disorders, research version, patient edition (SCID-I/P). New York: Biometrics Research, New York State Psychiatric Institute; 2002.
24. Zonneveld LN, van Rood YR, Timman R, Kooiman CG, Van't Spijker A, Busschbach JJ. Effective group training for patients with unexplained physical symptoms: a randomized controlled trial with a non-randomized one-year follow-up. PLoS One. 2012;7(8):e42629.
25. Speckens AEM, van Hemert AM, Spinhoven P, Hawton KE, Bolk JH, Rooijmans HGM. Cognitive behavioural therapy for medically unexplained physical symptoms: a randomised controlled trial. BMJ. 1995;311:1328-32.
26. Mynors-Wallis LM, Gath DH, Day A, Baker F. Randomised controlled trial of problem solving treatment, antidepressant medication, and combined treatment for major depression in primary care. BMJ. 2000;320:26-30.
27. Wilkinson P, Mynors-Wallis L. Problem-solving therapy in the treatment of unexplained physical symtpoms in primary care: a preliminary study. J Psychosom Res. 1994;38(6):591-8.
28. van der Zee KI, Sanderman R. [Measuring general health with the RAND-36, a guideline] (in Dutch). 2nd ed. Groningen: Noordelijk Centrum voor Gezondheidsvraagstukken, NCG; 2012.
29. Zigmond A, Snaith R. The Hospital Anxiety and Depression Scale. Acta Psychiatr Scand. 1983;67(6):361–70.
30. Dreer LE, Berry J, Rivera P, Snow M, Elliott TR, Miller D, et al. Efficient assessment of social problem-solving abilities in medical and rehabilitation settings: a Rasch analysis of the Social Problem-Solving Inventory-Revised. J Clin Psychol. 2009;65(7):653-69.
31. Pilowsky I. Dimensions of hypochondriasis. Br J Psychiatry. 1967;113(494):89-93.
32. Broadbent E, Petrie KJ, Main J, Weinman J. The brief illness perception questionnaire. J Psychosom Res. 2006;60(6):631-7.
33. Moss-Morris RE, Chalder T. Illness representations: where to from here? 16th Conference of the European Health Psychology Society; Kos, 2003.
34. Ryan EG, Vitoratou S, Goldsmith KA, Chalder T. Psychometric properties and factor structure of a long and shortened version of the cognitive and behavioural responses questionnaire. Psychosom Med. 2018;80(2):230-7.
35. Pearlin L, Schooler C. The structure of coping. J Health Soc Behav. 1978;19(1):2-21.
36. Hatcher RL, Gillaspy JA. Development and validation of a revised short version of the working alliance inventory. Psychother Res. 2006;16(1):12-25.
37. Cambpell MK, Fayers PM, Grimshaw JM. Determinants of the intracluster correlation coefficient in cluster randomized trials: the case of implementation research. Clin Trials. 2005;2:99-107.
38. Schulz KF, Altman DG, Moher D. CONSORT 2010 statement: updated guidelines for reporting parallel group randomised trials. Lancet 2010;375(9721):1136
39. Twisk JWR. Applied longitudinal data analysis for epidemiology. A practical guide. Second edition. Cambridge: Cambridge University Press; 2013.
40. Trimbos instituut. [Multidisciplinary guideline for MUPS and somatoform disorders] (In Dutch). 2011.
41. Krull JL, MacKinnon DP. Multilevel mediation modeling in group-based intervention studies. Eval Rev. 1999;23(4):418–44.
42. Heins MJ, Knoop H, Bleijenberg G. The role of the therapeutic relationship in cognitive behaviour therapy for chronic fatigue syndrome. Behav Res Ther. 2013;51(7):368-76.

43. Hulme K, Chilcot J, Smith MA. Doctor-patient relationship and quality of life in Irritable Bowel Syndrome: an exploratory study of the potential mediating role of illness perceptions and acceptance. Psychol Health Med. 2018;23(6):674-84.
44. Samsa G, Edelman D, Rothman ML, G.R. W, Lipscomb J, Matchar D. Determining clinically important differences in health status measures: a general approach with illustration to the Health Utilities Index Mark II. Pharmacoeconomics. 1999;02(2):141-55.
45. Woltmann E, Grogan-Kaylor A, Perron B, Georges H, Kilbourne AM, Bauwer MS. Comparative effectiveness of collaborative chronic care models for mental health conditions across primary, specialty, and behavioral health care settings: systematic review and meta-analysis. Am J Psychiatry. 2012;169:709-804.
46. Kroenke K, Unutzer J. Closing the false divide: sustainable approaches to integrating mental health services into primary care. J Gen Intern Med. 2017;32(4):404-10.
47. Liu J, Gill NS, Teodorczuk A, Li ZJ, Sun J. The efficacy of cognitive behavioural therapy in somatoform disorders and medically unexplained physical symptoms: A meta-analysis of randomized controlled trials. J Affect Disord. 2019;245:98-112.
48. Abbass A, Kisely S, Kroenke K. Short-term psychodynamic psychotherapy for somatic disorders. Systematic review and meta-analysis of clinical trials. Psychother Psychosom. 2009;78(5):265-74.
49. Wortman MS, Lucassen PL, van Ravesteijn HJ, Bor H, Assendelft PJ, Lucas C, et al. Brief multimodal psychosomatic therapy in patients with medically unexplained symptoms: feasibility and treatment effects. Fam Pract. 2016;33(4):346-53.
50. Laurant M, van der Biezen M, Wijers N, Watananirun K, Kontopantelis E, van Vught AJ. Nurses as substitutes for doctors in primary care. Cochrane Database Syst Rev. 2018;7:CD001271.
51. Hilton L, Hempel S, Ewing BA, Apaydin E, Xenakis L, Newberry S, et al. Mindfulness meditation for chronic pain: systematic review and meta-analysis. Ann Behav Med. 2017;51(2):199-213.
52. van Dessel NC, van der Wouden JC, Dekker J, van der Horst HE. Clinical value of DSM IV and DSM 5 criteria for diagnosing the most prevalent somatoform disorders in patients with medically unexplained physical symptoms (MUPS). J Psychosom Res. 2016;82:4-10.

Appendix A. Detailed results of the mixed models ITT analyses

Primary outcome **RAND-36 PCS**	**Intervention group (mean, sd)**	**Control group (mean, sd)**
Baseline	50.22 (9.89)	49.64 (9.81)
2 months	50.45 (9.67)	49.18 (9.01)
4 months	51.28 (10.57)	48.23 (8.77)
12 months	51.21 (9.93)	48.09 (10.53)
Overall effect	n/a	n/a
Secondary outcomes **RAND-36 MCS**	**Intervention group (mean, sd)**	**Control group (mean, sd)**
Baseline	49.80 (9.97)	50.22 (10.93)
2 months	50.05 (10.13)	50.12 (10.28)
4 months	49.67 (10.09)	50.25 (10.26)
12 months	49.68 (9.22)	50.44 (10.51)
Overall effect	n/a	n/a
RAND-36 Physical functioning	**Intervention group (mean, sd)**	**Control group (mean, sd)**
Baseline	62.78 (25.08)	59.25 (26.12)
2 months	61.68 (25.81)	57.79 (25.44)
4 months	66.42 (25.36)	58.71 (26.70)
12 months	67.77 (23.61)	60.63 (28.98)
Overall effect	n/a	n/a
RAND-36 Role functioning/ physical	**Intervention group (mean, sd)**	**Control group (mean, sd)**
Baseline	21.62 (32.59)	25.29 (32.84)
2 months	28.45 (36.00)	29.96 (35.71)
4 months	35.51 (41.47)	27.38 (34.20)
12 months	38.75 (41.13)	30.24 (38.18)
Overall effect	n/a	n/a
RAND-36 Role functioning/ emotional	**Intervention group (mean, sd)**	**Control group (mean, sd)**
Baseline	56.52 (44.13)	52.61 (46.01)
2 months	58.24 (44.96)	57.21 (44.12)
4 months	58.62 (43.42)	58.10 (45.64)
12 months	62.71 (43.76)	60.22 (45.10)
Overall effect	n/a	n/a
RAND-36 Social functioning	**Intervention group (mean, sd)**	**Control group (mean, sd)**
Baseline	55.63 (27.73)	57.76 (27.84)
2 months	64.49 (25.91)	62.87 (24.24)
4 months	64.77 (24.75)	63.19 (26.533)
12 months	67.59 (24.83)	63.10 (27.63)
Overall effect	n/a	n/a
RAND-36 Bodily pain	**Intervention group (mean, sd)**	**Control group (mean, sd)**
Baseline	46.24 (21.31)	45.79 (22.19)
2 months	52.85 (21.71)	49.55 (24.01)
4 months	56.77 (21.42)	50.17 (21.77)

Crude analyses		Adjusted analyses [a]	
B (95% CI)	**p-value**	**B (95% CI)**	**p-value**
1.25 (-0.81 to 3.30)	0.24	1.66 (-0.52 to 3.84)	0.14
2.08 (0.05 to 4.12)	0.045*	2.93 (0.77 to 5.09)	0.008*
2.10 (-0.02 to 4.23)	0.053	2.09 (-0.14 to 4.33)	0.066
1.80 (0.19 to 3.42)	0.029*	2.24 (0.51 to 3.97)	0.011*
Crude analyses		**Adjusted analyses [a]**	
B (95% CI)	**p-value**	**B (95% CI)**	**p-value**
-0.02 (-2.46 to 2.43)	0.99	0.02 (-2.37 to 2.42)	0.99
-0.69 (-3.10 to 1.73)	0.58	-0.67 (-3.10 to 1.76)	0.59
-1.02 (-3.54 to 1.51)	0.43	-0.44 (-2.97 to 2.08)	0.73
-0.55 (-2.47 to 1.37)	0.57	-0.35 (-2.22 to 1.52)	0.71
Crude analyses		**Adjusted analyses [a]**	
B (95% CI)	**p-value**	**B (95% CI)**	**p-value**
0.04 (-4.48 to 4.57)	0.99	0.56 (-4.16 to 5.28)	0.82
3.66 (-0.84 to 8.16)	0.11	5.00 (0.28 to 9.73)	0.038*
0.65 (-4.05 to 5.35)	0.79	1.35 (-3.63 to 6.13)	0.62
1.47 (-2.08 to 5.02)	0.42	2.33 (-1.40 to 6.06)	0.21
Crude analyses		**Adjusted analyses [a]**	
B (95% CI)	**p-value**	**B (95% CI)**	**p-value**
3.56 (-7.43 to 14.56)	0.53	7.18 (-4.33 to 18.69)	0.22
9.33 (-1.54 to 20.20)	0.093	13.57 (2.21 to 24.92)	0.019*
8.77 (-2.65 to 20.20)	0.13	11.51 (-0.35 to 23.37)	0.057
7.17 (-1.16 to 15.50)	0.091	10.82 (2.14 to 19.49)	0.015*
Crude analyses		**Adjusted analyses [a]**	
B (95% CI)	**p-value**	**B (95% CI)**	**p-value**
-3.21 (16.19 to 9.77)	0.63	0.03 (-13.26 to 13.31)	1.00
-2.45 (-15.26 to 10.36)	0.71	0.66 (-12.47 to 13.78)	0.92
-0.62 (-14.10 to 12.86)	0.93	3.55 (-10.21-17.31)	0.61
-2.14 (-11.63 to 7.36)	0.66	1.41 (-8.29 to 11.10)	0.78
Crude analyses		**Adjusted analyses [a]**	
B (95% CI)	**p-value**	**B (95% CI)**	**p-value**
2.44 (-4.54 to 9.41)	0.45	2.04 (-5.31 to 9.40)	0.59
1.85 (-5.05 to 8.75)	0.60	2.22 (-5.05 to 9.50)	0.55
3.78 (-3.42 to 10.98)	0.30	3.74 (-3.80 to 11.28)	0.33
2.65 (-2.85 to 8.14)	0.35	2.66 (-3.09 to 8.41)	0.37
Crude analyses		**Adjusted analyses [a]**	
B (95% CI)	**p-value**	**B (95% CI)**	**p-value**
3.26 (-2.22 to 8.74)	0.24	4.45 (-1.24 to 10.14)	0.13
4.98 (-0.44 to 10.40)	0.072	6.45 (0.83 to 12.08)	0.025*

12 months	57.19 (20.20)	51.21 (25.89)
Overall effect	n/a	n/a
RAND-36 Emotional well-being	**Intervention group (mean, sd)**	**Control group (mean, sd)**
Baseline	59.72 (17.04)	60.43 (20.23)
2 months	60.18 (17.43)	60.50 (20.29)
4 months	63.91 (18.61)	63.77 (19.30)
12 months	62.76 (16.84)	64.51 (19.81)
Overall effect	n/a	n/a
RAND-36 Energy/fatigue	**Intervention group (mean, sd)**	**Control group (mean, sd)**
Baseline	36.89 (16.90)	38.65 (15.95)
2 months	40.19 (17.18)	38.80 (15.81)
4 months	42.82 (18.12)	41.36 (14.32)
12 months	46.71 (18.70)	46.06 (17.81)
Overall effect	n/a	n/a
RAND-36 General health	**Intervention group (mean, sd)**	**Control group (mean, sd)**
Baseline	43.81 (17.64)	41.43 (16.25)
2 months	44.51 (17.86)	43.50 (16.67)
4 months	47.14 (20.38)	47.54 (18.06)
12 months	49.09 (19.37)	47.06 (17.47)
Overall effect	n/a	n/a
Anxiety symptoms (HADS-A)	**Intervention group (mean, sd)**	**Control group (mean, sd)**
Baseline	7.89 (3.80)	7.69 (4.52)
2 months	8.47 (4.22)	7.32 (4.28)
4 months	7.46 (4.07)	6.72 (4.43)
12 months	6.89 (4.09)	6.26 (4.41)
Overall effect	n/a	n/a
Depressive symptoms (HADS-D)	**Intervention group (mean, sd)**	**Control group (mean, sd)**
Baseline	7.00 (3.86)	7.52 (4.23)
2 months	6.93 (3.91)	7.26 (4.29)
4 months	6.29 (4.04)	6.62 (4.31)
12 months	5.83 (4.15)	6.29 (4.54)
Overall effect	n/a	n/a
Somatic symptom severity (PHQ-15)	**Intervention group (mean, sd)**	**Control group (mean, sd)**
Baseline	13.63 (4.89)	13.47 (4.43)
2 months	13.13 (5.36)	13.74 (5.23)
4 months	12.83 (5.20)	12.74 (4.95
12 months	11.86 (5.27)	12.03 (5.78)
Overall effect	n/a	n/a

Abbreviations: 95%CI: 95% Confidence Interval; HADS-A, Hospital Anxiety and Depression Scale-Anxiety subscale; HADS-D, Hospital Anxiety and Depression Scale-Depression subscale; ITT: Intention to Treat; MCS: Mental Component Summary Score; n/a: not applicable; PCS: Physical Component Summary Score; PHQ-15: Patient Health Questionnaire 15-item somatic symptom severity scale; sd: standard deviation
[a] adjusted for (if necessary): gender, age, level of education, duration of symptoms, somatic symptom severity (PHQ-15), anxiety symptoms (HADS-A), depressive symptoms (HADS-D), number of comorbid physical diseases, time interval baseline – 2-months follow-up, time interval baseline – 4-months follow-up
* $p<0.05$

3.69 (-1.97 to 9.34)	0.20	4.08 (-1.75 to 9.91)	0.17
3.98 (-0.31 to 8.27)	0.069	5.08 (0.58 to 9.57)	0.027*

Crude analyses		Adjusted analyses [a]	
B (95% CI)	**p-value**	**B (95% CI)**	**p-value**
0.04 (-3.91 to 4.00)	0.98	0.22 (-3.80 to 4.23)	0.92
-0.12 (-4.04 to 3.80)	0.95	0.85 (-3.14 to 4.85)	0.68
-2.36 (-6.43 to 1.71)	0.23	-1.68 (-5.79 to 2.44)	0.43
-0.77 (-3.88 to 2.35)	0.63	-0.13 (-3.28 to 3.02)	0.93

Crude analyses		Adjusted analyses [a]	
B (95% CI)	**p-value**	**B (95% CI)**	**p-value**
3.31 (-0.87 to 7.49)	0.12	2.08 (-2.15 to 6.31)	0.34
2.22 (-1.91 to 6.36)	0.29	1.76 (-2.44 to 5.96)	0.41
2.05 (-2.25 to 6.36)	0.35	1.90 (-2.43 to 6.24)	0.39
2.56 (-0.64 to 5.75)	0.12	1.98 (-1.24 to 5.20)	0.23

Crude analyses		Adjusted analyses [a]	
B (95% CI)	**p-value**	**B (95% CI)**	**p-value**
-0.60 (-3.82 to 5.02)	0.79	0.64 (-4.11 to 5.38)	0.79
-1.53 (-5.91 to 2.84)	0.49	-1.21 (-5.92 to 3.50)	0.62
0.28 (-4.23 to 4.79)	0.91	0.63 (-4.21 to 5.46)	0.80
-0.28 (-3.90 to 3.34)	0.88	0.05 (-3.91 to 4.02)	0.98

Crude analyses		Adjusted analyses [a]	
B (95% CI)	**p-value**	**B (95% CI)**	**p-value**
0.48 (-0.34 to 1.29)	0.25	0.51 (-0.29 to 1.31)	0.21
0.19 (-0.62 to 1.00)	0.65	0.17 (-0.63 to 0.98)	0.67
0.24 (-0.60 to 1.08)	0.58	0.32 (-0.50 to 1.15)	0.44
0.30 (-0.33 to 0.94)	0.35	0.33 (-0.29 to 0.94)	0.30

Crude analyses		Adjusted analyses [a]	
B (95% CI)	**p-value**	**B (95% CI)**	**p-value**
-0.05 (-0.83 to 0.72)	0.89	-0.12 (-0.93 to 0.68)	0.77
-0.06 (-0.84 to 0.71)	0.87	-0.21 (-1.02 to 0.60)	0.61
-0.05 (-0.85 to 0.75)	0.90	-0.34 (-1.17 to 0.48)	0.42
-0.06 (-0.66 to 0.55)	0.86	-0.23 (-0.89 to 0.43)	0.49

Crude analyses		Adjusted analyses [a]	
B (95% CI)	**p-value**	**B (95% CI)**	**p-value**
-1.06 (-2.25 to 0.14)	0.083	-1.19 (-2.42 to 0.04)	0.057
-0.05 (-1.25 to 1.14)	0.93	-0.33 (-1.56 to 0.91)	0.60
-0.04 (-1.62 to 0.89)	0.56	-0.47 (-1.75 to 0.80)	0.47
-0.51 (-1.43 to 0.40)	0.27	-0.69 (-1.64 to 0.24)	0.15

Appendix B. Results of mixed models analyses split up per group for effect modifiers: symptom duration and physical comorbidity

Primary outcome **RAND-36 PCS**	**Symptom duration below median (n=70-98)**		**Symptom duration above median (n=72-98)**	
	Intervention group (mean, sd)	**Control group (mean, sd)**	**Intervention group (mean, sd)**	**Control group (mean, sd)**
Baseline	51.52 (9.80)	51.20 (10.05)	48.78 (9.97)	48.20 (9.57)
2 months	52.08 (9.94)	49.95 (9.69)	48.88 (9.25)	48.52 (8.65)
4 months	54.90 (9.87)	50.74 (8.97)	47.41 (10.15)	46.00 (8.21)
12 months	54.13 (9.02)	48.63 (12.00)	48.01 (10.13)	47.42 (9.38)
Overall effect	n/a	n/a	n/a	n/a
Secondary outcomes **RAND-36 Bodily pain**	**Symptom duration below median (n=70-98)**		**Symptom duration above median (n=72-98)**	
	Intervention group (mean, sd)	**Control group (mean, sd)**	**Intervention group (mean, sd)**	**Control group (mean, sd)**
Baseline	48.35 (22.74)	45.51 (22.36)	43.92 (19.76)	45.56 (22.30)
2 months	57.05 (23.05)	50.95 (26.21)	48.68 (19.58)	48.34 (22.81)
4 months	62.59 (19.86)	52.68 (22.67)	50.58 (21.77)	46.94 (21.03)
12 months	62.34 (18.00)	50.80 (27.53)	51.83 (21.44)	50.78 (24.86)
Overall effect	n/a	n/a	n/a	n/a
RAND -36 General health	**Symptom duration below median (n=70-98)**		**Symptom duration above median (n=72-98)**	
	Intervention group (mean, sd)	**Control group (mean, sd)**	**Intervention group (mean, sd)**	**Control group (mean, sd)**
Baseline	44.53 (17.67)	43.33 (18.00)	43.17 (17.88)	39.77 (14.74)
2 months	46.36 (16.51)	40.69 (18.50)	42.68 (19.38)	45.45 (15.10)
4 months	49.64 (19.41)	47.90 (20.77)	43.69 (20.78)	47.04 (16.02)
12 months	53.10 (18.21)	45.00 (19.53)	22.87 (20.15)	48.09 (15.52
Overall effect	n/a	n/a	n/a	n/a
Primary outcome **RAND-36 PCS**	**0-2 comorbid physical diseases (n=69-91)**		**3 or more comorbid physical diseases (n=75-106)**	
	Intervention group (mean, sd)	**Control group (mean, sd)**	**Intervention group (mean, sd)**	**Control group (mean, sd)**
Baseline	53.64 (9.38)	55.07 (8.78)	53.64 (9.38)	55.07 (8.78)
2 months	53.65 (8.48)	52.85 (8.20)	53.65 (8.48)	52.85 (8.20)
4 months	55.24 (9.63)	51.47 (8.51)	55.24 (9.63)	51.47 (8.51)
12 months	55.38 (8.20)	50.85 (10.21)	55.38 (8.20)	50.85 (10.21)
Overall effect	n/a	n/a	n/a	n/a
Secondary outcome **RAND-36 General health**	**0-2 comorbid physical diseases (n=69-91)**		**3 or more comorbid physical diseases (n=75-106)**	
	Intervention group (mean, sd)	**Control group (mean, sd)**	**Intervention group (mean, sd)**	**Control group (mean, sd)**
Baseline	49.01 (16.64)	46.41 (14.49)	49.01 (16.64)	46.41 (14.49)
2 months	49.39 (17.25)	46.21 (16.18)	49.39 (17.25)	46.21 (16.18)
4 months	54.97 (19.96)	53.50 (16.20)	54.97 (19.96)	53.50 (16.20)
12 months	55.49 (18.57)	49.31 (15.91)	55.49 (18.57)	49.31 (15.91)
Overall effect	n/a	n/a	n/a	n/a

Abbreviations: 95%CI: 95% Confidence Interval; n/a: not applicable; PCS: Physical Component Summary Score; sd: standard deviation

* p<0.05

Crude difference			
Symptom duration below median		Symptom duration above median	
B (95% CI)	p-value	B (95% CI)	p-value
n/a	n/a	n/a	n/a
n/a	n/a	n/a	n/a
n/a	n/a	n/a	n/a
3.83 (1.57 to 6.09)	0.001*	-0.18 (-2.40 to 2.03)	0.87

Crude difference			
Symptom duration below median		Symptom duration above median	
B (95% CI)	p-value	B (95% CI)	p-value
n/a	n/a	n/a	n/a
n/a	n/a	n/a	n/a
n/a	n/a	n/a	n/a
6.94 (1.05 to 12.84)	0.021*	0.68 (-5.41 to 6.78)	0.83

Crude difference			
Symptom duration below median		Symptom duration above median	
B (95% CI)	p-value	B (95% CI)	p-value
n/a	n/a	n/a	n/a
n/a	n/a	n/a	n/a
n/a	n/a	n/a	n/a
5.74 (0.96 to 10.52)	0.019*	-5.98 (-11.09 to -0.86)	0.022*

Crude difference			
0-2 comorbid physical diseases		3 or more comorbid physical diseases	
B (95% CI)	p-value	B (95% CI)	p-value
n/a	n/a	n/a	n/a
n/a	n/a	n/a	n/a
n/a	n/a	n/a	n/a
3.55 (1.13 to 5.97)	0.004*	-0.02 (2.09 to 2.05)	0.99

Crude difference			
0-2 comorbid physical diseases		3 or more comorbid physical diseases	
B (95% CI)	p-value	B (95% CI)	p-value
n/a	n/a	n/a	n/a
n/a	n/a	n/a	n/a
n/a	n/a	n/a	n/a
2.38 (-2.60 to 7.36)	0.35	-3.09 (-8.16 to 1.98)	0.23

Appendix C. Results of the mixed models per protocol analyses with the group of intervention patients who attended all 6 sessions with the MHNP

	Crude analyses		Adjusted analyses [a]	
	B (95% CI)	p-value	B (95% CI)	p-value
Primary outcome				
RAND-36 PCS	2.23 (0.39 to 4.08)	0.018*	2.20 (0.26 to 4.14)	0.026*
Secondary outcomes				
RAND-36				
MCS	-0.44 (-2.64 to 1.76)	0.70	-0.52 (-2.63 to 1.600)	0.63
Physical functioning	2.21 (-2.01 to 6.43)	0.30	2.54 (-1.68 to 6.76)	0.24
Role functioning/physical	7.87 (-1.72 to 17.47)	0.11	9.60 (-0.31 to 19.51)	0.058
Role functioning/emotional	-0.03 (-10.95 to 10.90)	1.00	0.29 (-10.62 to 11.20)	0.96
Social functioning	3.10 (-3.04 to 9.24)	0.32	1.88 (-4.27 to 8.04)	0.55
Bodily pain	4.29 (-0.53 to 9.11)	0.081	4.29 (-0.53 to 9.11)	0.081
Emotional well-being	-1.09 (-4.79 to 2.61)	0.56	-0.80 (4.44 to 2.85)	0.67
Energy/fatigue	3.03 (-0.60 to 6.65)	0.10	2.00 (-1.52 to 5.51)	0.27
General health	1.15 (-3.13 to 5.42)	0.60	0.15 (-4.46 to 4.75)	0.95
Anxiety symptoms (HADS-A)	0.19 (-0.53 to 0.91)	0.60	0.36 (-0.32 to 1.04)	0.30
Depressive symptoms (HADS-D)	-0.23 (-0.95 to 0.49)	0.53	-0.21 (-0.96 to 0.53)	0.58
Somatic symptom severity (PHQ-15)	-0.92 (-1.95 to 0.11)	0.080	-0.54 (-1.51 to 0.43)	0.28

Abbreviations: 95%CI: 95% Confidence Interval; HADS-A: Hospital Anxiety and Depression Scale-Anxiety subscale; HADS-D: Hospital Anxiety and Depression Scale-Depression subscale; MCS: Mental Component Summary Score; MHNP: mental health nurse practitioner; PCS: Physical Component Summary Score; PHQ-15: Patient Health Questionnaire 15-item somatic symptom severity scale

[a] adjusted for (if necessary): gender, age, level of education, duration of symptoms, somatic symptom severity (PHQ-15), anxiety symptoms (HADS-A), depressive symptoms (HADS-D), number of comorbid physical diseases, time interval baseline – 2-months follow-up, time interval baseline – 4-months follow-up

* $p < 0.05$

Appendix D. Results of the mixed models per protocol analyses with the group of intervention patients who attended all 6 or less sessions with the MHNP because their goals were achieved before the 6th session

	Crude analyses		Adjusted analyses [a]	
	B (95% CI)	p-value	B (95% CI)	p-value
Primary outcome				
RAND-36 PCS	2.16 (0.37 to 3.96)	0.018*	2.09 (0.24 to 3.94)	0.027*
Secondary outcomes				
RAND-36				
MCS	0.05 (-2.10 to 2.21)	0.96	-0.25 (-2.34 to 1.84)	0.81
Physical functioning	1.98 (-2.13 to 6.09)	0.35	1.71 (-2.38 to 5.80)	0.41
Role functioning/physical	8.70 (-0.62 to 18.03)	0.067	9.99 (0.30-19.68)	0.043*
Role functioning/emotional	1.22 (-9.35 to 11.79)	0.82	1.19 (-9.18 to 11.55)	0.82
Social functioning	4.14 (-1.85 to 10.12)	0.18	2.55 (-3.46 to 8.56)	0.41
Bodily pain	4.37 (-0.27 to 9.02)	0.065	3.85 (-0.65 to 8.36)	0.093
Emotional well-being	-0.45 (-4.07 to 3.17)	0.81	-0.57 (-4.16 to 3.01)	0.76
Energy/fatigue	3.97 (0.39 to 7.54)	0.030*	2.66 (-0.84 to 6.15	0.14
General health	1.15 (-2.99 to 5.29)	0.59	0.08 (-4.41 to 4.56)	0.97
Anxiety symptoms (HADS-A)	0.10 (-0.59 to 0.80)	0.77	0.38 (-0.30 to 1.05)	0.27
Depressive symptoms (HADS-D)	-0.31 (-1.00 to 0.39)	0.39	-0.23 (-0.99 to 0.52)	0.54
Somatic symptom severity (PHQ-15)	-1.04 (-2.03 to -0.04)	0.041*	-0.58 (-1.52 to 0.36)	0.23

Abbreviations: 95%CI: 95% Confidence Interval; HADS-A: Hospital Anxiety and Depression Scale-Anxiety subscale; HADS-D: Hospital Anxiety and Depression Scale-Depression subscale; MCS: Mental Component Summary Score; MHNP: mental health nurse practitioner; PCS: Physical Component Summary Score; PHQ-15: Patient Health Questionnaire 15-item somatic symptom severity scale

[a] adjusted for (if necessary): gender, age, level of education, duration of symptoms, somatic symptom severity (PHQ-15), anxiety symptoms (HADS-A), depressive symptoms (HADS-D), number of comorbid physical diseases, time interval baseline – 2-months follow-up, time interval baseline – 4-months follow-up

* p <0.05

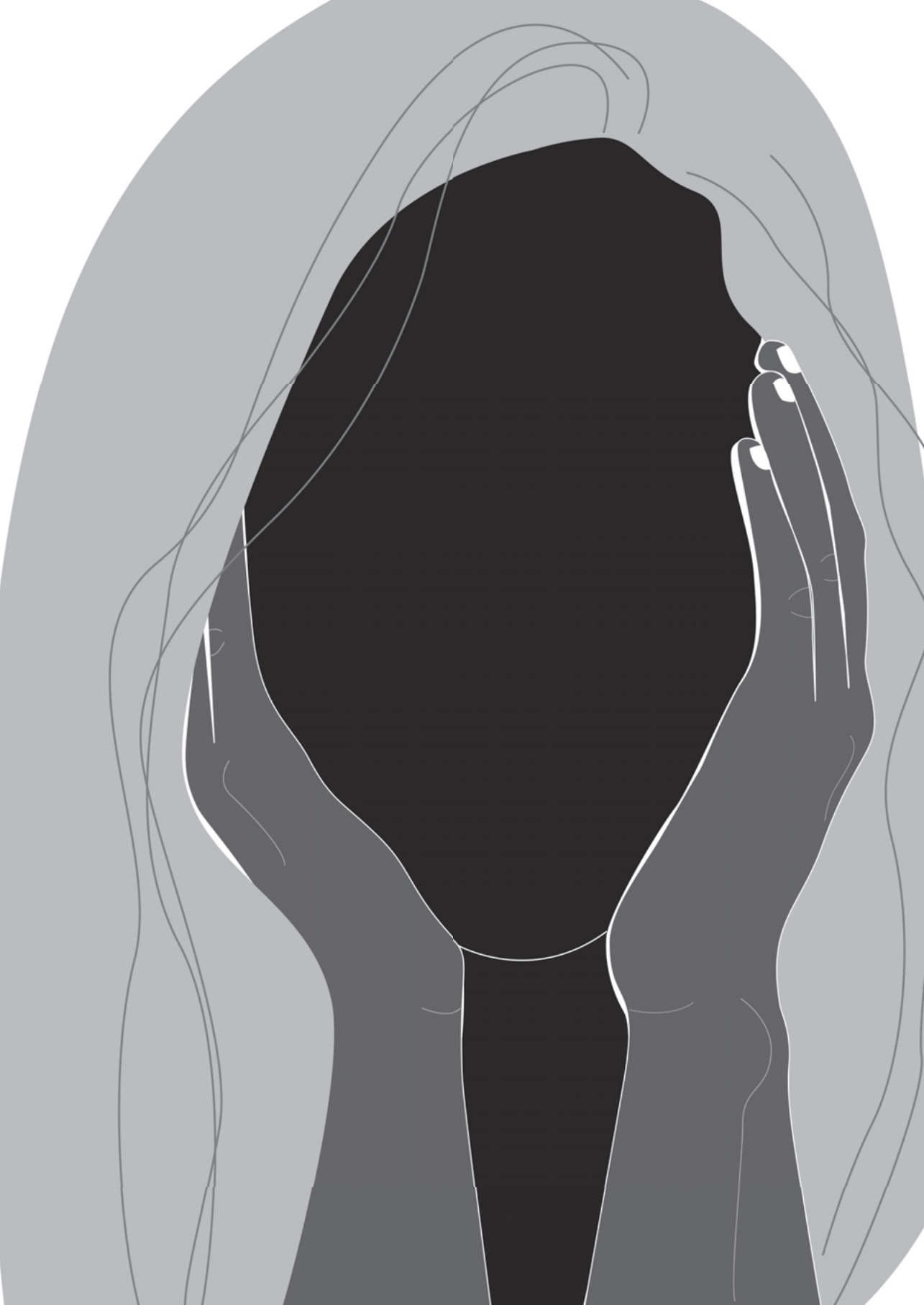
Appendix E. Results of the mixed models per protocol analyses with the group of intervention patients who attended at least 4 of the 6 sessions with the MHNP

	Crude analyses		Adjusted analyses [a]	
	B (95% CI)	p-value	B (95% CI)	p-value
Primary outcome				
RAND-36 PCS	2.25 (0.52 to 3.98)	0.011*	2.34 (0.55 to 4.14)	0.011*
Secondary outcomes				
RAND-36				
MCS	-0.09 (-2.14 to 1.97)	0.93	-0.55 (-2.59 to 1.50)	0.60
Physical functioning	2.04 (-1.78 to 5.86)	0.30	2.10 (-1.82 to 6.01)	0.29
Role functioning/physical	8.57 (-0.32 to 17.47)	0.059	9.51 (-0.35 to 18.66)	0.042*
Role functioning/emotional	0.55 (-9.61 to 10.71)	0.92	-0.23 (-10.65 to 10.18)	0.97
Social functioning	3.65 (-2.18 to 9.49)	0.22	1.17 (-4.63 to 6.98)	0.69
Bodily pain	4.82 (0.23 to 9.41)	0.040*	4.82 (0.23 to 9.41)	0.040*
Emotional well-being	-0.32 (-3.68 to 3.05)	0.85	-0.63 (-4.03 to 2.76)	0.72
Energy/fatigue	3.38 (-0.07 to 6.84)	0.055	1.88 (-1.46 to 5.22)	0.27
General health	0.86 (-3.02 to 4.74)	0.67	0.47 (-3.53 to 4.47)	0.82
Anxiety symptoms (HADS-A)	0.08 (-0.58 to 0.73)	0.82	0.42 (-0.22 to 1.06)	0.20
Depressive symptoms (HADS-D)	-0.20 (-0.86 to 0.45)	0.54	-0.16 (0.89 to 0.56)	0.66
Somatic symptom severity (PHQ-15)	-0.87 (-1.82 to 0.09)	0.074	-0.54 (-1.45 to 0.38)	0.25

Abbreviations: 95%CI: 95% Confidence Interval; HADS-A: Hospital Anxiety and Depression Scale-Anxiety subscale; HADS-D: Hospital Anxiety and Depression Scale-Depression subscale; MCS: Mental Component Summary Score; MHNP: mental health nurse practitioner; PCS: Physical Component Summary Score; PHQ-15: Patient Health Questionnaire 15-item somatic symptom severity scale

[a] adjusted for (if necessary): gender, age, level of education, duration of symptoms, somatic symptom severity (PHQ-15), anxiety symptoms (HADS-A), depressive symptoms (HADS-D), number of comorbid physical diseases, time interval baseline – 2-months follow-up, time interval baseline – 4-months follow-up

* $p < 0.05$



Chapter 6

A brief cognitive behavioural intervention is cost-effective for primary care patients with medically unexplained physical symptoms compared to usual care

Kate Sitnikova
Aureliano P Finch
Stephanie S Leone
Judith E Bosmans
Harm WJ van Marwijk
Henriëtte E van der Horst
Johannes C van der Wouden

ABSTRACT

Objective: To assess the cost-effectiveness of a brief cognitive behavioural intervention for patients with medically unexplained physical symptoms (MUPS) provided by a mental health nurse practitioner (MHNP) in primary care in comparison with usual care.

Methods: We performed an economic evaluation from a societal perspective alongside a cluster randomized controlled trial with 12 months follow-up. The primary outcome was quality-adjusted life-years (QALYs). Secondary outcomes were the RAND-36 physical component summary score (PCS), somatic symptom severity (Patient Health Questionnaire (PHQ-15), and anxiety and depression symptoms (Hospital Anxiety and Depression Scale (HADS)). Missing data were imputed using multiple imputation. We used non-parametric bootstrapping to estimate statistical uncertainty. The bootstrapped cost-effect pairs were used to estimate cost-effectiveness planes and cost-effectiveness acceptability curves.

Results: Mean total costs in the intervention group were significantly lower than in the usual care group (mean difference -2300€, 95% CI -3257 to -134). The mean difference in QALYs was 0.01 (95% CI -0.01 to 0.04), in PCS 2.46 (95% CI 1.44 to 3.47), in PHQ-15 -0.26 (95% CI -0.81 to 0.28), and in HADS -0.07 (-0.81 to 0.67). At a willingness to pay of 0 € per additional unit of effect, the probability of the intervention being cost-effective was 0.93 for QALYs and 0.92 for PCS, PHQ-15 and HADS scores.

Conclusion: Our intervention is cost-effective compared to usual care for patients with MUPS. Implementation of the intervention has the potential to result in a significant decline in costs. However, large scale implementation would require increased deployment of MHNPs.

INTRODUCTION

Patients with medically unexplained physical symptoms (MUPS) are frequently encountered in all healthcare settings and particularly in primary care (1-3). If MUPS persist, the symptoms may be considered severe enough to be classified as a DSM-IV somatoform disorder (4), or as a somatic symptom disorder according to the DSM-5 (5). MUPS substantially affect health-related quality of life (HRQoL), cause high levels of functional impairment and are associated with mental disorders, such as anxiety or depression (1, 3, 6). Moreover, MUPS are associated with substantial costs (7). Total societal costs among patients with MUPS amounted to €21.2 billion in a 2010 European study (8).

According to various systematic reviews, cognitive behavioural therapy (CBT) is the most effective form of treatment for MUPS (9, 10). After receiving CBT, a substantial number of patients report less physical symptoms, disability and psychological distress. However, patients do not always turn to healthcare providers who deliver CBT, as these are commonly situated in a mental healthcare setting outside of general practice. Both patients and general practitioners (GPs) may not feel comfortable turning to mental healthcare for physical symptoms and only patients with severe symptoms and high functional impairment are referred (11, 12).

Patients with mild to moderate MUPS may benefit from referral to a mental health nurse practitioner (MHNP) which is advocated in the current Dutch GP guideline (13). The MHNP is typically a mental health nurse or psychologist within the general practice, who provides counselling to patients with mild psychosocial problems. The MHNP was introduced in 2014 by the Dutch government to decrease the growing mental healthcare costs, to decrease GPs' workload, and to offer more accessible mental health services within the familiar surroundings of a general practice (14).

To establish the effectiveness of CBT for patients with mild to moderate MUPS in primary care, we recently conducted a randomised controlled trial, the Cognitive behavioural Intervention in PRimary care for Undifferentiated Somatoform disorder (CIPRUS) study comparing a CBT-based intervention provided by MHNPs in addition to usual care, to usual care alone (15). The intervention was effective in improving physical functioning and in decreasing pain and limitations due to physical symptoms. It was particularly effective in patients with symptoms that had been present for a limited number of years and who had few or no comorbid physical diseases.

Since healthcare resources are scarce and MUPS are associated with substantial costs (7, 8), it is important to also evaluate the cost-effectiveness of interventions to treat MUPS, besides evaluating their clinical effectiveness. Two systematic reviews (7, 16) show that CBT is cost-effective compared to pharmacological and non-pharmacological treatment and waiting-list controls. However, only a minority of the included studies included both clinical effects and quality adjusted life years (QALYs) as outcome measures. QALYs are generally considered as the most important outcome in health economic evaluations, since this outcome reflects the societal desirability of a specific health state. Additionally, QALYs can be compared across diseases. Traditionally, this was done using life-years gained, but considering that MUPS is not a life-threatening condition this is not an appropriate outcome measure in this context. Therefore, the aim of this study is to conduct an economic evaluation of a cognitive behavioural intervention delivered by MHNPs for MUPS patients on top of usual care, from a societal perspective. This evaluation includes both QALYs and physical functioning as outcome measures.

METHODS

Trial design

We conducted a cluster-randomised controlled trial with 12 months follow-up in the Netherlands between August 2015 and May 2018. The VU University Medical Center Ethics Committee approved the study (number 2014.305, 9 July 2014, amendment 5 August 2016). The design of the trial is described in more detail elsewhere (17). The trial is registered in the Dutch Trial Registry, www.trialregister.nl under NTR4686.

Treatment allocation, participants and procedures

Cluster randomisation was used to avoid contamination between treatment groups. Since a MHNP could be affiliated with more than one general practice, a cluster was defined by the MHNP (n=31) rather than the participating general practice (n=85). An independent epidemiologist carried out concealed random allocation of clusters to the intervention or usual care condition using a computer-generated randomisation list.

Participants were recruited from 85 general practices throughout the Netherlands. Participants were eligible for the trial if they were 18 years old and above, and had MUPS. MUPS was operationalised as fitting the DSM-IV classification criteria for undifferentiated somatoform disorder (USD). We chose this operationalisation to ensure we selected patients who had complaints for at least 6 months and were significantly impaired by these complaints.

Exclusion criteria were: having a medical or psychological disorder that explained the reported symptoms that would allow patients to participate in our study; having a severe psychiatric disorder (e.g. psychotic disorder); currently receiving psychological help for MUPS; having poor language skills or physical handicaps that would prevent patients from understanding the intervention or questionnaires.

GPs selected patients from their electronic databases who had consulted them with one or more symptoms from the 'Robbins' list (18) at least twice in the previous 3 months. The Robbins list consists of 23 physical symptoms that are associated with functional somatic syndromes. Potentially eligible patients who met the inclusion criteria received concise information about the study and the Patient Health Questionnaire 15-item somatic symptom severity scale (PHQ-15) (19) from their GP. Interested patients with a PHQ-15 score of at least 5 (low symptom severity) were provided with information on the study and invited to participate in a clinical interview (Structured Clinical Interview for DSM-IV Axis I Disorders (SCID-I)) (20). Patients meeting the DSM-IV criteria for USD and giving informed consent were included in the study.

Interventions

The intervention consisted of six individual 30-minute sessions with a MHNP in the general practice in addition to usual care. Before delivering the intervention, MHNPs followed two group training sessions lasting 3 to 3.5 hours each. The session were led by a clinical psychologist specialised in treating somatoform disorders. MHNPs also received an intervention manual describing each session in detail. Supervision by the clinical psychologist who had trained them was provided if needed.

The intervention consisted of a combination of two CBT methods: the modified version of the consequences model for somatoform disorders (21) and Problem-Solving Treatment (PST) (22). The consequences model focuses on the consequences or problems that arise due to physical symptoms, rather than on their possible (unknown) cause(s). PST is a cognitive behavioural problem-solving approach consisting of seven steps. During the sessions, patients first identified the consequences or problems they experienced in daily life due to MUPS in collaboration with their MHNP. The identified consequences or problems were then tackled using the seven PST steps. The goal was to enhance patients' problem solving skills in order to deal with the consequences of their physical symptoms and other problems that may arise in daily life. The intervention is described in more detail elsewhere (17).

The usual care group did not receive any additional intervention other than the care they would usually receive from their GP or any other healthcare providers they were referred to for their USD symptoms. Usual care is generally based on the applicable GP guideline and multidisciplinary guideline for management of MUPS and somatoform disorders (13, 23).

Resource use and unit costs

Information on resource use was retrospectively collected at baseline and 4, 8 and 12 months of follow-up using an adapted version of the Trimbos and iMTA questionnaire on Costs associated with Psychiatric Illness (Tic-P) (24). Intervention costs were calculated using a micro-costing bottom-up approach and included costs of training sessions for the MHNPs and the six intervention sessions of 30 minutes. Healthcare costs included primary care costs such as visits to the GP, MHNP, physiotherapist, complementary medicine and psychologists; secondary care costs such as medical specialists, psychotherapists and diagnostics; and medication costs (both prescribed and over-the-counter medication). Other costs included productivity losses resulting from absenteeism and presenteeism, and paid or unpaid help, for instance with domestic work.

Healthcare costs were estimated by multiplying healthcare utilisation with the standard prices reported in the Dutch costing guidelines (25). Costs of medication were calculated using prices of the Royal Dutch Society for Pharmacy (26).

Absenteeism from paid work was assessed by asking participants how many of their working days they had called in sick during the previous period of 4 months. Costs of presenteeism were assessed by asking participants how many of their working hours would have to be replaced due to reduced productivity while being present at work. Costs of absenteeism and presenteeism were calculated using sex-specific mean wages of the Dutch population (27). Absenteeism costs from paid work were estimated according to the friction cost approach. The friction cost approach assumes that a sick employee is replaced by another employee after a certain period of time i.e. the friction period. Productivity losses are assumed to occur during this friction period only. A friction period of 85 days (12 weeks) was used in our analysis.

All costs were indexed for the year 2016. Discounting was not necessary because the time horizon of the economic evaluation was limited to 12 months.

Outcome measures

Primary outcome

The primary outcome of this study was quality adjusted life years (QALYs). QALYs are an index, i.e. a utility, summarizing the length of life and HRQoL (28). HRQoL was measured with the EQ-5D-5L (29) at baseline and at 2, 4, 8 and 12 months after baseline. The EQ-5D-5L is the most frequently used preference-based HRQoL instrument in health technology assessment (30) and has been shown to be valid and responsive across multiple conditions (31). It describes health in terms of five dimensions (mobility, self-care, usual activities, pain/discomfort and anxiety/depression). Each dimension has 5 levels (from 'no problems with…' to 'unable to…'). The health state indicated by patients on the EQ-5D-5L was converted to a utility score using the Dutch EQ-5D-5L tariff (32). The EQ-5D-5L utility scores at different time points were used to calculate QALYs using the area under the curve method. Changes between health states at different time points were considered linear.

Secondary outcomes

All secondary outcomes were assessed at baseline and at 2, 4 and 12 months after baseline. The improvement in patients' physical functioning during the total 12-months follow-up period was measured by the physical component summary score (PCS) of the RAND-36 questionnaire (33). Higher scores indicate better physical functioning. The PHQ-15 somatic symptom severity scale (19) was used to measure the severity of the somatic symptoms. Anxiety and depressive symptoms were measured using the Hospital Anxiety and Depression Scale (HADS) (34). Higher scores on the PHQ-15 and HADS indicate more severe symptoms.

Sample size

In the clinical trial, we aimed to detect a clinically relevant effect size of 0.4 standard deviations (SD) on the primary outcome (the PCS of the RAND-36), using a two-sided significance level of 5%, a power of 80%, and an allocation ratio of 1:1. We assumed a correlation coefficient of 0.5 for repeated measurements. Using linear mixed models with these assumptions required a sample size of 74 patients per condition. After correcting for the cluster design (assuming an average cluster size of 4 and an intracluster correlation coefficient (ICC) of 0.05) (35) and accounting for a potential dropout rate of 20%, we aimed to include 106 patients in each condition.

Statistical analysis

The analyses were conducted according to the intention-to-treat principle. Missing data on costs and outcomes were imputed using multiple imputations with chained equations

(MICE) and stratified by treatment group, using predictive mean matching. An imputation model was created that contained all variables in the analysis models, characteristics differing between groups at baseline, variables related to missing data and variables related to outcome variables. Twenty imputed datasets were created, resulting in a loss of efficiency of less than 5% (36). The imputed datasets were each analysed separately and the results of the analyses were pooled using Rubin's rules (37).

Bivariate regression analyses were used to estimate differences in costs and effects. Incremental Cost-Effectiveness Ratios (ICERs) were calculated by dividing the difference in costs between the groups by the difference in effects. Non-parametric bias-corrected and accelerated bootstrapping with 5000 replications was used to estimate statistical uncertainty. The bootstrapped cost-effect pairs were plotted on a cost-effectiveness plane. Cost-effectiveness acceptability curves (CEACs) were estimated, showing the probability that the intervention was cost-effective compared to usual care at different ceiling ratios (38). A ceiling ratio represents the maximum amount of money society is willing to pay to gain one unit of effect on the outcome measure. Effect differences for the PHQ-15 and the HADS were multiplied by -1 to enhance interpretability of results. All analyses were performed with Stata SE/14.

Sensitivity analyses

In order to assess the robustness of our results, we conducted multiple sensitivity analyses. First, we performed the economic evaluation from a healthcare perspective, which is the recommended perspective in countries such as the United Kingdom (39). Within the healthcare perspective, only healthcare costs are taken into account. Because the effectiveness of the intervention was particularly pronounced in patients with symptoms below the median duration of symptoms and in patients with fewer comorbid physical diseases (15), we also performed subgroup analyses for patients with a duration of symptoms below and above the median duration of MUPS complaints, and with either 0-2 or 3 or more comorbid physical diseases.

RESULTS

Participants

An overview of patient enrolment, allocation and follow-up is provided in Appendix A. Recruitment took place between August 2015 and March 2017. Invitations were sent by mail to 1806 potential participants. Of these, 234 (13%) expressed an interest to participate in the study and fulfilled the criteria for USD. In total, 117 people were enrolled in the intervention group and 96 in the usual care group.

Patients' baseline characteristics are presented in Table 1. The mean age in both arms was 51.5 years (SD 16.3) and 74.5% of the whole sample were female. Clinically relevant differences were found for gender, education and neurological symptoms.

Table 1. Patients' baseline characteristics and outcomes

Characteristics	Intervention group (n=111)	Control group (n=89)	Total sample (n=200) [a]
Age, mean (sd)	53.00 (15.47)	49.69 (17.13)	51.53 (16.3)
Female	78 (70.3%)	71 (79.8%)	149 (74.5%)
Both parents born in the Netherlands	90 (81.1%)	72 (80.9%)	162 (81.0%)
Educational level			
Low	8 (7.3%)	7 (8.0%)	15 (7.6%)
Medium	58 (52.3%)	54 (62.1%)	112 (56.9%)
High education	44 (40.0%)	26 (29.8%)	70 (35.5%)
Work status*			
Employed	43 (38.7%)	36 (40.4.%)	79 (39.5%)
Unemployed	68 (61.3%)	53 (59.6%)	121 (60.5%)
Living situation			
Alone	28 (25.2%)	23 (25.8%)	51 (25.5%)
Not alone	83 (74.8%)	66 (74.2%)	149 (74.5%)
Symptom duration in years (self-report), mean (sd)	10.56 (11.28)	11.46 (12.79)	10.96 (11.95)
Most prominent symptoms*			
Musculoskeletal	78 (70.3%)	66 (74.2%)	144 (72.0%)
General and unspecified	41 (36.9%)	36 (40.4%)	77 (38.5%)
Neurological	39 (35.1%)	16 (18.0%)	55 (27.5%)
Psychological	19 (17.1%)	18 (20.2%)	37 (18.5%)
Digestive	11 (9.9%)	7 (7.9%)	18 (9.0%)
Number of somatic comorbidities, mean (sd)	3.16 (2.50)	3.34 (2.41)	3.24 (2.46)
Most reported somatic comorbidities*			
Back problems	79 (71.2%)	63 (70.8%)	142 (71.0%)
Pulmonary	40 (36.0%)	28 (31.5%)	68 (34.0%)
Neurological	35 (31.5%)	31 (34.8%)	66 (33.0%)
Number of self-report psychiatric comorbidities, mean (sd)	0.69 (0.91)	0.71 (1.19)	0.70 (1.04)
Most reported self-report psychiatric comorbidities*			
Distress/burn-out	27 (24.5%)	18 (20.9%)	45 (23.0%)
Depression	26 (23.4%)	17 (19.5%)	43 (21.7%)
Anxiety	19 (17.1%)	16 (17.4%)	34 (17.3%)
Utility (EQ-5D-5L (0-1)), mean (sd)	0.62 (0.22)	0.57 (0.26)	0.60 (0.24)
RAND-36 PCS (primary outcome), mean (sd)	50.22 (9.89)	49.64 (9.81)	49.97 (9.83)

Table 1. Continued

Characteristics	Intervention group (n=111)	Control group (n=89)	Total sample (n=200) [a]
Anxiety (HADS-A) , mean (sd)	7.89 (3.80)	7.69 (4.52)	7.80 (4.11)
Depression (HADS-D), mean (sd)	7.00 (3.86)	7.52 (4.23)	7.22 (4.02)
Somatic symptom severity, PHQ-15, mean (sd)	13.63 (4.89)	13.47 (4.43)	13.56 (4.69)

Results are expressed as *n* (%) unless stated otherwise. Abbreviations: HADS: Hospital Anxiety and Depression Scale; MCS: Mental Component Summary Score; MUS: medically unexplained symptoms; PCS: Physical Component Summary Score; PHQ-15: Patient Health Questionnaire 15-item somatization scale; sd: standard deviation

[a] Due to missing values the available *n* ranged from 188-200.

* More than one answer option was permitted, so numbers do not necessarily add up to 100%

Costs and effects

Table 2 provides the mean costs and effects over 12 months. Total societal costs in the intervention group were lower (mean difference -€2300, 95% CI -3257 to -134) than in the usual care group and this difference was statistically significant. The main contributors to the cost differences between the two groups were paid and unpaid help, primary care costs and productivity losses in the form of absenteeism.

The mean number of QALYs (primary outcome) was 0.66 (SE 0.01) in the intervention group and 0.65 (SE 0.02) in the usual care group. This difference (0.01) was not statistically significant (95% CI -0.01 to 0.04). The mean PCS was 2.46 (95% CI 1.44 to 3.47) points higher in the intervention group than in the usual care group after 12 months. The mean PHQ-15 score was 0.26 (95% CI -0.28 to 0.81) points higher, and the mean HADS score was 0.07 (95% CI -0.67 to 0.81) points higher in the intervention group than in the usual care group after 12 months. Both differences were not statistically significant.

Cost-effectiveness analysis

Primary outcome

The results of the cost utility analysis (CUA) and cost effectiveness analysis (CEA) are presented in Table 3. The ICER for QALYs was -149,775, which indicates that a gain in 1 QALY is associated with cost savings of €149,775 in the intervention group as compared to usual care. The CE-plane for QALYs (Figure 1a) shows that the majority of the bootstrapped cost-effect pairs (66%) is located in the southeast quadrant (intervention dominant over usual care, i.e. less expensive and more effective). About a quarter (26%) of the cost-effect pairs is located in the southwest quadrant (less effective, less costly). At a willingness to pay of 0 €/QALY gained, the probability that the intervention is cost-effective compared to usual care was 0.93 (Figure 1b). The CEAC is a decreasing function

of willingness to pay, because costs in the intervention group were lower than in the usual care group.

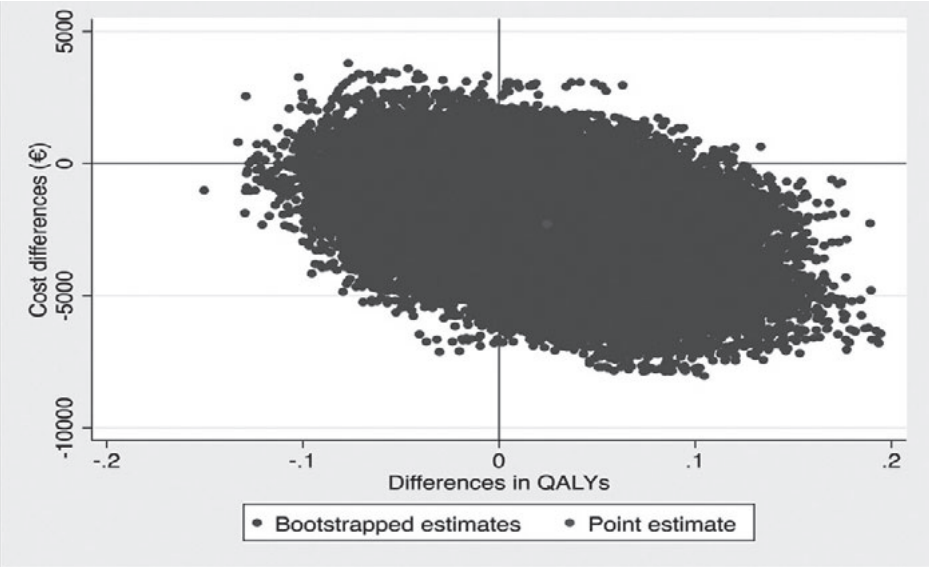


Figure 1a. Cost-effectiveness plane QALYs
Abbreviation: QALYs: quality adjusted life years

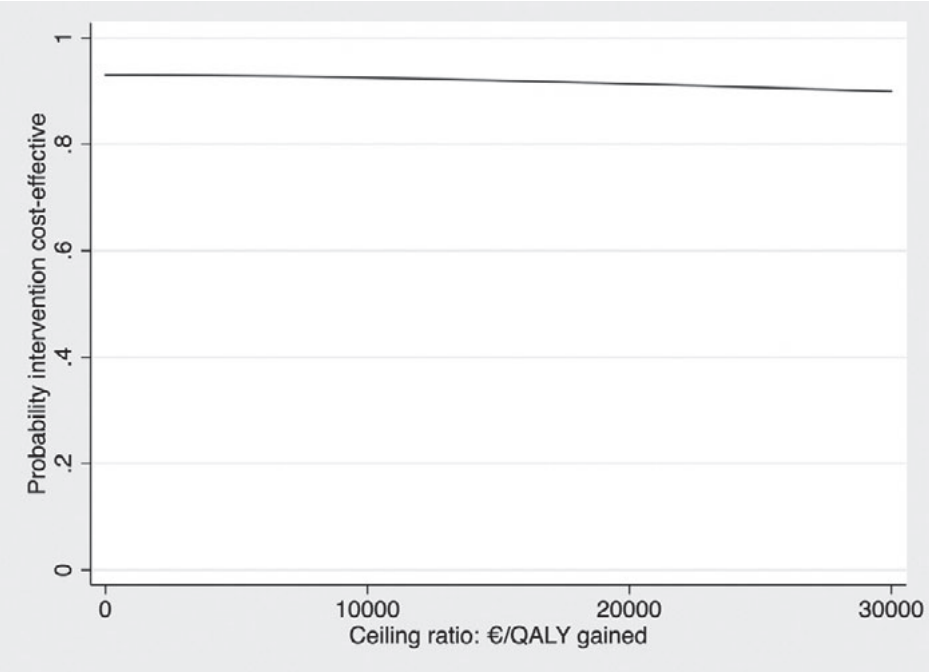


Figure 1b. Cost-effectiveness acceptability curve QALYs
Abbreviation: QALYs: quality adjusted life years

Secondary outcomes

The ICER for the PCS was -934, which indicates that one point of improvement on the PCS is associated with cost savings of €934 in the intervention group compared to usual care. The CE-plane for PCS (Figure 2a) shows that the vast majority of the bootstrapped cost-effect pairs (91%) is located in the southeast quadrant (intervention dominant over usual care). At a willingness to pay of €0, the probability of the intervention being cost-effective is 0.92 (Figure 2b).

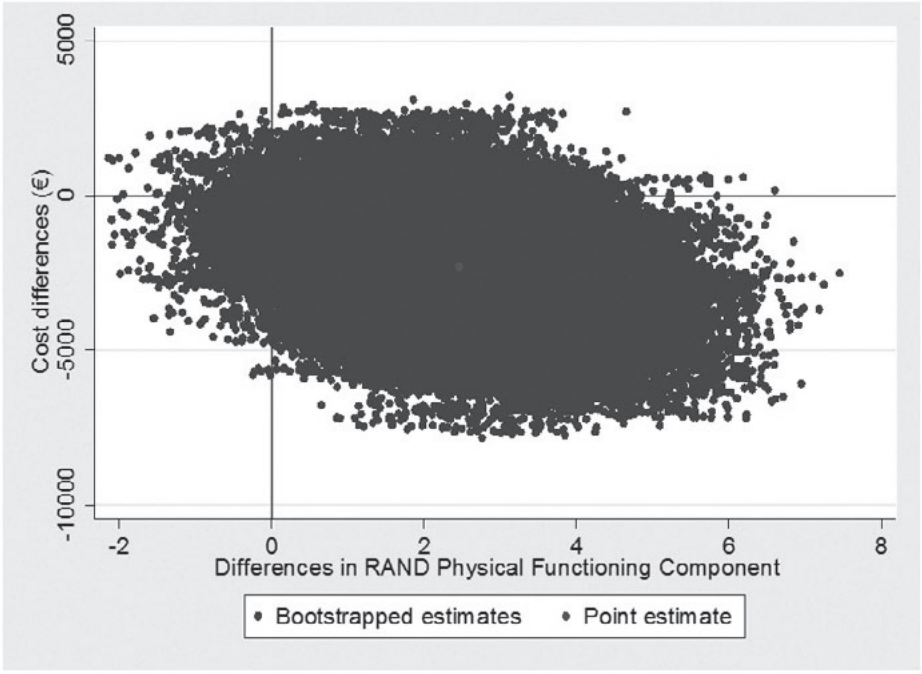


Figure 2a. Cost-effectiveness plane RAND-36 Physical Component Summary Score (PCS)
Abbreviation: PCS: Physical Component Summary Score

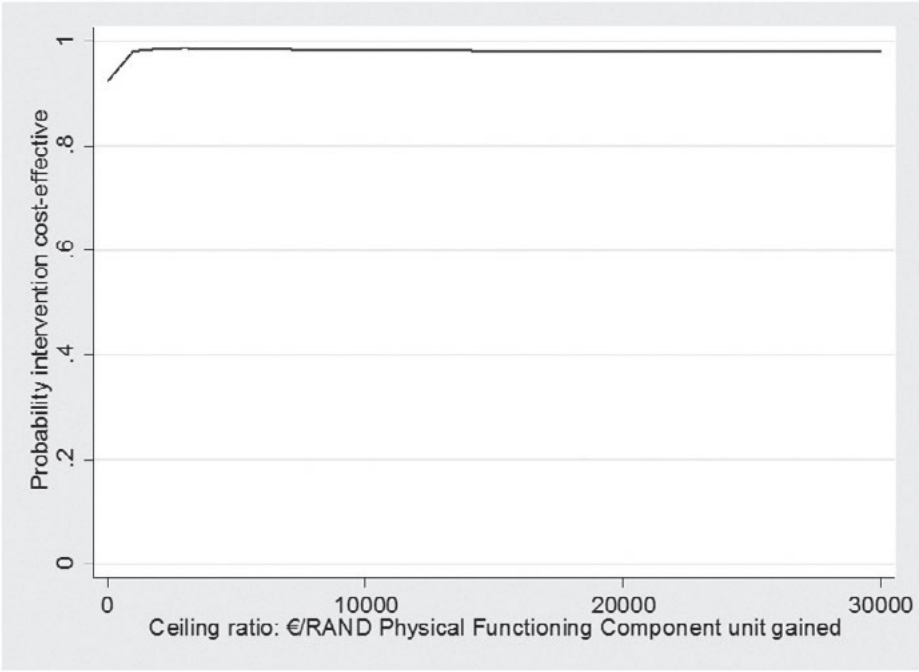


Figure 2b. Cost-effectiveness acceptability curve RAND-36 Physical Component Summary Score (PCS)
Abbreviation: PCS: Physical Component Summary Score

The ICERs for the PHQ-15 and HADS (PHQ-15: 8,708 and HADS: 32,427) indicate that the intervention did not significantly improve these outcomes, but saved money compared to usual care. The CE-planes and CEACs for these variables can be found in Appendices B and C. The majority of the bootstrapped cost-effect pairs (PHQ-15: 62% and HADS: 49%) were located in the southwest quadrant (less effective, less costly), but a substantial percentage (PHQ-15: 31% and HADS: 44%) were also located in the southeast quadrant (more effective, less costly). At a willingness to pay of €0/unit of effect, the probability of the intervention being cost-effective compared to usual care was 0.92 for both the PHQ-15 and the HADS (Appendices B and C).

Sensitivity analyses

Healthcare perspective

The results of the analyses from the healthcare perspective are similar to the results from the societal perspective analysis (Tables 2 and 3, and Appendices D). For all outcome measures, the probability of the intervention being cost-effective compared to usual care is 0.78 at a willingness-to-pay of 0 €/incremental unit of effect. The probability is

lower than from the societal perspective (0.92), because the cost difference is smaller than from the societal perspective.

Table 2. Mean costs (€) and outcomes over 12-month follow-up in the intervention and control group

	Intervention group (mean, SE) (n=117)	**Control group (mean, SE) (n=96)**	**Difference**	**95% CI**
Clinical outcomes*				
Primary outcome				
QALYs (range 0-1)	0.66 (0.01)	0.65 (0.02)	0.01	-0.01 to 0.04
Secondary outcomes				
RAND-36 PCS	51.25 (0.85)	48.79 (0.87)	2.46	1.44 to 3.47
Symptom severity (PHQ-15)	12.66 (0.44)	12.40 (0.41)	0.26	-1.48 to 0.95
Depressive and anxiety symptoms (HADS)	13.37 (0.61)	13.30 (0.75)	0.07	-0.86 to 0.72
Annual costs (€)				
Intervention costs	289 (0)	0 (0)	289	289 to 289
Primary care	1093 (128)	1617 (478)	-524	-2387 to 79
Secondary care	602 (440)	792 (297)	-190	-1551 to 641
Diagnostics	415 (50)	539 (131)	-124	-577 to 63
Medication	442 (118)	647 (196)	-205	-723 to 166
Alternative medicine	18 (6)	9 (4)	9	-5 to 27
Productivity losses - Absenteeism	484 (135)	1005 (342)	-521	-1523 to 36
Productivity losses - Presenteeism	678 (245)	657 (187)	21	-500 to 729
Paid or unpaid help	2963 (588)	4017 (826)	-1054	-3198 to 749
Total costs:				
Societal perspective	**6987 (862)**	**9287 (1326)**	**-2300**	**-3257 to -1342**
Healthcare perspective	**2841 (404)**	**3595 (850)**	**-754**	**-1365 to -143**

Abbreviations: 95% CI: 95% confidence interval; HADS: Hospital Anxiety and Depression Scale; SE: standard error; PCS: Physical Component Summary Score; PHQ-15: Patient Health Questionnaire 15-item symptom severity scale; QALYs: quality adjusted life years

* The means provided for the clinical outcomes are pooled estimates

Table 3. Difference in costs and effects, ICER and distribution on the CE-plane

Outcomes	Costs Δ (95%CI)	Effect Δ (95%CI)	ICER	Distribution CE-plane (%)			
				NE	SE	SW	NW
ITT analysis (societal perspective)							
QALYs (Range: 0 - 1)	-2300 (-3257 to -1342)	0.01 (-0.01 to 0.04)	-149,775	3	66	26	5
RAND-36 PCS	-2300 (-3257 to -1342)	2.46 (1.44 to 3.47)	-934	8	91	1	0
PHQ-15	-2300 (-3257 to -1342)	-0.26 (-1.48 to 0.95)	8,708	1	31	62	6
HADS	-2300 (-3257 to -1342)	-0.07 (-0.86 to 0.72)	32,427	2	44	49	5
ITT analysis (healthcare perspective)							
QALYs (Range: 0 - 1)	-754 (-1365 to -143)	0.01 (-0.01 to 0.04)	-49,151	12	58	20	10
RAND-36 PCS	-754 (-1365 to -143)	2.46 (1.44 to 3.47)	-132	21	77	1	1
PHQ-15	-754 (-1365 to -143)	-0.26 (-1.48 to 0.95)	2,857	4	28	50	18
HADS	-754 (-1365 to -143)	-0.07 (-0.86 to 0.72)	10,641	7	39	39	15

Abbreviations: 95% CI: 95% confidence interval; CE-plane: cost-effectiveness plane; HADS: Hospital Anxiety and Depression Scale; ICER: incremental cost effectiveness ratio; ITT: intention-to-treat; NE: northeast; NW: northwest; PCS: Physical Component Summary Score; PHQ-15: Patient Health Questionnaire 15-item symptom severity scale; QALYs: quality adjusted life years; SE: southeast; SW: southwest

PHQ-15 and HADS differences in effect were multiplied by -1 to make interpretation easier.

Subgroup analyses

The results of the subgroup analyses are presented in Table 4 and Appendix E. For the PCS, the intervention was significantly more effective than usual care in the subgroups with shorter symptom duration and less comorbid diseases. All of the differences in the other outcomes (for QALYs, PHQ-15 and HADS) were not statistically significant. Total societal costs in the intervention group were significantly lower than in the usual care group in the subgroups with longer symptom duration and less comorbid diseases. In the opposite subgroups (shorter symptom duration and more comorbid diseases), the difference in total societal costs was much smaller and not statistically significant. As a consequence, due to the significantly lower societal costs, the intervention was considered dominant over usual care in the subgroups with shorter symptom duration and more comorbid diseases, but less so in the other subgroups.

DISCUSSION

Main findings

We investigated whether a cognitive behavioural intervention for patients with MUPS was cost-effective compared to current usual care in the Netherlands. Total mean healthcare and societal costs were significantly lower in the intervention group compared to the usual care group. Although the difference in QALYs was in favour of the intervention group, this difference was not statistically significant. At a willingness to pay of 0 € per QALY gained, the probability of the intervention being cost-effective was 0.93 from the societal perspective. Therefore, we can consider our intervention to be dominant over usual care.

Even though the difference in the primary outcome, the QALYs, was not statistically significant, the PCS score in the intervention group was significantly higher (meaning that physical functioning was significantly more improved) than in the usual care group. Furthermore, costs were significantly lower in the subgroups with a shorter symptom duration and fewer comorbid diseases as compared to the subgroups with a longer symptom duration and many comorbid diseases. Finally, the intervention was cost-effective from both the societal and health care perspective.

In both groups, the main contributor to the difference in societal costs were costs of paid and unpaid help. This can probably be explained by the fact that patients with MUPS often experience severe physical limitations due to their symptoms, and are therefore not able to carry out daily tasks on their own (40). The improvement in PCS is thus reflected by an improvement in the ability to carry out these daily tasks. This is also indicated

Table 4. Difference in costs and effects, ICER and distribution on the CE-plane for subgroup analyses

Outcomes		Costs Δ (95%CI)	Effect Δ (95%CI)	ICER	Distribution CE-plane (%)			
ITT analysis (societal perspective)	**Subgroup**				**NE**	**SE**	**SW**	**NW**
QALYs (Range: 0 - 1)	Symptom duration below median (0-2092 days) [a]	-351 (-2223 to 1521)	0.04 (-0.01 to 0.08)	-10,373	0.29	0.49	0.06	0.16
	Symptom duration above median (2093+ days) [b]	-3983 (-5858 to 2107)	0.00 (-0.04 to 0.03)	681,428	0.01	0.44	0.53	0.02
	0-2 comorbid physical diseases [c]	-3265 (-5017 to -1514)	0.00 (-0.04 to 0.05)	-791,377	0.02	0.52	0.43	0.03
	3 or more comorbid physical diseases [d]	-738 (-2802 to 1325)	0.01 (-0.05 to 0.08)	-170.077	0.14	0.39	0.24	0.23
RAND-36 PCS	Symptom duration below median (0-2092 days) [a]	-351 (-2223 to 1521)	3.67 (1.92 to 5.43)	-95	0.44	0.54	0.00	0.02
	Symptom duration above median (2093+ days) [b]	-3983 (-5858 to 2107)	0.99 (-0.59 to 2.57)	-4,029	0.02	0.71	0.25	0.02
	0-2 comorbid physical diseases [c]	-3265 (-5017 to -1514)	3.10 (1.39 to 4.82)	-1,051	0.05	0.93	0.02	0.00
	3 or more comorbid physical diseases [d]	-738 (-2802 to 1325)	0.58 (-0.93 to 2.09)	-1.272	0.19	0.44	0.19	0.18
PHQ-15	Symptom duration below median (0-2092 days) [a]	-351 (-2223 to 1521)	0.14 (-0.98 to 1.26)	-2,506	0.20	0.36	0.19	0.25
	Symptom duration above median (2093+ days) [b]	-3983 (-5858 to 2107)	-0.70 (-3.71 to 2.31)	5,652	0.00	0.19	0.78	0.03
	0-2 comorbid physical diseases [c]	-3265 (-5017 to -1514)	0.11 (-0.87 to 1.10)	-28,705	0.02	0.53	0.42	0.03
	3 or more comorbid physical diseases [d]	-738 (-2802 to 1325)	-1.15 (-5.98 to 3.67)	639	0.02	0.07	0.56	0.35

Table 4. Continued

Outcomes		Costs Δ (95%CI)	Effect Δ (95%CI)	ICER	Distribution CE-plane (%)			
ITT analysis (societal perspective)	**Subgroup**				**NE**	**SE**	**SW**	**NW**
HADS	Symptom duration below median (0-2092 days) [a]	-351 (-2223 to 1521)	0.83 (-2.85 to 4.51)	-422	0.27	0.45	0.10	0.18
	Symptom duration above median (2093+ days) [b]	-3983 (-5858 to 2107)	-0.96 (-5.15 to 3.23)	4135	0.00	0.23	0.74	0.03
	0-2 comorbid physical diseases [c]	-3265 (-5017 to -1514)	0.36 (-1.61 to 2.34)	-8,946	0.03	0.59	0.36	0.02
	3 or more comorbid physical diseases [d]	-738 (-2802 to 1325)	-0.78 (-4.34 to 2.77)	940	0.08	0.21	0.40	0.30

Abbreviations: 95% CI: 95% confidence interval; CE-plane: cost-effectiveness plane; HADS: Hospital Anxiety and Depression Scale; ICER: incremental cost effectiveness ratio; NE: northeast; NW: northwest; PCS: Physical Component Summary Score; PHQ-15: Patient Health Questionnaire 15-item symptom severity scale; QALYs: quality adjusted life years; SE: southeast; SW: southwest

[a] n=45 (usual care group), n=60 (intervention group)

[b] n=51 (usual care group), n=57(intervention group)

[c] n=43 (usual care group), n=64 (intervention group)

[d] n=53 (usual care group), n=53 (intervention group)

by the much lower absenteeism costs in the intervention group, which are nearly half the absenteeism costs in the usual care group. This suggests that the intervention also enables patients to become more productive at work. In this study, primary care costs were higher than secondary care costs in both groups, although secondary care services are typically more expensive than primary care. Overall, we observed substantial reductions in healthcare costs, which indicates that the intervention is also beneficial from the healthcare perspective.

Comparison to previous studies

Our findings confirm the results from several previous studies evaluating the cost-effectiveness of psychological interventions for patients with MUPS. A group CBT intervention for somatoform disorders also was dominant (less expensive and more effective) compared to a waiting-list control group (41). The findings from our study may be somewhat more robust as we compared our intervention against usual care, which may be considered an active treatment, as opposed to a waiting-list control group. A group intervention, based on CBT and psychodynamic therapy, provided by a GP in collaboration with a 'psychosomatic specialist' (physician or psychologist), on top of enhanced medical care, was more effective than enhanced medical care alone, for patients with functional somatic syndromes (42). However, it did not lead to a statistically significant decrease in costs. Finally, two studies investigating the cost-utility of interventions consisting of psycho-educational and CBT techniques for patients with fibromyalgia, who tend to have comparable complaints to MUPS patients, showed that these interventions were cost-effective (lower costs and more effective) compared to usual care (43, 44).

Strengths and limitations

A first strength of this study is that it was designed as a pragmatic trial, thereby mirroring daily practice as much as possible. In addition, we used a societal perspective that included a broad range of costs making it possible to identify potential cost shifts between sectors. Another strength is that we evaluated the impact of the intervention on multiple outcomes measures, such as QALYs, physical functioning (RAND-36 PCS), somatic symptom severity (PHQ-15) and anxiety and depression (HADS). Finally, we carried out several sensitivity and subgroup analyses.

A limitation of the current analysis is that the original power calculation was not based on the EQ-5D-5L or costs, but on the PCS of the RAND-36, which was the primary outcome in the effectiveness trial. However, it may be considered unethical to include more patients than necessary to demonstrate clinical effectiveness (45). Despite this limitation we

were able to demonstrate a significant effect on the costs of the intervention. The observed difference in QALYs based on the EQ-5D-5L was only 0.01 which is rather small, while cost differences were rather large (-€2300). The EQ-5D-5L may not have been sensitive enough to pick up the changes in our sample due to the absence of important dimensions of health, relevant for this population, such as relationships, energy and sleep (46). Another limitation is that the recall period of the questionnaire used to measure costs, the Tic-P, was 4 months. Patients reported that they found this rather long and had difficulty estimating the precise use of healthcare and medication in the previous period. This may have led to a less precise calculation of costs. However, we expect that this potential bias is present in both groups. Thus, this probably did not affect our estimations of the differences between the treatment groups. Moreover, there is evidence that recall up to 6 months is reliable (47).

A final point of consideration is that our intervention was designed and tested for and within the Dutch healthcare setting, where an MHNP is routinely available in general practice. Our findings may be less generalizable to countries with other, very different, healthcare systems or where resources are allocated in another way. Our findings may also be less generalisable to those MUPS populations that do not fulfil the DSM-IV criteria for undifferentiated somatoform disorder.

Implications for practice

Although the Dutch GP guideline for MUPS recommends GPs to refer patients with mild to moderate MUPS to a MHNP, this is still not common practice (48). Considering that the workload of the MHNP is already high, delivering our intervention to all eligible MUPS patients, on top of the MHNPs' current activities, means that deployment of MNHPs in general practice should be increased. The results of our economic evaluation show that in the long run, it would result in a decrease of costs in primary care, so that such an investment may be considered efficient. Moreover, our intervention provides GPs with an efficient and practical strategy for dealing with patients with MUPS.

CONCLUSION

Based on the current study, the cognitive behavioural intervention delivered by MHNPs, is considered dominant to usual care. Based on our findings, the intervention could be a valuable addition to usual care in general practice, and in particular for patients with a shorter duration of symptoms and few comorbid physical diseases. However, to implement the intervention on a wider scale, may mean that the deployment of MHNPs needs to be increased.

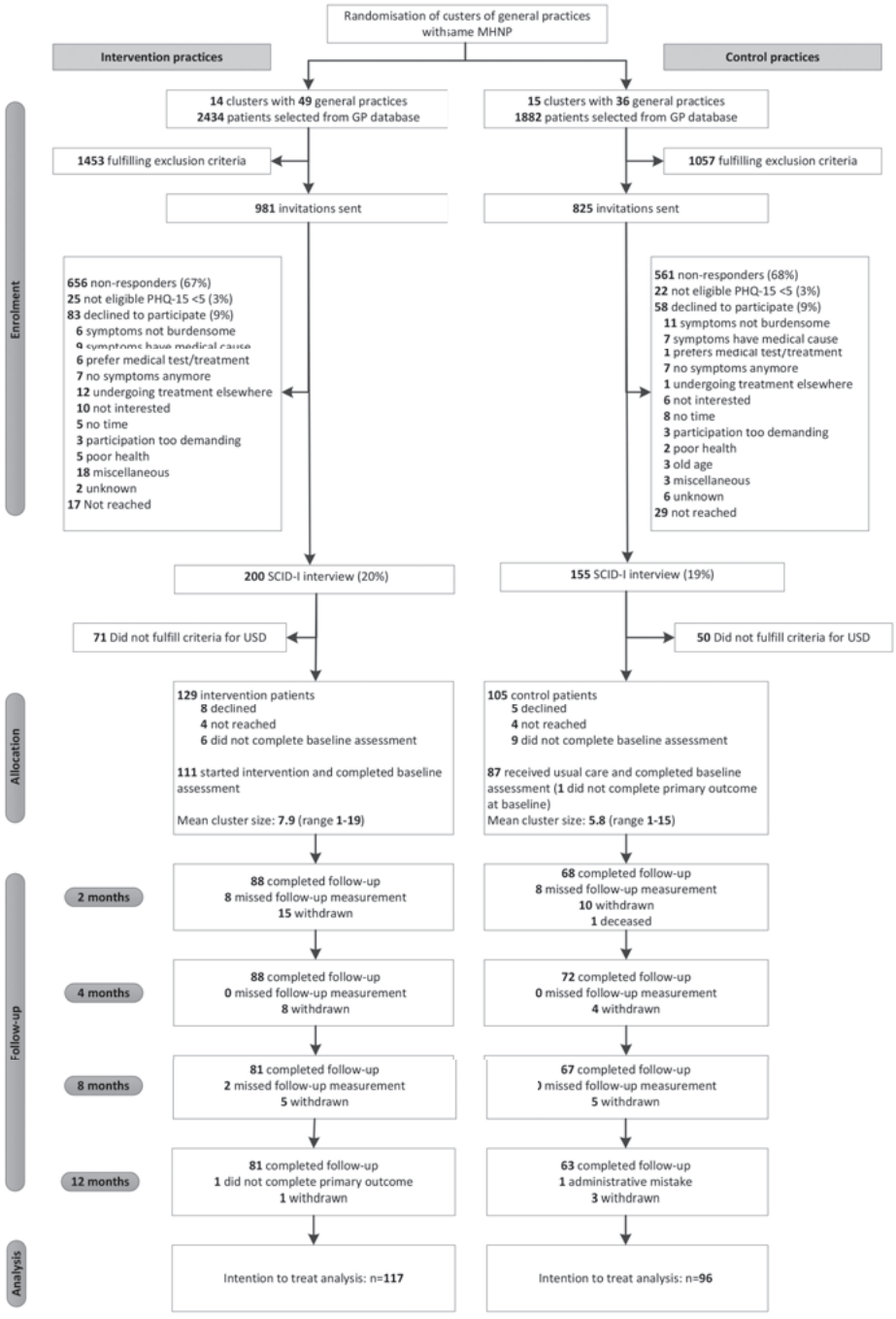
REFERENCES

1. De Waal MWM, Arnold IA, Eekhof JAH, Van Hemert AM. Somatoform disorders in general practice: prevalence, functional impairment and comorbidity with anxiety and depressive disorders. British Journal of Psychiatry. 2004;184(6):470-6.
2. Steibrecher N, Koerber S, Frieser D, Hiller W. The prevalence of medically unexplained symptoms in primary care. Psychosomatics. 2011;52:263-71.
3. Haftgoli N, Favrat B, Verdon F, Vaucher P, Bischoff T, Burnand B, et al. Patients presenting with somatic complaints in general practice: depression, anxiety and somatoform disorders are frequent and associated with psychosocial stressors. BMC Fam Pract. 2010;11(67).
4. APA. Diagnostic and statistical manual of mental disorders, 4th ed. (DSM-IV). Washington DC: American Psychiatric Association; 1994.
5. APA. Diagnostic and statistical manual of mental disorders, 5th ed. (DSM-5). Arlington, VA: American Psychiatric Publishing; 2013.
6. van Eck van der Sluijs J, Ten Have M, Rijnders C, van Marwijk H, de Graaf R, van der Feltz-Cornelis C. Medically unexplained and explained physical symptoms in the general population: association with prevalent and incident mental disorders. PLoS One. 2015;10(4):e0123274.
7. Konnopka A, Schaefert R, Heinrich S, Kaufmann C, Luppa M, Herzog W, et al. Economics of medically unexplained symptoms: a systematic review of the literature. Psychother Psychosom. 2012;81(5):265-75.
8. Gustavsson A, Svensson M, Jacobi F, Allgulander C, Alonso J, Beghi E, et al. Cost of disorders of the brain in Europe 2010. Eur Neuropsychopharmacol. 2011;21(10):718-79.
9. van Dessel N, den Boeft M, van der Wouden JC, Kleinstäuber M, Leone SS, Terluin B, et al. Non-pharmacological interventions for somatoform disorders and medically unexplained physical symptoms (MUPS) in adults Cochrane Database Syst Rev. 2014;11:CD01114.
10. Jing L, Gill NS, Teodorczuk A, Li ZJ, Sun J. The efficacy of cognitive behavioural therapy in somatoform disorders and medically unexplained physical symptoms: A meta-analysis of randomized controlled trials. J Affect Disord. 2018;245:98-112.
11. Henningsen P, Zipfel S, Sattel H, Creed F. Management of functional somatic syndromes and bodily distress. Psychother Psychosom. 2018;87(1):12-31.
12. Mack S, Jacobi F, Gerschler A, Strehle J, Hofler M, Busch MA, et al. Self-reported utilization of mental health services in the adult German population--evidence for unmet needs? Results of the DEGS1-Mental Health Module (DEGS1-MH). Int J Methods Psychiatr Res. 2014;23(3):289-303.
13. Olde Hartman T, Blankenstein A, Molenaar A, Bentz van den Berg D, van der Horst H, Arnold I, et al. [NHG guideline on medically unexplained symptoms (MUS)] (in Dutch). Huisarts Wet. 2013;56(5):222-30.
14. Magnee T, de Beurs DP, de Bakker DH, Verhaak PF. Consultations in general practices with and without mental health nurses: an observational study from 2010 to 2014. BMJ Open. 2016;6(7):e011579.
15. Sitnikova K, Leone SS, van Marwijk HWJ, Twisk J, van der Horst HE, van der Wouden JC. Effectiveness of a cognitive behavioural intervention for patients with undifferentiated somatoform disorder: Results from the CIPRUS cluster randomized controlled trial in primary care. J Psychosom Res. 2019:109745.
16. Wortman MSH, Lokkerbol J, van der Wouden JC, Visser B, van der Horst HE, Olde Hartman TC. Cost-effectiveness of interventions for medically unexplained symptoms: A systematic review. PLoS One. 2018;13(10):e0205278.
17. Sitnikova K, Leone SS, Zonneveld LNL, van Marwijk HWJ, Bosmans JE, van der Wouden JC, et al. The CIPRUS study, a nurse-led psychological treatment for patients with undifferentiated somatoform disorder in primary care: study protocol for a randomised controlled trial. Trials. 2017;18(1):206.
18. Robbins J, Kirmayer L, Hemami S. Latent variable models of functional

somatic distress. J Nerv Ment Dis. 1997;185(10):606-15.
19. Kroenke K, Spitzer RL, Williams JBW. The PHQ-15: validity of a new measure for evaluating the severity of somatic symptoms. Psychosom Med. 2002;64:258-66.
20. First M, Spitzer R, Gibbon M, Williams J. Structured Clinical Interview for DSM-IV-TR Axis I Disorders, Research Version, Patient Edition (SCID-I/P). New York: Biometrics Research: New York State Psychiatric Institute; 2002.
21. Zonneveld LN, van Rood YR, Timman R, Kooiman CG, Van't Spijker A, Busschbach JJ. Effective group training for patients with unexplained physical symptoms: a randomized controlled trial with a non-randomized one-year follow-up. PLoS One. 2012;7(8):e42629.
22. Mynors-Wallis L. Problem-solving treatment for anxiety and depression: a practical guide. Oxford and New York: Oxford University Press; 2005.
23. Trimbos instituut. [Multidisciplinary guideline for MUPS and somatoform disorders] (In Dutch). 2011.
24. Hakkaart-van Roijen L, van Straten A, Donker M, Tiemens B. Manual Trimbos/IMTA questionnaire for costs associated with psychiatric illness (TIC-P). Rotterdam: Rotterdam: Institute for Medical Technology Assessment, Erasmus Univeristy; 2002.
25. Hakkaart-van Roijen L, Tan SS, Bouwmans CAM. [Manual for costs-research: methods and standard cost prices for economic evaluations in healthcare] (in Dutch). Diemen: College voor Zorgverzekeringen; 2010.
26. Z-index. G-Standaard. Den Haag, 2010 [Available from: https://www.z-index.nl/g-standaard.
27. Oostenbrink J, Bouwmans CAM, Koopmanschap M, Rutten F. Guideline for economic evaluations in healthcare: Institute for Medical Technology Assessment; 2004.
28. Brazier J, Ratcliffe J, Salomon JA, Tsuchiya A. Measuring and valuing health benefits for economic evaluation. Second ed. Oxford: Oxford University Press; 2017.
29. Herdman M, Gudex C, Lloyd A, Janssen M, Kind P, Parkin D, et al. Development and preliminary testing of the new five-level version of EQ-5D (EQ-5D-5L). Qual Life Res. 2011;20(10):1727-36.
30. Devlin NJ, Brooks R. EQ-5D and the EuroQol group: past, present and future. Appl Health Econ Health Policy. 2017;15(2):127-37.
31. Finch AP, Brazier JE, Mukuria C. What is the evidence for the performance of generic preference-based measures? A systematic overview of reviews. Eur J Health Econ. 2018;19(4):557-70.
32. Versteegh MM, Vermeulen KM, Evers SMAA, de Wit GA, Prenger R, Stolk EA. Dutch tariff for the five-level version of EQ-5D. Value Health. 2016;19(4):343-52.
33. van der Zee KI, Sanderman R. [Measuring general health with the RAND-36, a guideline] (in Dutch). 2nd ed. Groningen: Noordelijk Centrum voor Gezondheidsvraagstukken, NCG; 2012.
34. Zigmond A, Snaith R. The Hospital Anxiety and Depression Scale. Acta Psychiatr Scand. 1983;67(6):361–70.
35. Cambpell MK, Fayers PM, Grimshaw JM. Determinants of the intracluster correlation coefficient in cluster randomized trials: the case of implementation research. Clin Trials. 2005;2:99-107.
36. Sterne JAC, White IR, Carlin JB, Spratt M, Royston P, Kenward MG, et al. Multiple imputation for missing data in epidemiological and clinical research: potential and pitfalls. BMJ. 2009;338(jun29 1):b2393-b.
37. Rubin DB. Multiple imputation for nonresponse in surveys. New York: Wiley; 1987.
38. Fenwick E, O'Brien BJ, Briggs A. Cost-effectiveness acceptability curves--facts, fallacies and frequently asked questions. Health Econ. 2004;13(5):405-15.
39. National Institute for Health and Care Excellence (NICE). Developing NICE guidelines: the manual. 2014.
40. Joustra ML, Janssens KA, Bultmann U, Rosmalen JG. Functional limitations in functional somatic syndromes and well-defined medical diseases. Results from the general population cohort LifeLines. J Psychosom Res. 2015;79(2):94-9.
41. Visser MS, Zonneveld LN, Van't Spijker

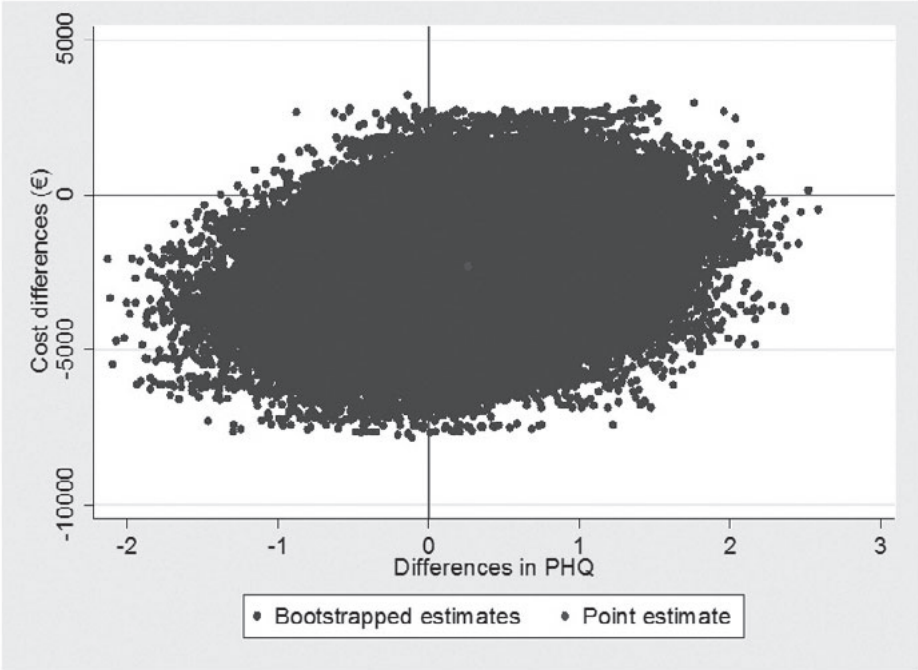
A, Hunink MG, Busschbach JJ. The cost-effectiveness of cognitive-behavioral group training for patients with unexplained physical symptoms. Value Health. 2015;18(5):570-7.
42. Konnopka A, Konig HH, Kaufmann C, Egger N, Wild B, Szecsenyi J, et al. Cost-utility of a specific collaborative group intervention for patients with functional somatic syndromes. J Psychosom Res. 2016;90:43-50.
43. Luciano J, Sabes-Figuera R, Cardenosa E, Penarrubia-Maria MT, Fernandez-Vergel R, Garcia-Campayo J, et al. Cost-utility of a psychoeducational intervention in fibromyalgia patients compared with usual care an economic evaluation alongside a 12-month randomized controlled trial. Clin J Pain. 2013;29:702-11.
44. Luciano J, D'Amico F, Cerda-Lafont M, Penarrubia-Maria MT, Knapp M, Cuesta-Vargas AI, et al. Cost-utility of cognitive behavioral therapy versus U.S. Food and Drug Administration recommended drugs and usual care in the treatment of patients with fibromyalgia: an economic evaluation alongside a 6-month randomized controlled trial. Arthritis Research & Therapy. 2014;16:451.
45. Briggs A. Economic evaluation and clinical trials: size matters. BMJ. 2000;321:1362-3.
46. Finch AP, Brazier JE, Mukuria C, Bjorner JB. An exploratory study on using principal-component analysis and confirmatory factor analysis to identify bolt-on dimensions: the EQ-5D case study. Value Health. 2017;20(10):1362-75.
47. van den Brink M, van den Hout WB, Stiggelbout AM, Putter H, van de Velde CJ, Kievit J. Self-reports of health-care utilization: diary or questionnaire? Int J Technol Assess Healthcare. 2005;21(3):298-304.
48. Sitnikova K, Pret-Oskam R, Dijkstra-Kersten SMA, Leone SS, van Marwijk HWJ, van der Horst HE, et al. Management of patients with persistent medically unexplained symptoms: a descriptive study. BMC Fam Pract. 2018;19(1):88.

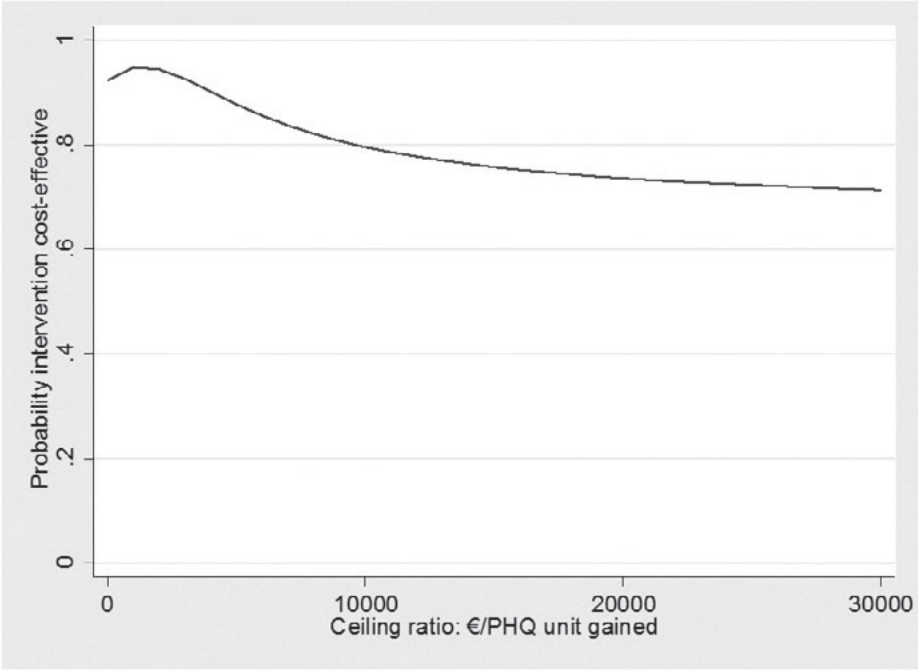
Appendix A. Flow of study participants



GP general practitioner; *MHNP* mental health nurse practitioner; *PHQ-15* Patient Health Questionnaire 15-item somatic symptom scale; *SCID-I* Structured Clinical Interview for DSM-IV Axis I Disorders

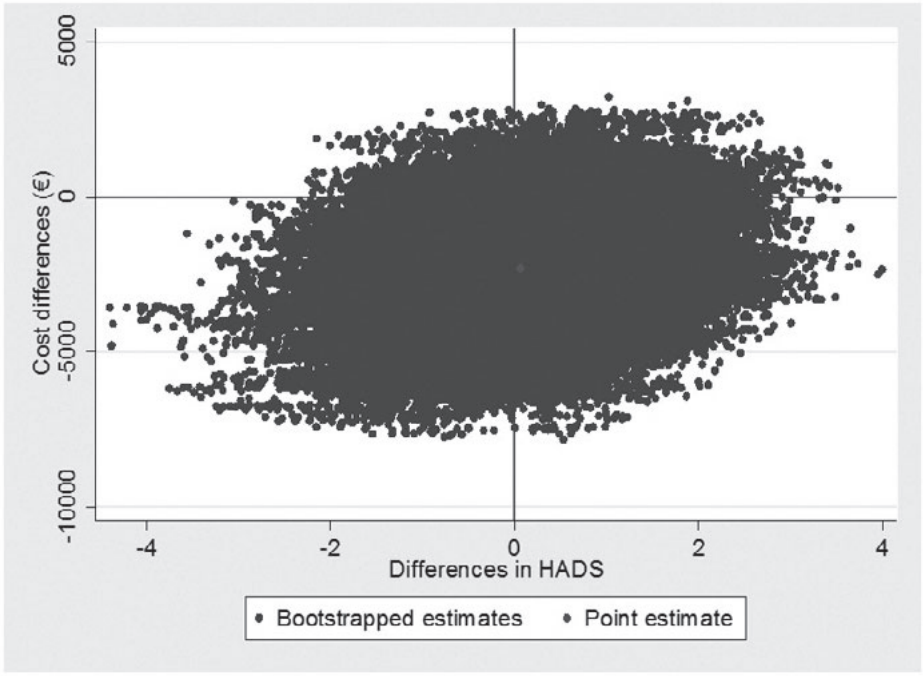
Appendix B. Cost-effectiveness plane and acceptability curve for PHQ-15 somatic symptom severity scale

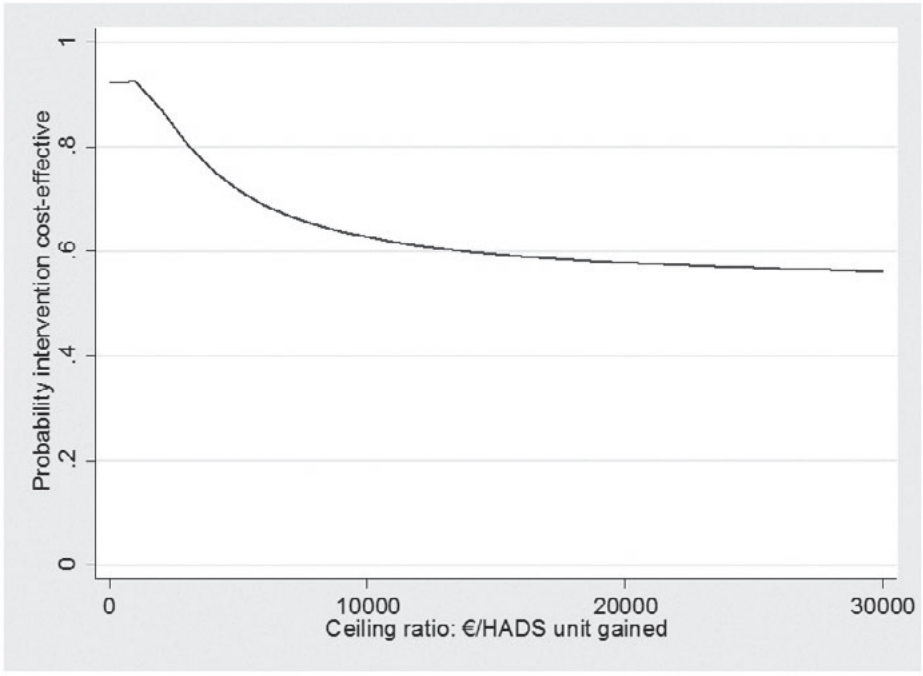




Abbreviation: PHQ-15: Patient Health Questionnaire 15-item somatization scale

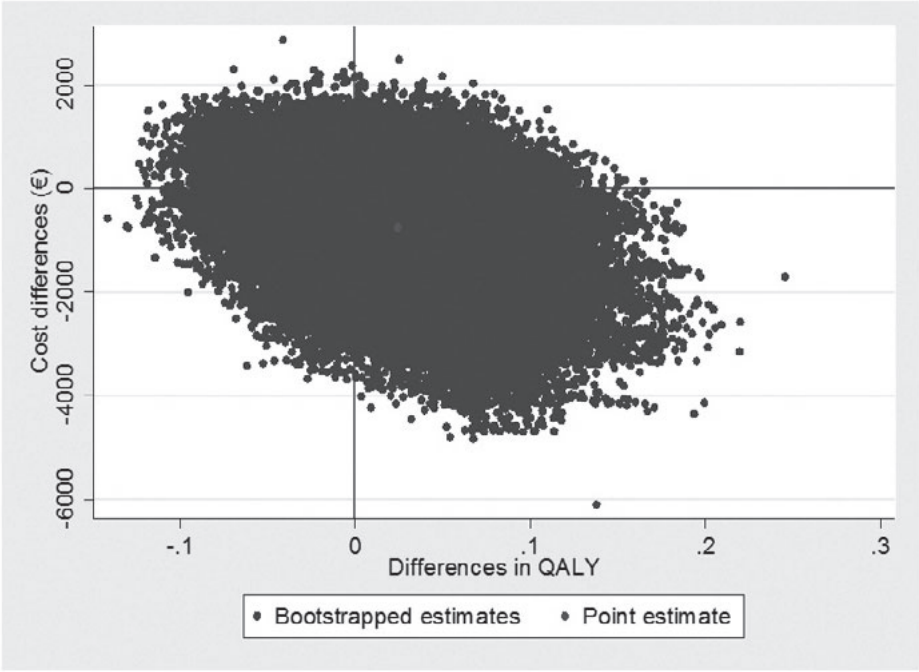
Appendix C. Cost-effectiveness plane and acceptability curve for HADS

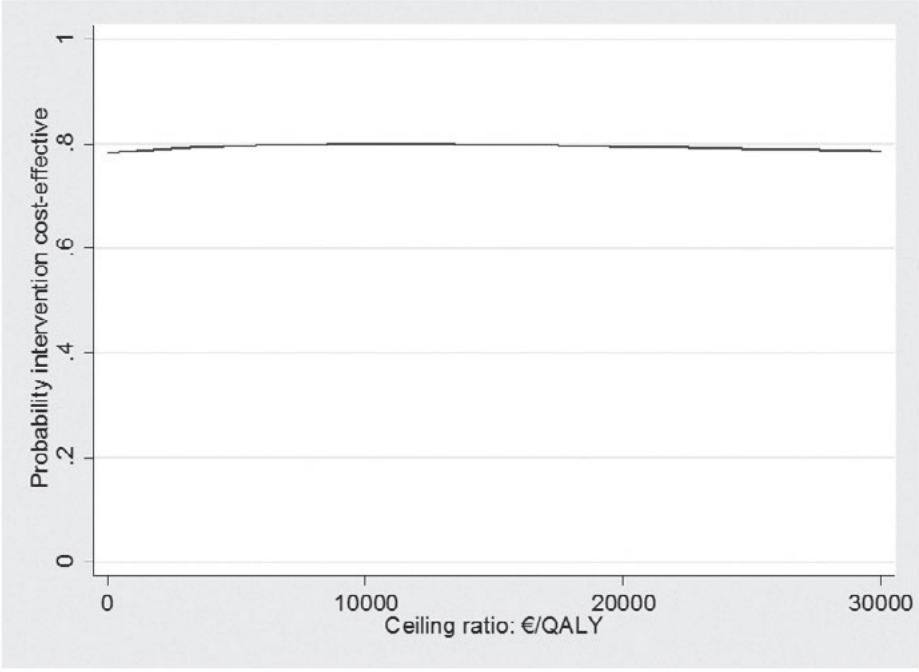




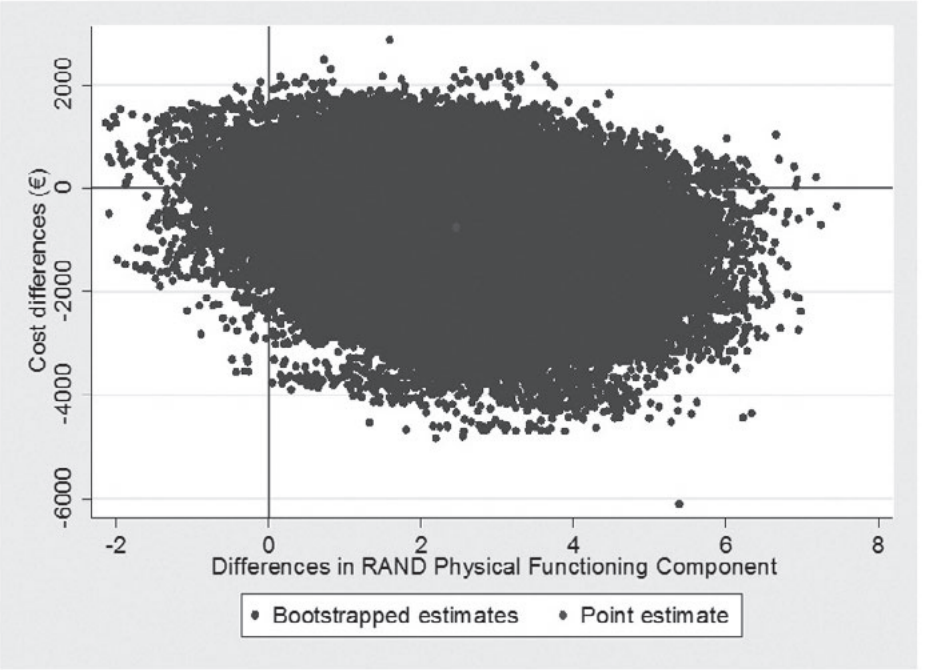
Abbreviation: HADS: Hospital Anxiety and Depression Scale

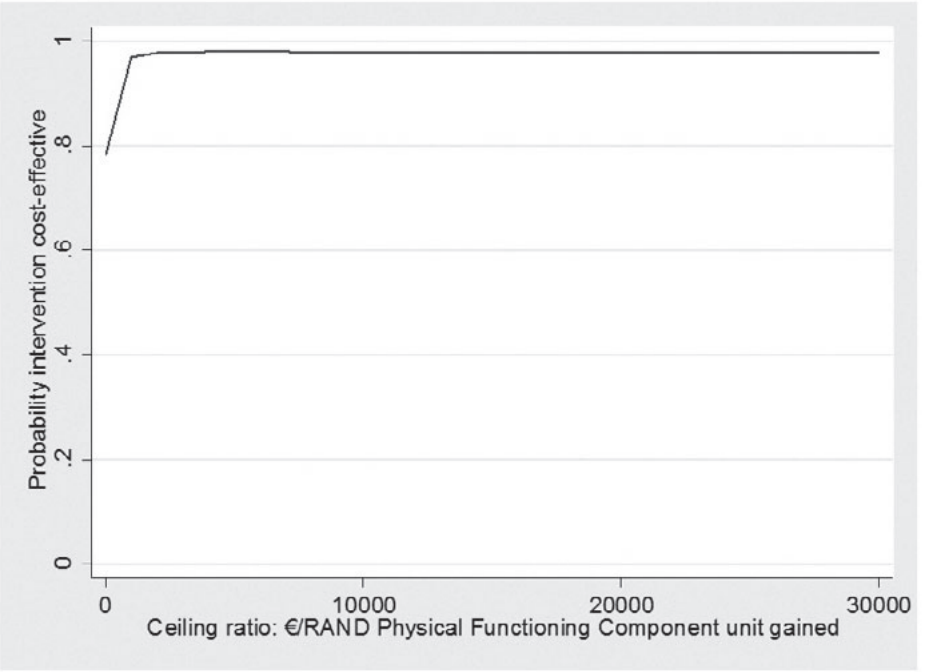
Appendix D. Results of the analyses from the healthcare perspective
Cost-effectiveness plane and acceptability curve for QALYs



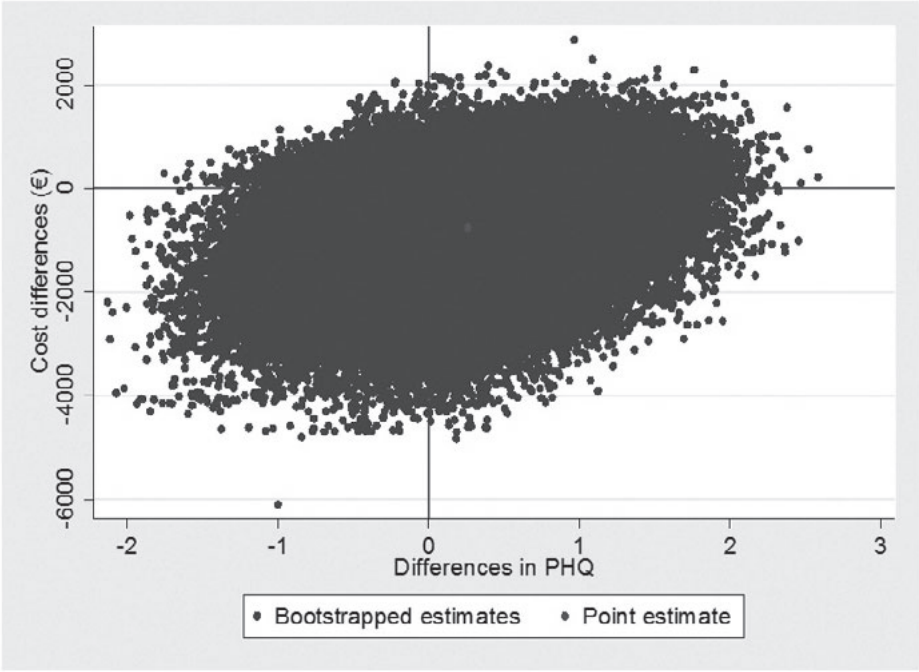


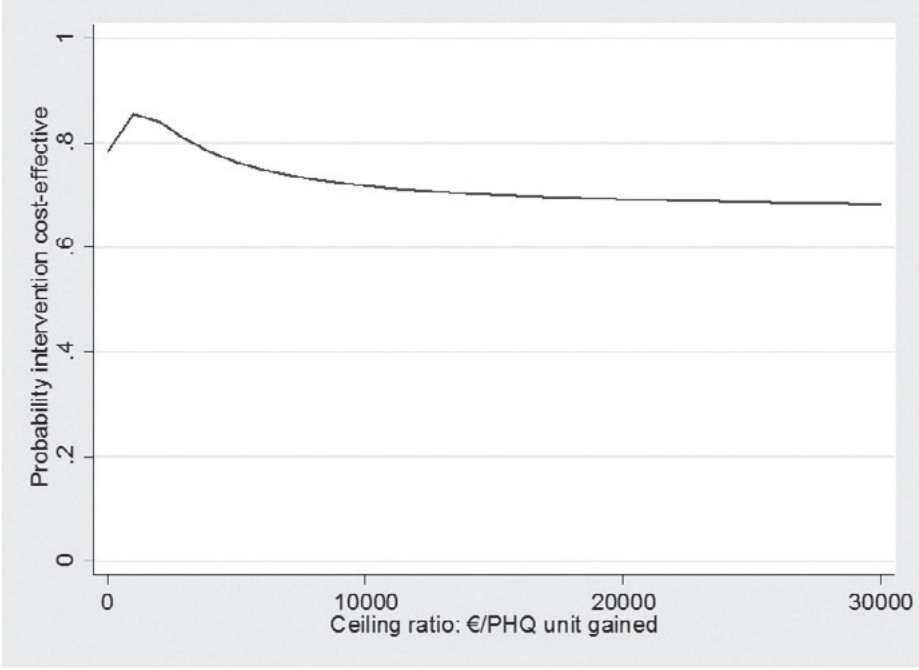
Cost-effectiveness plane and acceptability curve for RAND-36 PCS



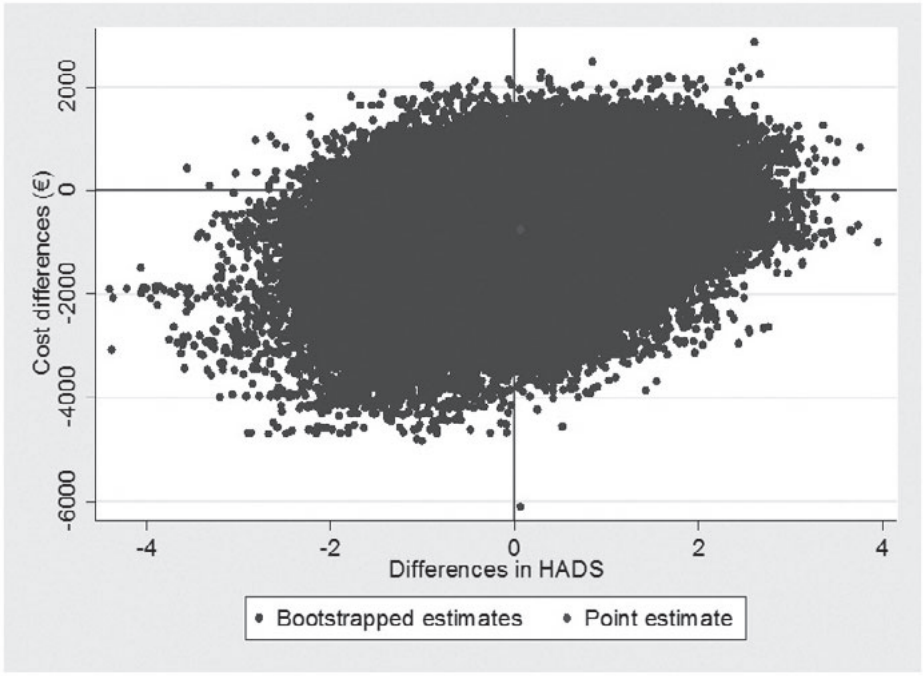


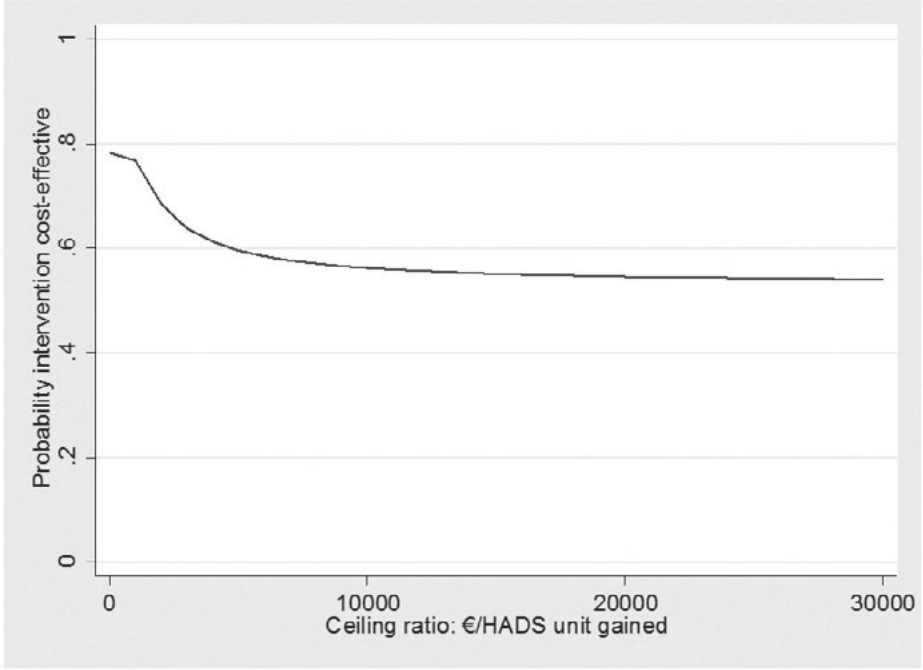
Cost-effectiveness plane and acceptability curve for PHQ-15 somatic symptom severity scale





Cost-effectiveness plane and acceptability curve for HADS

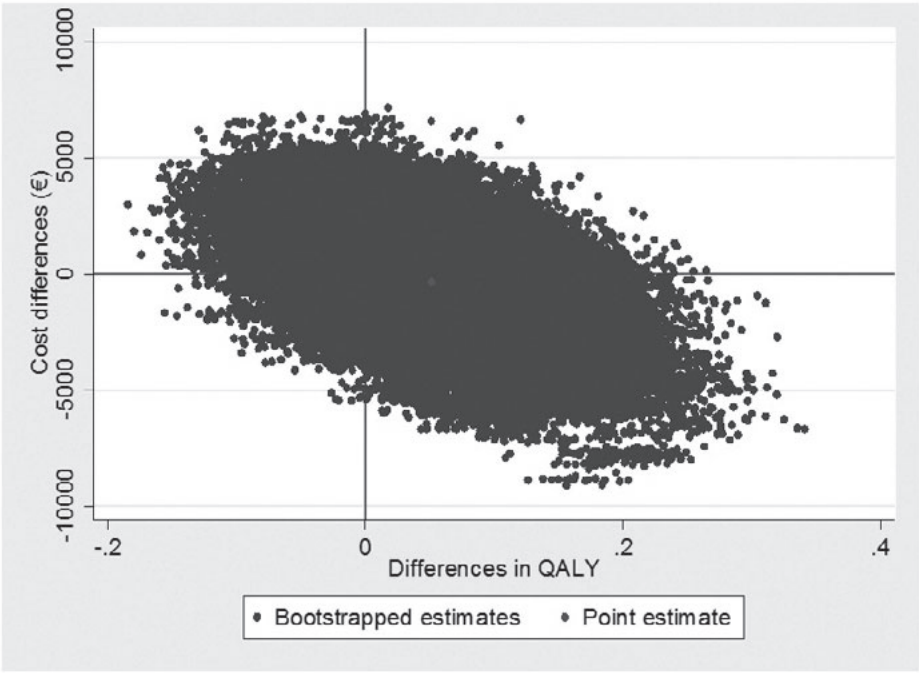


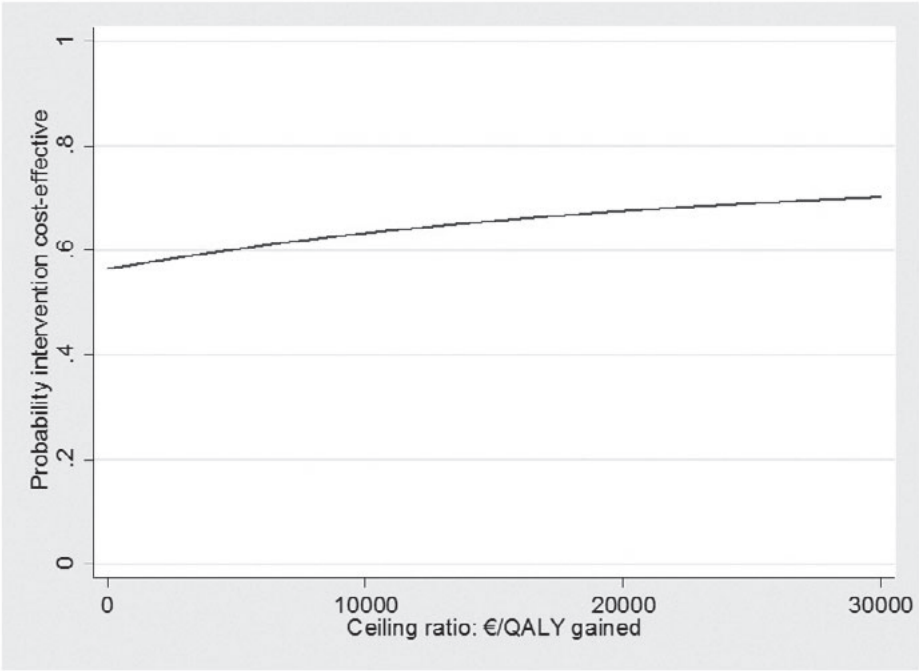


Abbreviations: HADS: Hospital Anxiety and Depression Scale; PCS: Physical Component Summary Score; PHQ-15: Patient Health Questionnaire 15-item somatization scale; QALYs: quality adjusted life years

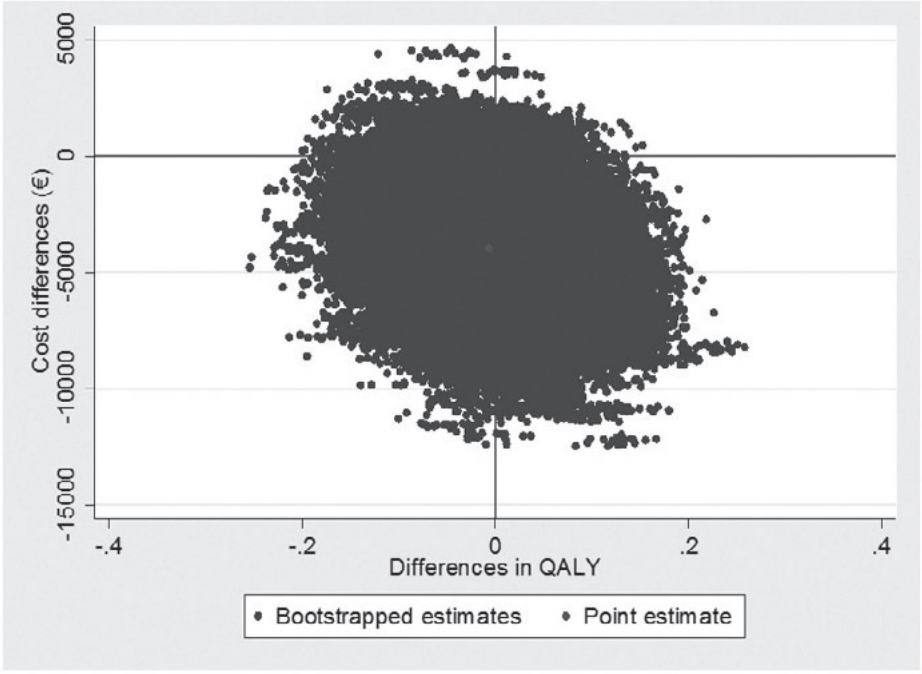
Appendix E. Results of the subgroup analyses

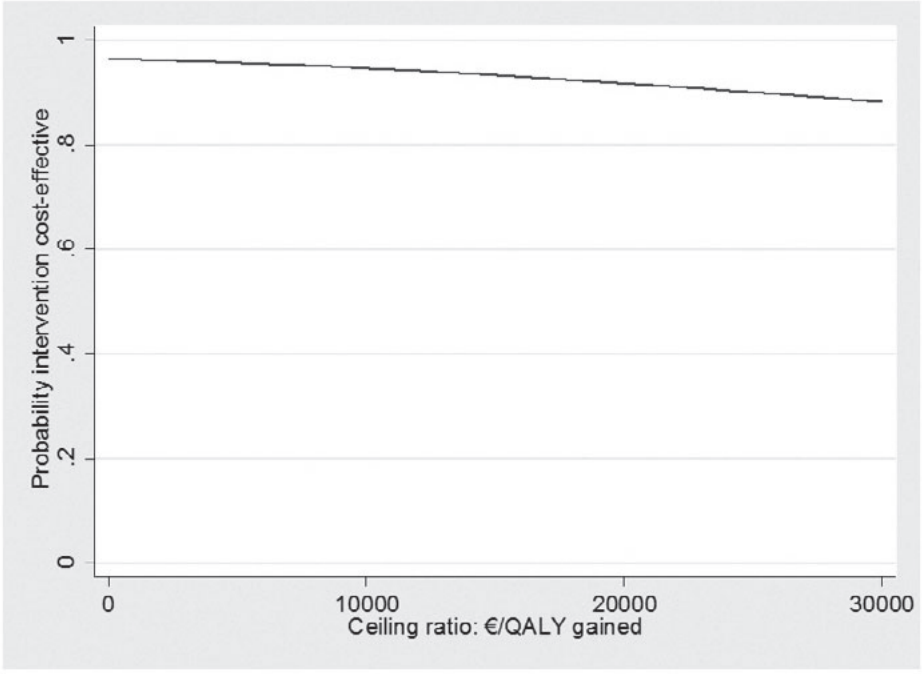
QALYs - Symptom duration below median (0-2092 days), cost-effectiveness plane and acceptability curve



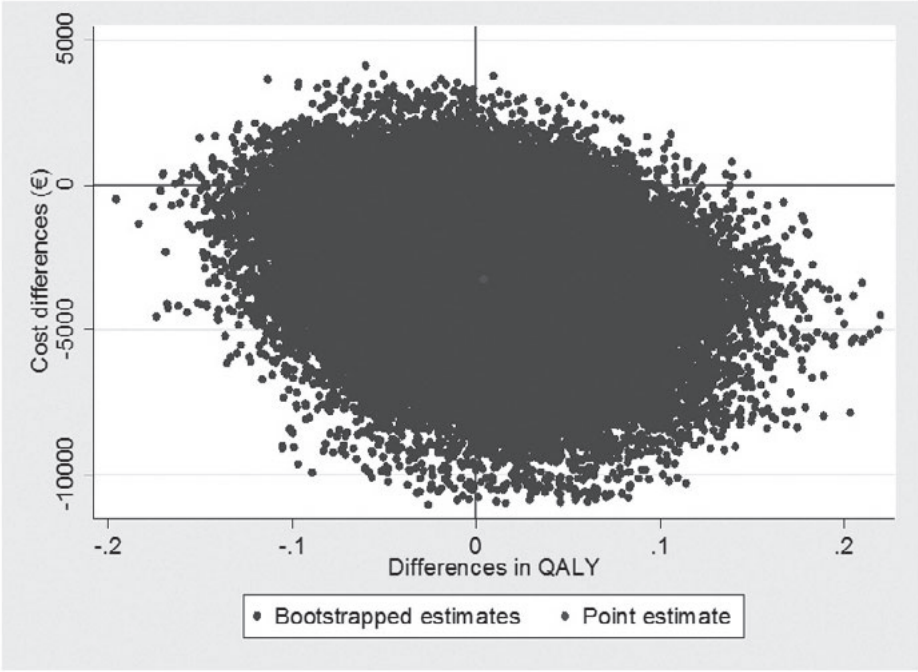


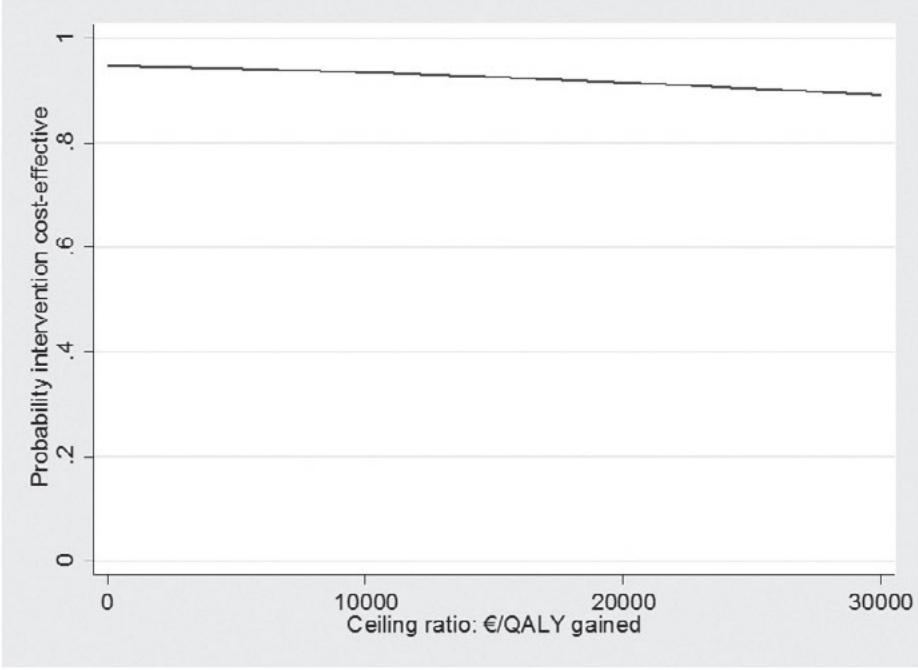
QALYs - Symptom duration above median (2093+ days), cost-effectiveness plane and acceptability curve



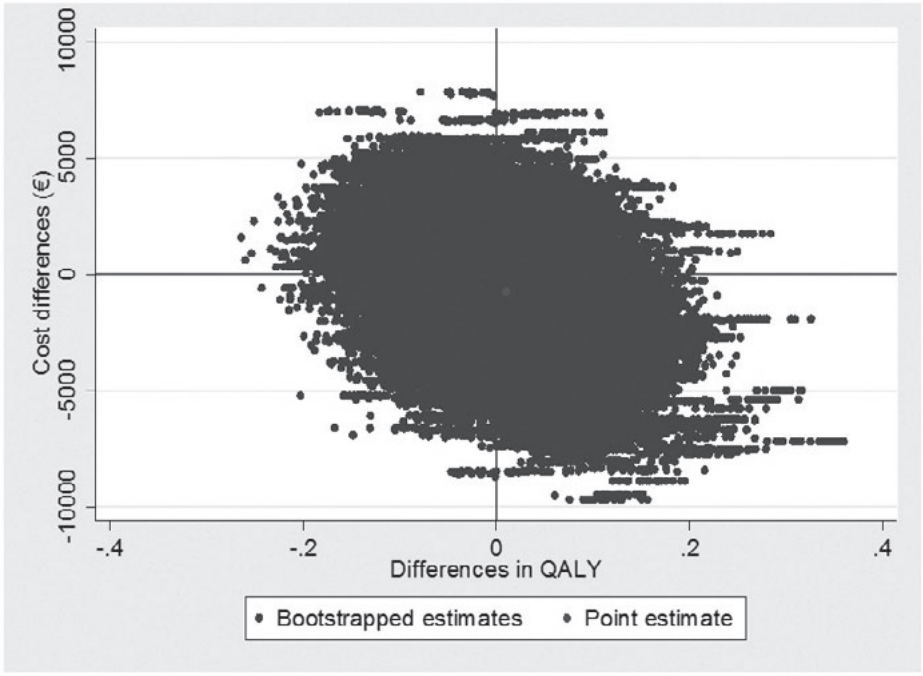


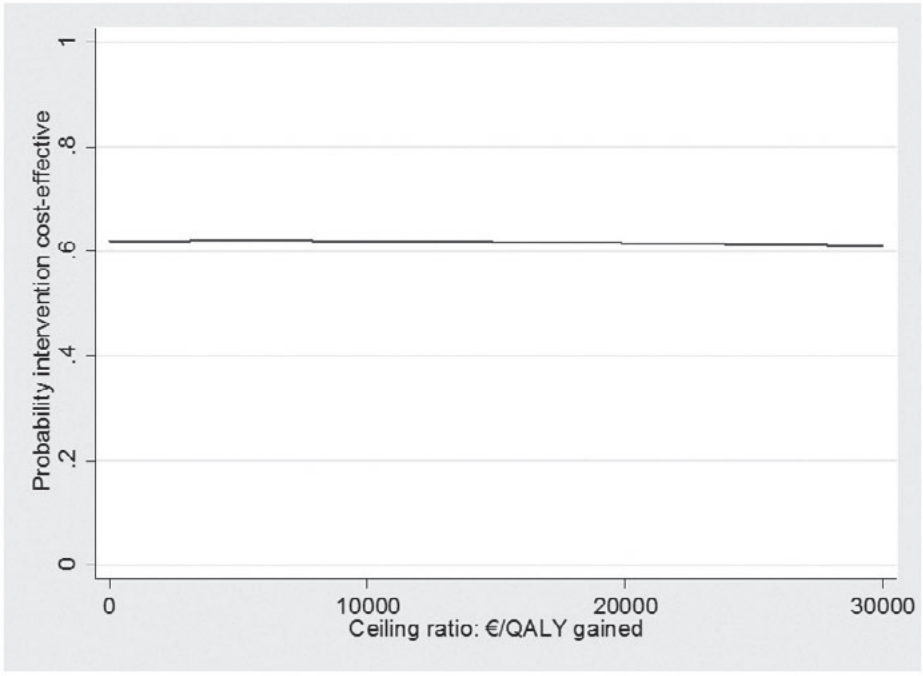
QALYs - 0-2 comorbid physical diseases, cost-effectiveness plane and acceptability curve



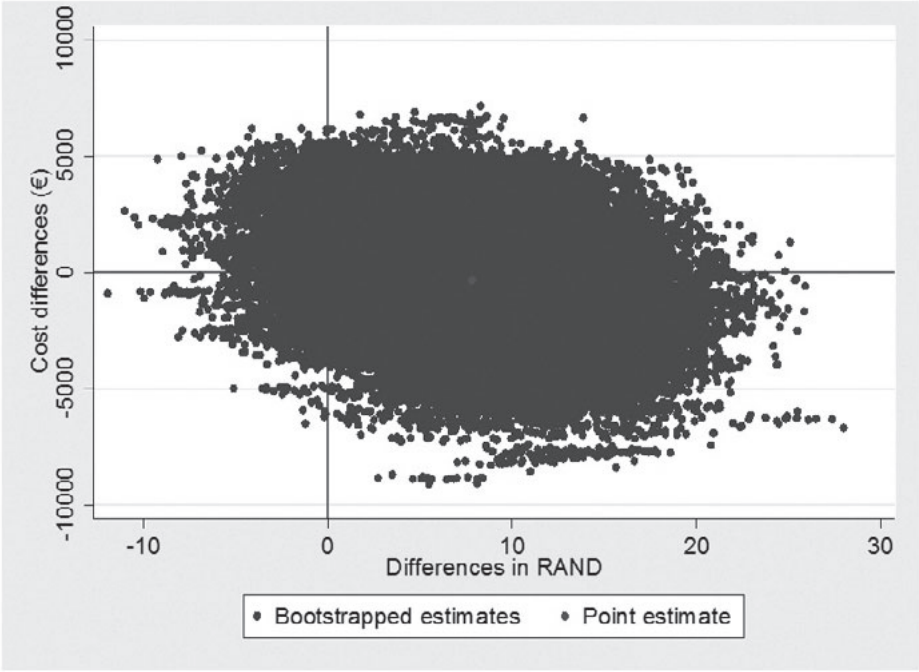


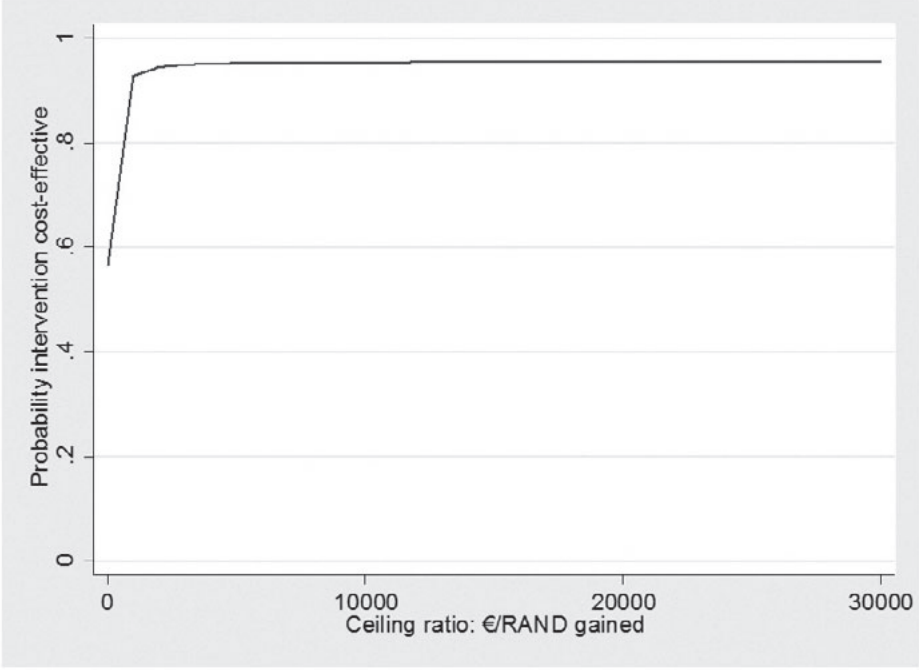
QALYs - 3 or more comorbid physical diseases, cost-effectiveness plane and acceptability curve



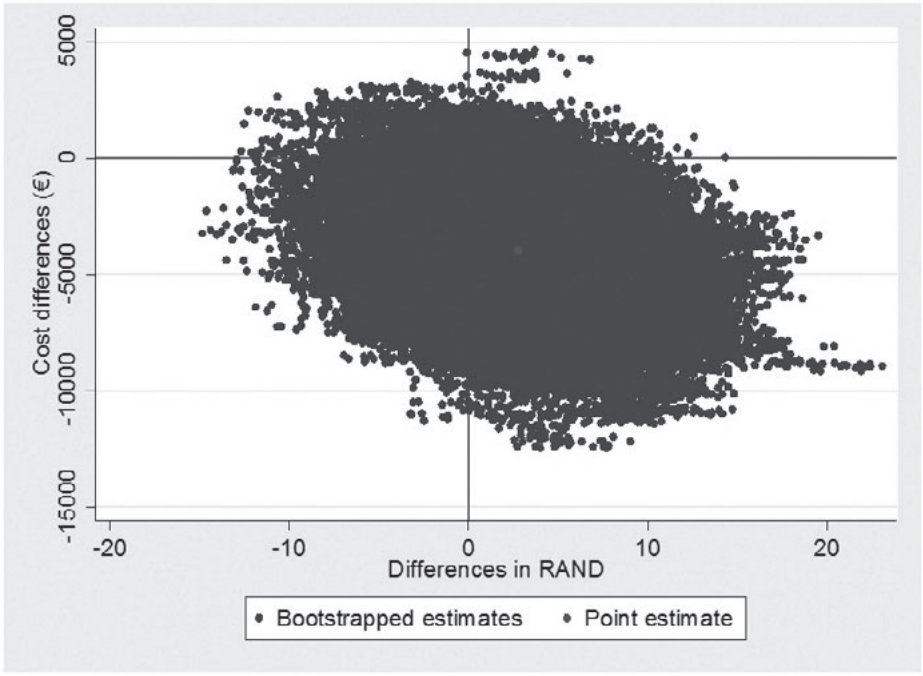


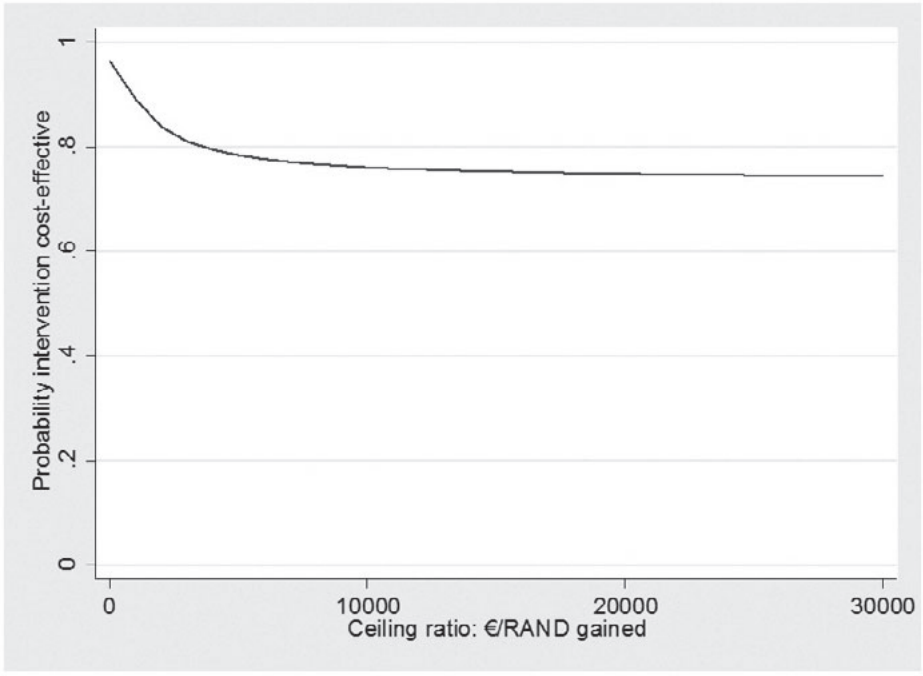
RAND-36 PCS - Symptom duration below median (0-2092 days), cost-effectiveness plane and acceptability curve



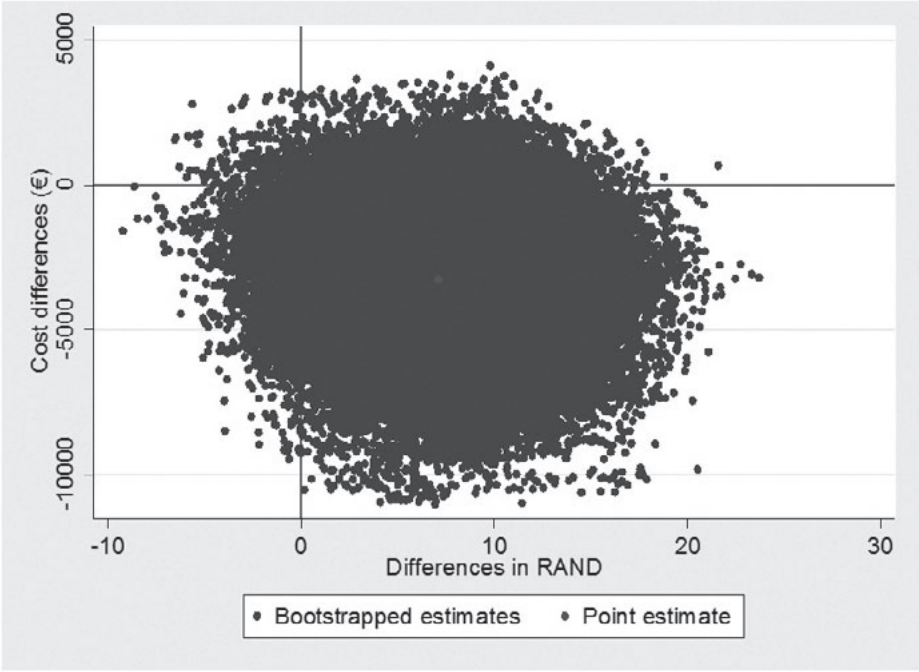


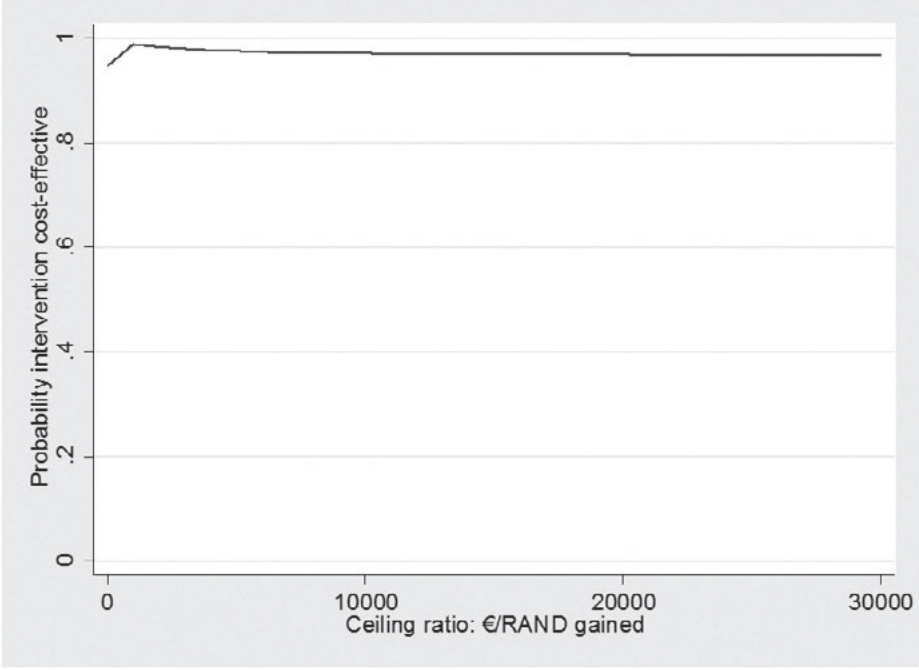
RAND-36 PCS – Symptom duration above median (2093+ days), cost-effectiveness plane and acceptability curve



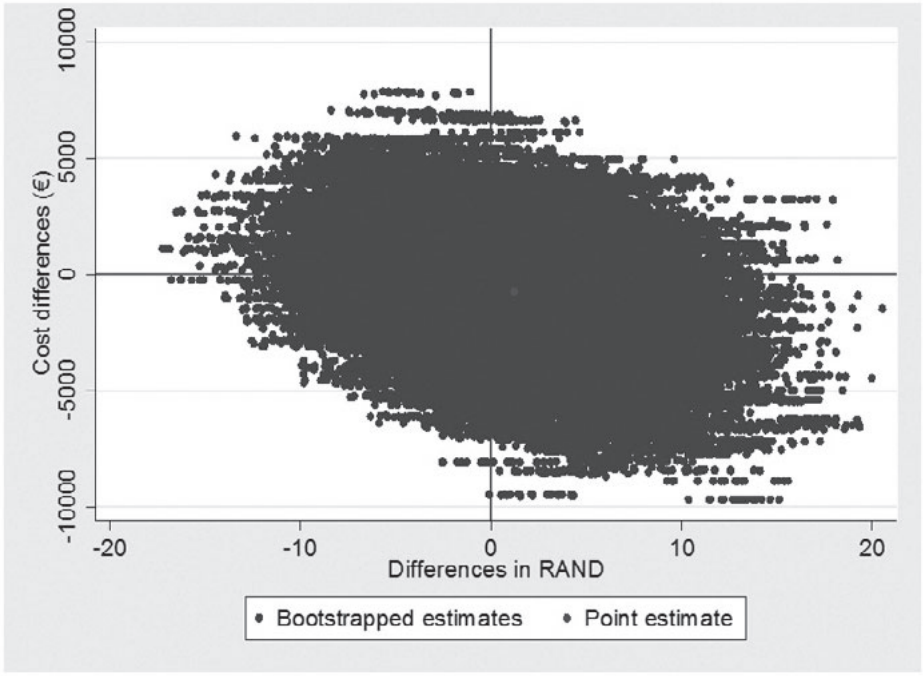


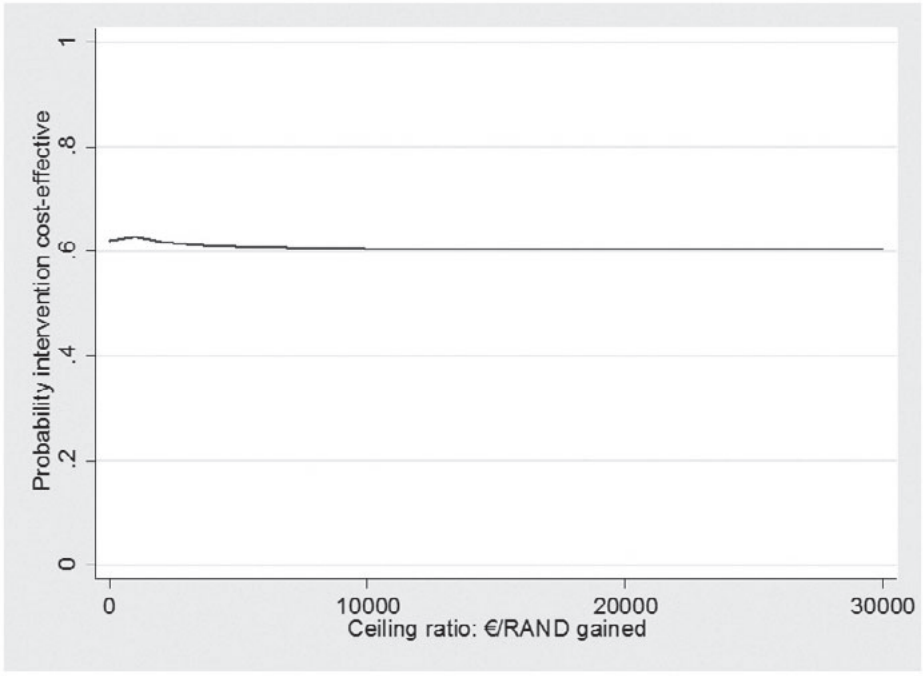
RAND-36 PCS – 0-2 comorbid physical diseases, cost-effectiveness plane and acceptability curve



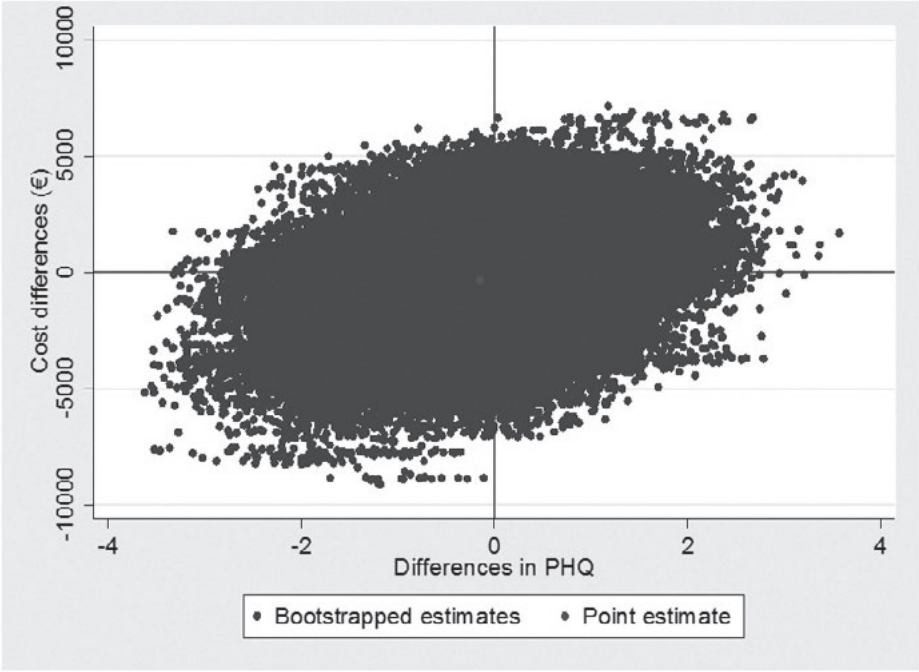


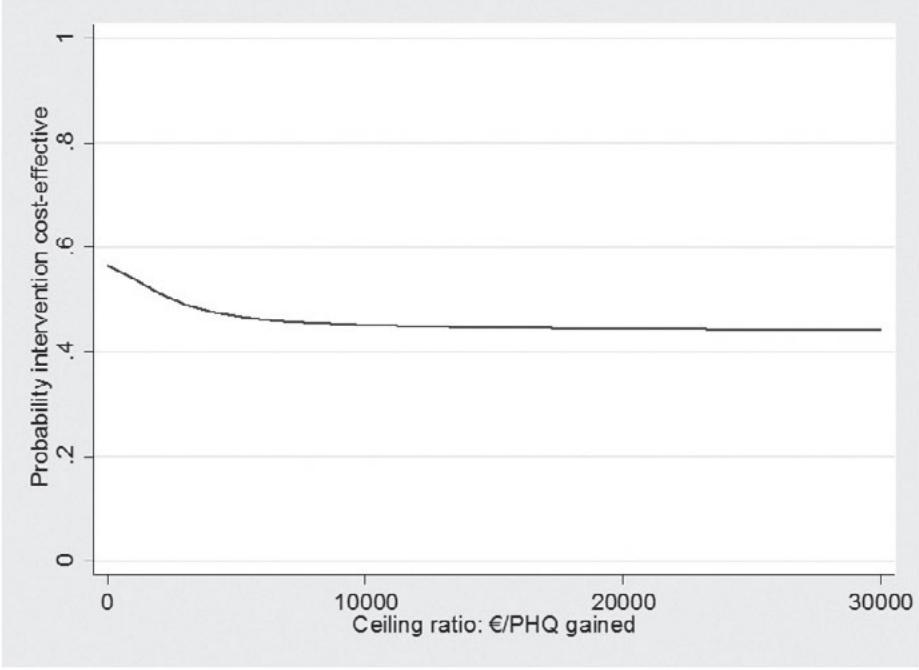
RAND-36 PCS, 3 or more comorbid physical diseases, cost-effectiveness plane and acceptability curve



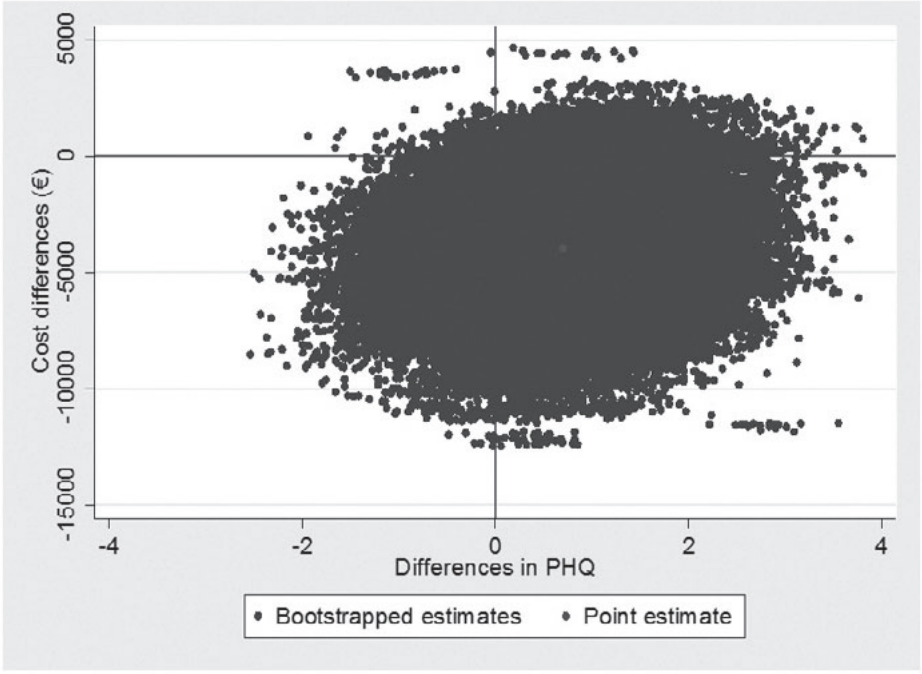


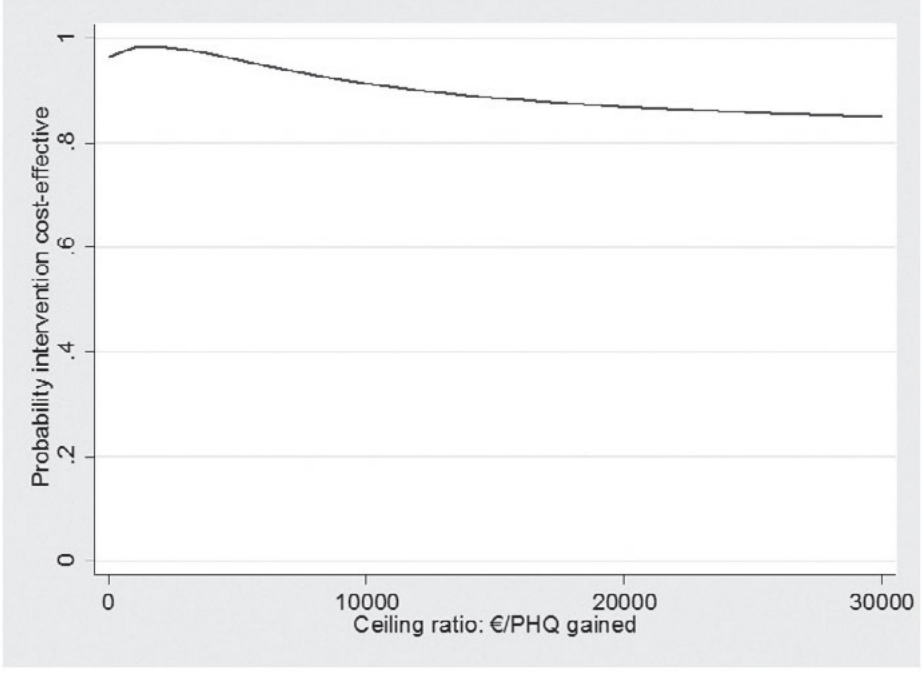
PHQ-15 somatic symptom severity scale - Symptom duration below median (0-2092 days), cost-effectiveness plane and acceptability curve



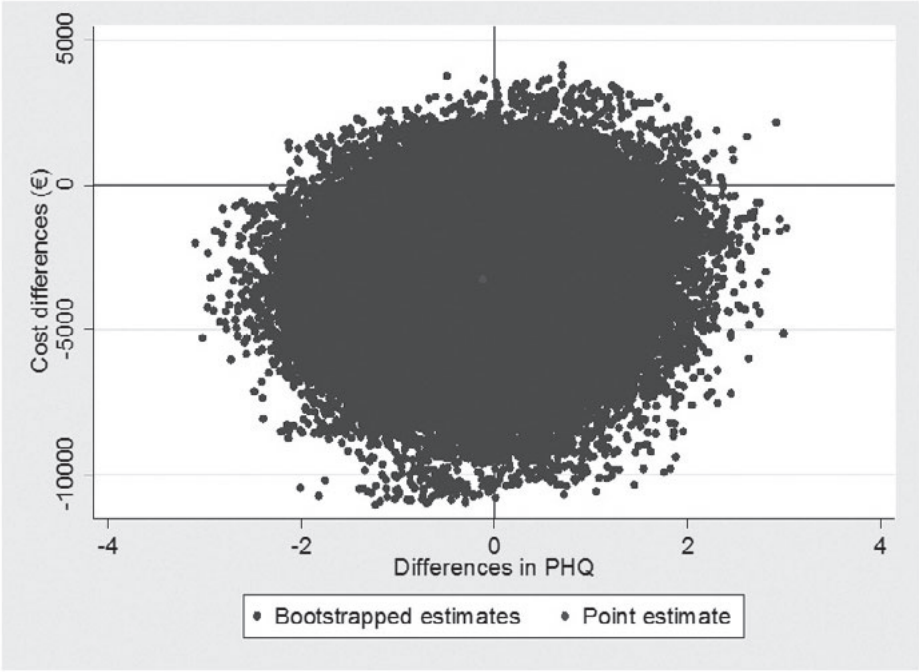


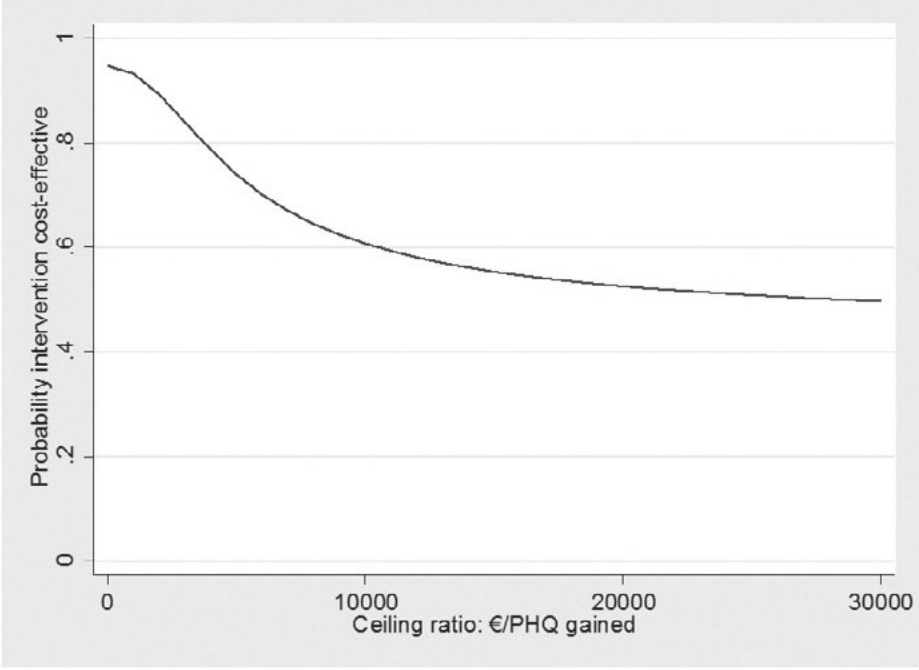
PHQ-15 somatic symptom severity scale - Symptom duration above median (2093+ days), cost-effectiveness plane and acceptability curve



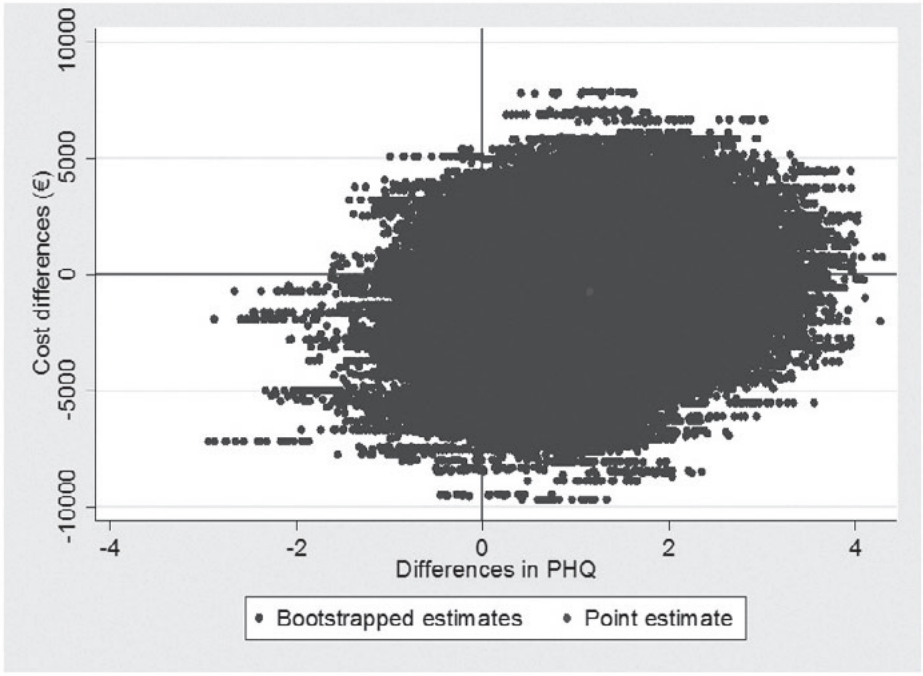


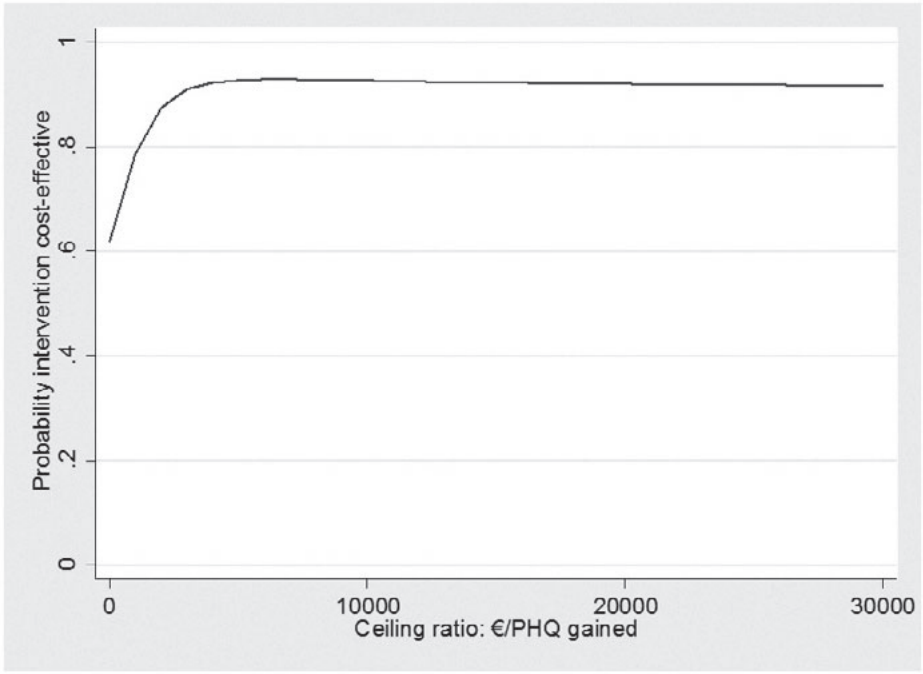
PHQ-15 somatic symptom severity scale - 0-2 comorbid physical diseases, cost-effectiveness plane and acceptability curve



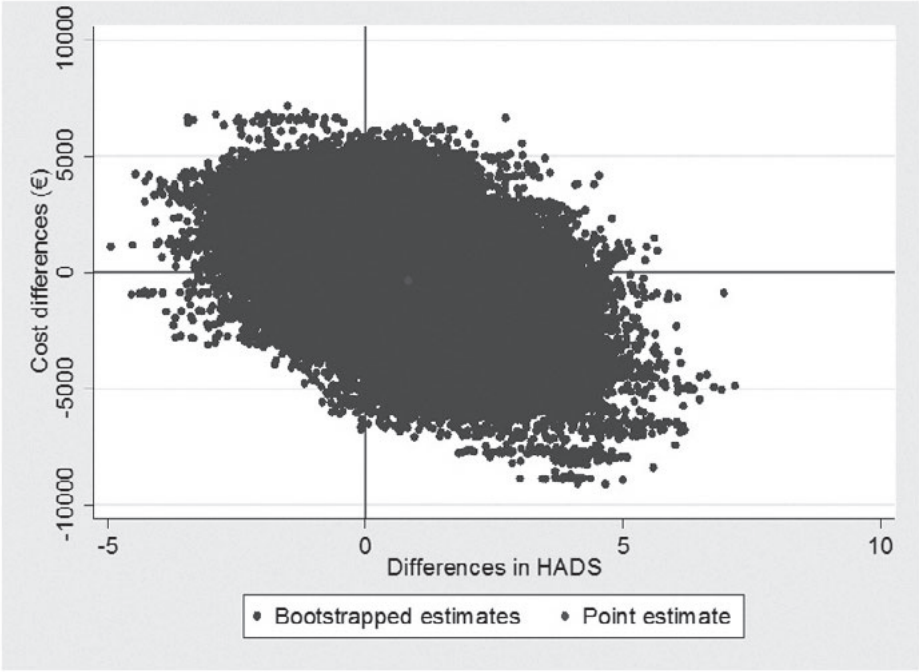


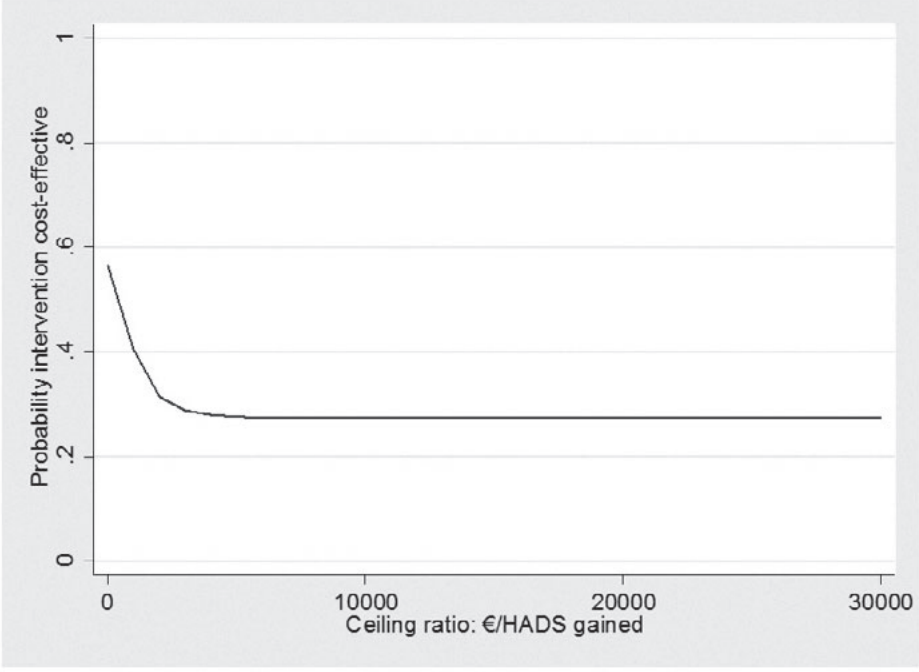
PHQ-15 somatic symptom severity scale - 3 or more comorbid physical diseases, cost-effectiveness plane and acceptability curve



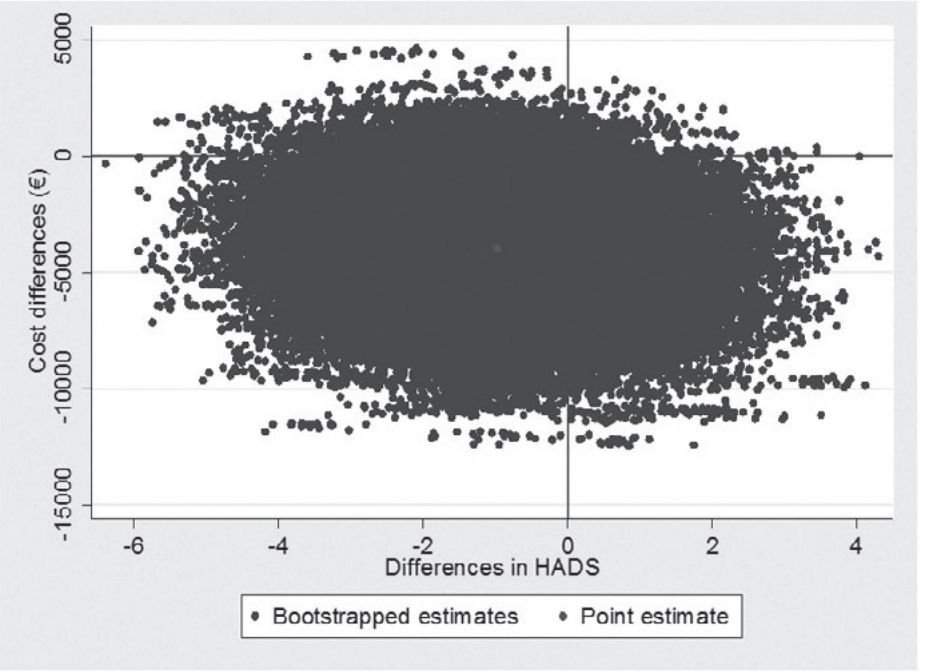


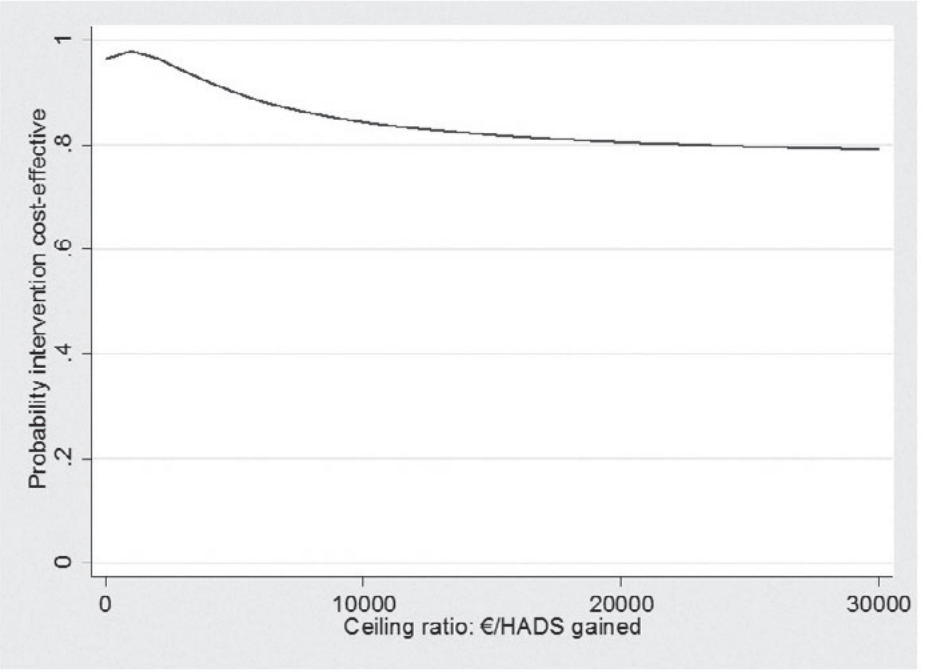
HADS - Symptom duration below median (0-2092 days), cost-effectiveness plane and acceptability curve



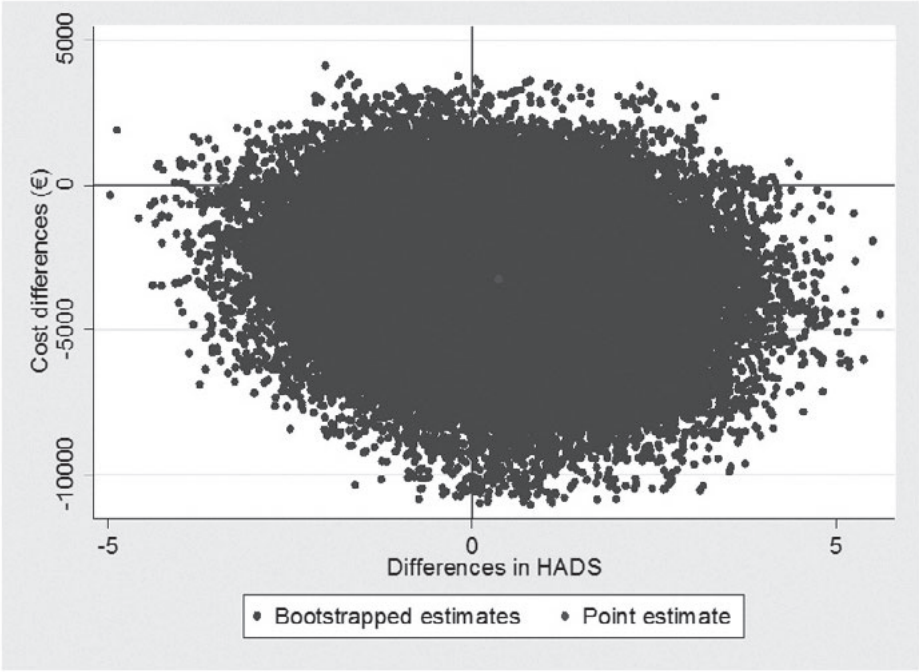


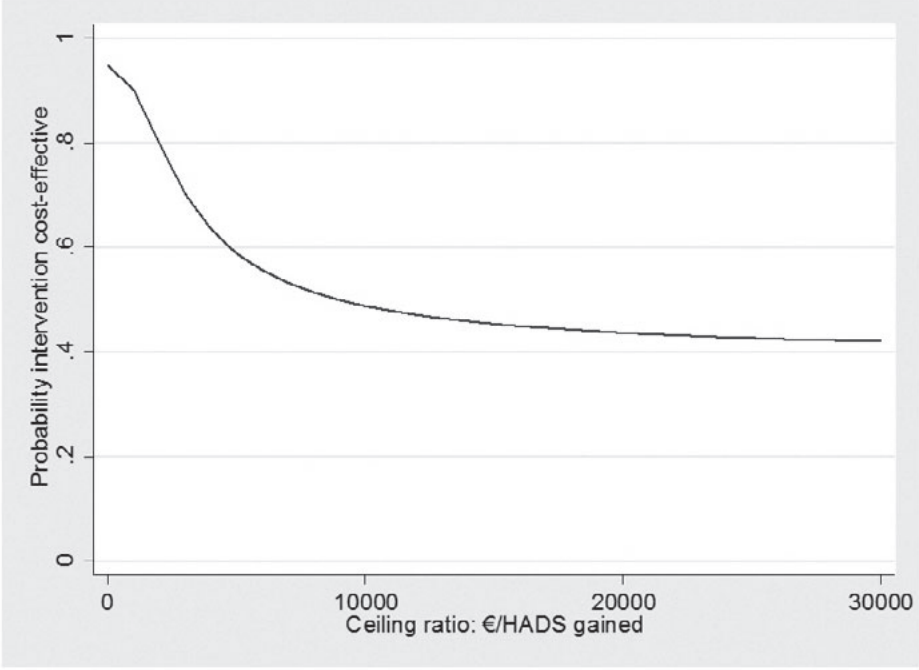
HADS - Symptom duration above median (2093+ days), cost-effectiveness plane and acceptability curve



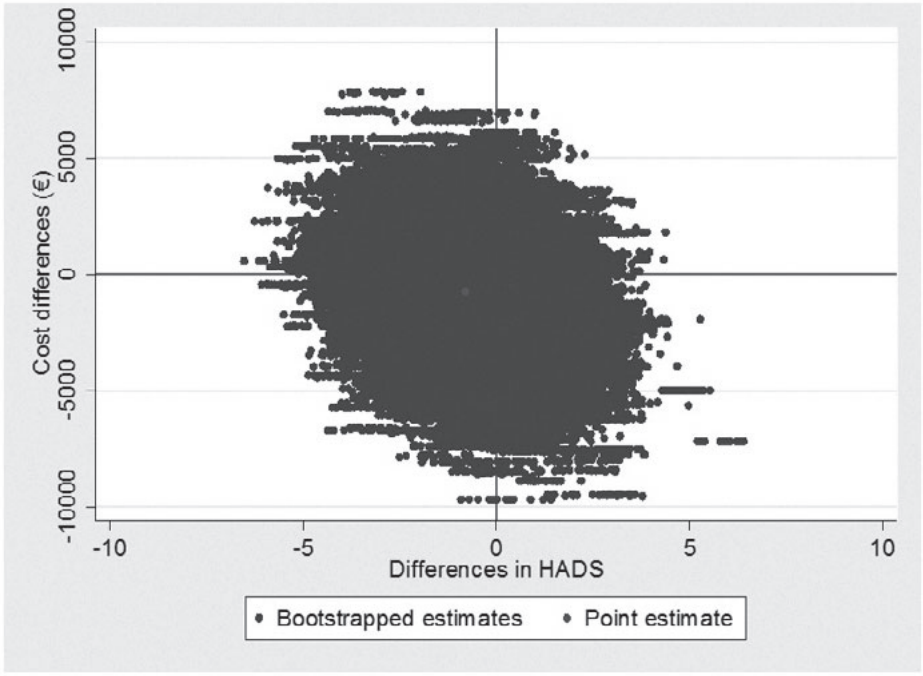


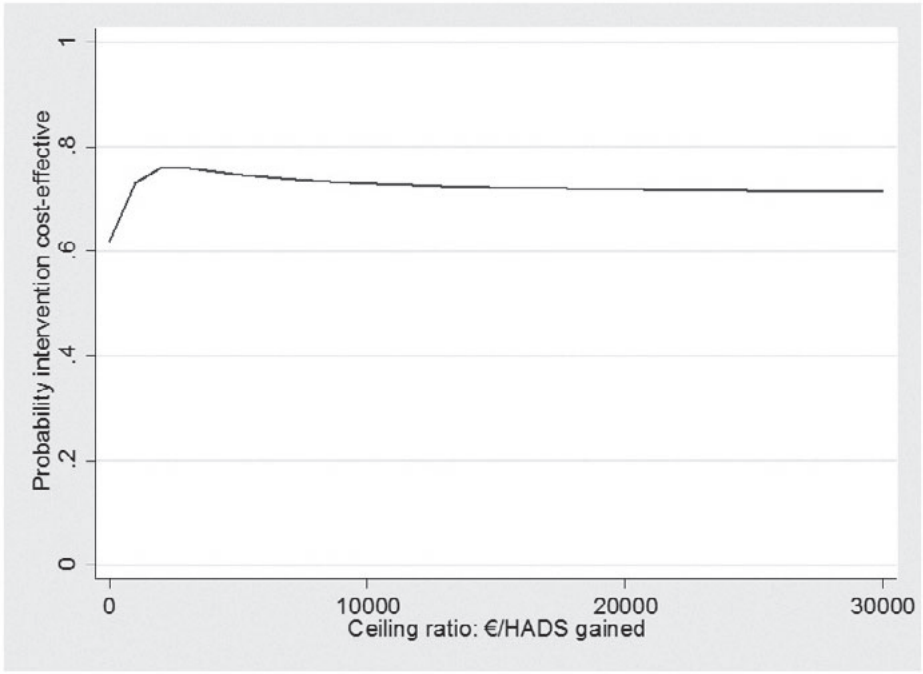
HADS – 0-2 comorbid physical diseases, cost-effectiveness plane and acceptability curve



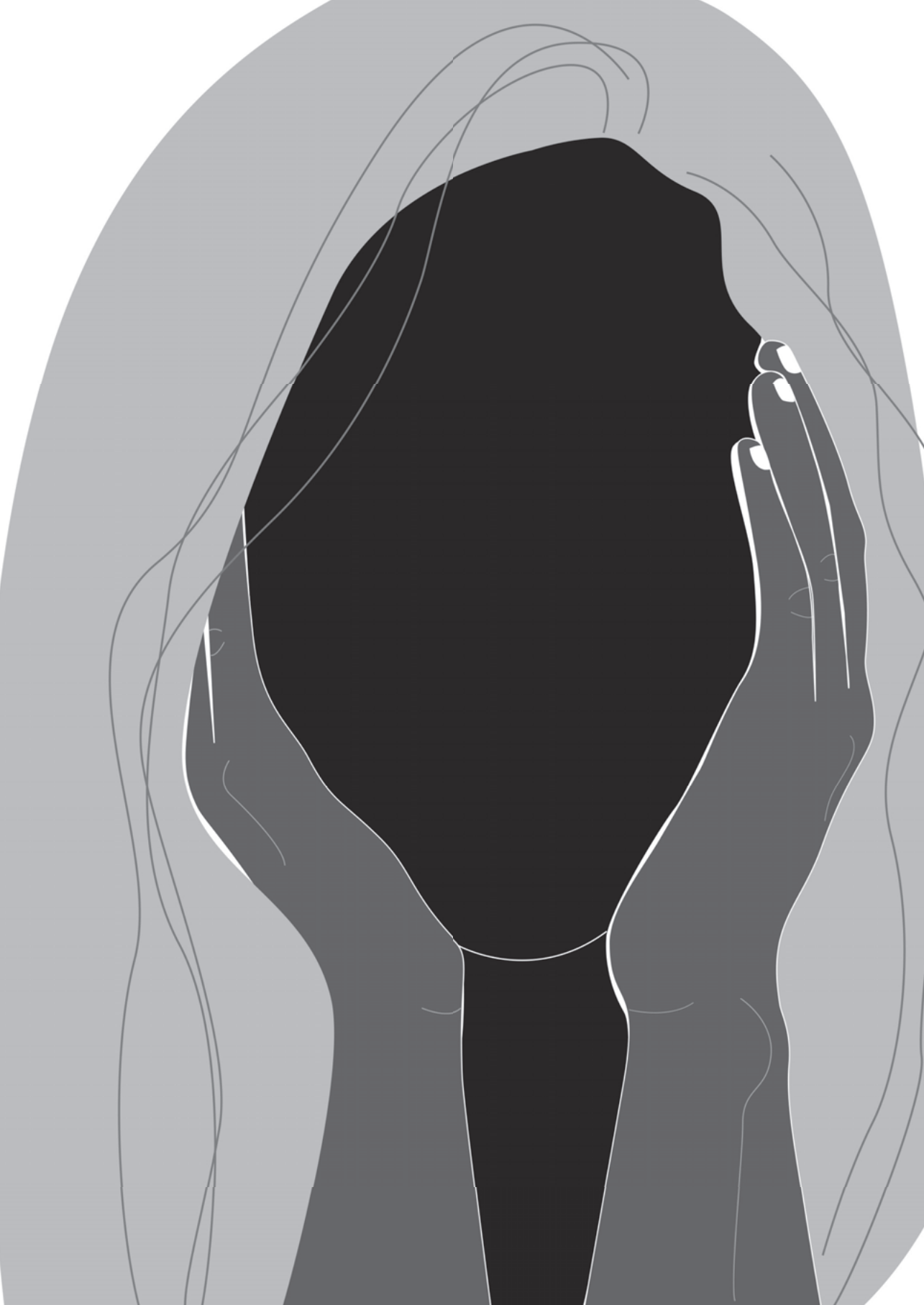


HADS - 3 or more comorbid physical diseases, cost-effectiveness plane and acceptability curve





Abbreviations: HADS: Hospital Anxiety and Depression Scale; PCS: Physical Component Summary Score; PHQ-15: Patient Health Questionnaire 15-item somatization scale; QALYs: quality adjusted life years



Chapter 7

Process evaluation of a cognitive behavioural intervention for undifferentiated somatoform disorder in Dutch primary care: mental health nurse practitioners' and patients' experiences

Kate Sitnikova
Harm WJ van Marwijk
Stephanie S Leone
Henriëtte E van der Horst
Johannes C van der Wouden

ABSTRACT

Background: Patients with undifferentiated somatoform disorder (USD) are common in primary are, but general practitioners (GPs) report having difficulty adequately helping them, due to the unexplained nature of the symptoms. Mental health nurse practitioners (MHNPs) are increasingly helping patients with psychosocial problems in Dutch surgeries. As part of a cluster randomized controlled trial that we conducted, investigating the effectiveness of a new psychological intervention for patients with USD provided by MHNPs, we also conducted a process evaluation**.** The aim of the current paper is to report on the process evaluation of the trial by gaining more insight into the experiences of the MHNPs when carrying out the intervention and into patients' experiences in receiving the intervention.

Methods: a) All MHNPs who were involved in applying the study intervention were interviewed using semi-structured interviews based on a topic list developed beforehand. All interviews were audiotaped and transcribed verbatim. Data analyses were done by two researchers independently. Identified themes were based on the topic list, interview guide and data from the interviews. b) Patients participating in the intervention arm of the trial completed a written evaluation questionnaire.

Results: a) MHNPs reported they generally adhered to the intervention manual, but often adjusted the length of the sessions, because they considered 30 minutes to be too short. They did not feel comfortable following the text strictly when applying the treatment manual and would have liked more flexibility. MHNPs indicated that the intervention was less appropriate for at least a third of the participating patients. Most common reasons were psychological/psychiatric comorbidity (especially personality disorders and psychological trauma), psychosocial problems, low IQ and older age. MHNPs generally felt that the intervention was effective, especially for participants with less comorbidity and other problems, and those who were open to change. According to the MHNPs, patients acquired problem-solving tools and were able to deal with their symptoms better. The majority of the MHNPs also reported that the intervention had provided them with specific tools for treating patients with USD. The main barrier to implementation was the MHNPs' currently limited amount of time. Facilitators were cooperation of the GPs and proper scheduling and organization within the surgery. b) Half of the patients felt that the intervention helped them at least somewhat. A third of all patients were positive about undergoing the intervention again in the future, if needed.

Conclusion: Overall MHNPs felt the intervention was helpful for patients and gave them useful tools to work with patients with USD, while about half of the patients also reported that the intervention was helpful for them. Patients with a duration of symptoms shorter than 2 years were more likely to report that the intervention helped them. MHNPs would have liked more freedom and flexibility in applying the treatment manual, more time for the sessions and would like to introduce other treatment techniques when deemed necessary. MHNPs thought that pre-selecting patients with less comorbidity might have contributed to a more successful treatment. For future research, MHNPs and GPs could play a larger role in selecting patients for such an intervention.

INTRODUCTION

Medically unexplained physical symptoms (MUPS) are presented commonly in primary care (1-4). If MUPS persist and cluster, they may be severe enough to classify as a somatoform disorder according to the DSM-IV, the most common of which is the undifferentiated somatoform disorder (USD) (5). There is evidence that psychological interventions, and in particular cognitive behavioural therapy (CBT), help patients with somatoform complaints (6). However, patients and GPs may be opposed to turning to mental healthcare when it concerns physical symptoms and patients may not receive the proper care they need (7-10).

In 2014 the 'mental health nurse practitioner' (MHNP) has been introduced in Dutch general practice, due to a reform introduced by the government. Since then MHNPs increasingly offer help for mental health issues instead of the GP (11, 12). The MHNPs' new position in primary rather than secondary care (where they would, for instance, have more focused supervision), makes this shift towards service provision in primary care an interesting implementation case.

We evaluated the effects of a cognitive behavioural intervention for patients with USD delivered by MHNPs within general practice in a recent randomized controlled trial, called the CIPRUS study (13). The main findings of the CIPRUS study are presented in chapter 5. In short, the new intervention was effective in improving physical functioning, reducing limitations due to physical symptoms, and physical pain, compared to usual care. The intervention was particularly effective for patients whose physical symptoms had lasted for a limited period of time and with few comorbid physical diseases (13).

To gain a better understanding of our results, we can learn from the MHNPs' and patients' experiences. From an implementation point of view, it is also essential to understand the barriers and facilitators of delivering the intervention. The aim of the current study was, therefore, to deepen our understanding of the effects of the intervention by gaining insight into the experiences of the MHNPs when carrying out the intervention and patients' experiences in receiving the intervention.

METHODS

Research team

KS (MSc, trained psychologist) and SDK (MSc, health scientist) conducted all interviews. KS was the main researcher, SDK was the research assistant, both are female and followed courses in epidemiology. KS and/or SDK had met the MHNPs previously while recruiting MHNPs for participation in the study. Also, KS was one of the trainers of the training that MHNPs received in order to carry out the intervention. In case MHNPs had any questions during the study period, they could contact KS or SDK. MHNPs reported the dates on which they saw the patients to SDK by e-mail.

CIPRUS study

Trial design

The MHNPs and patients that were included in the current process evaluation participated in the intervention arm of the CIPRUS study (14). The CIPRUS study was a cluster randomized controlled trial, that investigated treatment of patients with MUPS severe enough to be classified as undifferentiated somatoform disorder (USD), delivered by MHNPs within general practice. A detailed account of the methods, design, treatment and results of the trial is described in chapters 3, 5 and 6 (13-15).

Eligible patients were adults, aged 18 and above, who had consulted their GP at least twice, with one or more physical symptoms associated with functional somatic syndromes, in the previous 3 months; whose symptoms were identified as unexplained by their GP; who had a score of at least 5 on the symptom severity scale Patient Health Questionnaire 15 (PHQ-15) (16) (indicating at least mild symptom severity); and who fulfilled the DSM-IV criteria for undifferentiated somatoform disorder (5) according to the Structured Clinical Interview for DSM-IV Axis I Disorders (SCID-I) (17). Exclusion criteria were: having a medical or psychological disorder that explained the symptoms based on which patients would participate in our study; having a severe psychiatric disorder (e.g. psychotic disorder); currently receiving psychological help for USD; having poor

language skills or physical handicaps interfering with understanding the intervention or questionnaires.

Intervention

The treatment in the intervention group consisted of 6 individual sessions with the MHNP, of 30 minutes each. The aim of the treatment was to improve patients' physical functioning by helping them cope with consequences of their physical symptoms and daily problems in general. The treatment was a new method based on a combination of two techniques: the consequences model for somatoform disorders (18, 19) and Problem-Solving Treatment (PST) (20). In the consequences model the focus is taken away from possible causes of the unexplained symptoms, and is turned to the various consequences that the patient experiences due to the physical symptoms. PST consists of seven steps that help define and solve the problem/consequence in the patient's life. During the intervention, the physical complaints and their consequences were defined first, and then tackled using augmented PST, tailored specifically for MUPS. MHNPs received a two-day group training lasting 3 to 3,5 hours each from a clinical psychologist and one of the researchers (KS) before starting. All but one participating MHNPs followed the group training. One MHNP was not able to participate in the group training and received the training individually from the main researcher (KS).

During the training, MHNPs received a copy of the intervention manual in which each session was described in detail. The MHNPs were requested to follow the sessions as described in the manual as closely as possible. In session 1 the MHNP introduced and explained the treatment, the patient told the MHNP about his/her physical symptoms and consequences or problems that arose in the patient's life due to the physical symptoms. In session 2 the MHNP explained the PST goals and steps. In sessions 3-6, the MHNP and the patient addressed a single consequence/problem per session, using the steps. Patients received homework to apply the steps at home. If not all steps were covered during one session, this was continued during the following session. MHNPs were offered supervision while they were carrying out the treatment.

Participant selection

MHNPs

Participating MHNPs and patients in the intervention group were involved in this process evaluation. 48 surgeries and 17 MHNPs were included in the intervention group. In 4 surgeries where 2 MHNPs worked, no patients were recruited, so these MHNPs did not carry out the intervention. The remaining 15 MHNPs were invited to participate in the interviews by e-mail. Researchers stated the reasons for conducting this research in the

e-mail and once more at the beginning of the interview. Those who did not reply were sent an e-mail reminder and then contacted by phone. Before the interviews, all participating MHNPs signed an informed consent form, giving the researchers permission to make audio-recordings.

Patients

All patients in the intervention group (n=111) of the CIPRUS study were eligible and invited to complete the evaluation questionnaire.

Data collection

MHNPs

KS or SDK conducted all interviews with the MHNPs between June and August 2017. One interview was conducted by KS while SDK listened. No one else was present at any of the interviews. Interviews were held in the surgeries where the MHNP worked when possible, or else by telephone. The interviews lasted about 30 minutes each, were audio-recorded and transcribed verbatim.

MHNPs were explicitly asked to answer the questions truthfully, so that the intervention could be improved based on their feedback. Before the interview, all MHNPs were requested to complete a 9-item questionnaire with pencil and return it to the researcher (Appendix A). Three items concerned their professional background and 6 items concerned the evaluation of the intervention. These latter 6 items were scored on a 0-5 Likert scale.

Data was then collected using a semi-structured interview. In order to structure the interviews and warrant uniformity among the different MHNPs, we developed a topic list (Appendix B), from which an interview guide was derived (Appendix C). We based the topic list on the RE-AIM framework (21); a framework for process evaluation for cluster-randomized trials of complex interventions developed by Grant (22); and the guideline for assessing health promotion program implementation developed by Saunders (23). The final topics used were ‘reach’, ‘effect’, ‘implementation’ and ‘maintenance’. We also added topics ‘general’ about participation in the trial and applying the treatment protocol, and ‘context’ about the general practice setting.

After the first interview, we marginally adjusted the interview guide. We did not carry out repeat interviews with the same MHNP. Where applicable, we typed out field notes and saved these in the same document as the interview with the MHNP. As all available MHNPs were interviewed, data saturation was not an issue. We did not return transcripts

to participants for correction, but sent a draft of this manuscript to the MHNPs, with a request to provide feedback on our findings.

Patients

For the present purpose, 13 items concerning the evaluation of, and satisfaction with the intervention were added to the assessment sent to patients after they completed their 6 sessions with the MHNP (i.e. 4-month follow-up). Most of the evaluation and satisfaction items were scored on a 1-5 Likert scale (Appendix D). Patients completed the questionnaires online or on paper.

Data analysis

We analyzed the interview data according to a framework approach (24). Two of the authors (KS and SDK) transcribed the interviews. They then both coded and analyzed the transcriptions of all interviews independently. The transcribed interviews were entered into Atlas.ti version 7.5.18 software, and codes were attached to every quotation. KS and SDK compared their codes and discrepancies were discussed until consensus was reached. If needed, a third researcher, JCW was consulted. We partly identified the themes in advance based on the topic list (Appendix B) and the interview guide (Appendix C), and partly derived them from the collected data.

For this chapter, we used descriptive statistics regarding questionnaires completed by MHNPs and patients, and the COREQ guidelines for reporting qualitative research regarding interviews with the MHNPs (25). We analyzed data from patients' questionnaires using descriptive statistics in SPSS version 22.

RESULTS

Quantitative results

MHNPs

All 15 MHNPs participating in the CIPRUS study were invited to complete the 9-item questionnaire and take part in the interviews. Thirteen agreed. Two MHNPs could not be reached: one of them did not work in the surgery anymore and the other one was absent from work due to long-term illness. We conducted eleven interviews face-to-face and two by telephone. Characteristics of the participating MHNPs are presented in table 1. Table 2 provides an overview of the answers MHNPs provided on the items of the MHNP-questionnaire (Appendix A), regarding their evaluation of the trial. From here on we will refer to all MHNPs as 'she' in the text, to avoid recognizability.

Table 1. Characteristics of participating MHNPs

MHNP ID	Professional background	Gender	Years of experience as MHNP	Years of relevant experience in mental health	Number of CIPRUS study patients assigned to MHNP	Number of CIPRUS study patients seen by the MHNP for at least 1 session
MHNP 1	Nurse	Female	1,5	20	5	5
MHNP 2	Social psychiatric nurse	Female	3	25	6	6
MHNP 3	Social psychiatric nurse	Female	9	30	9	9
MHNP 4	Nurse and MHNP	Female	1,5	2	5	5
MHNP 5	Social worker, MHNP	Female	5	12	3	3
MHNP 6	Applied psychologist	Male	3	7	3	3
MHNP 7	Social worker, gestalt therapist	Male	3	2	19	17
MHNP 8	Psychiatric nurse	Female	4	27	18	17
MHNP 9	Social worker	Female	4	6	10	9
MHNP 10	Psychologist	Male	2	2	1	1
MHNP 11	Mental health nurse	Female	3	28	5	5
MHNP 12	Psychologist	Female	1,5	2,5	13	11
MHNP 13	Psychologist	Male	3	1	5	5

Abbreviation: CIPRUS: Cognitive-behavioural Intervention in PRimary care for Undifferentiated Somatoform disorder; MHNP: mental health nurse practitioner

Table 2. MHNPs' answers to the MHNP evaluation questionnaire (n=13)

	0 (Not at all)	**1**	**2**	**3**	**4**	**5 (Completely/ very much so)**
1. How satisfied are you with the CIPRUS study manual?	0	1	0	3	8	1
2. To what extent were the patients that you saw for the CIPRUS study suitable for this study?	0	0	4	4	5	0
3. Did the manual help patients to deal with their MUPS better?	0	1	2	5	3	2
4. Did the manual provide you with tools for dealing with MUPS patients?	0	0	2	4	4	3
5. Did you adhere to the manual?	0	0	0	2	10	1
6. Would you use the manual again with patients with MUPS in the future?	0	0	4	1	4	3

The numbers in the cells represent the how many MHNPs selected this answer to the question.
Abbreviation: CIPRUS: Cognitive-behavioural Intervention in PRimary care for Undifferentiated Somatoform disorder; MHNP: mental health nurse practitioner; MUPS: medically unexplained physical symptoms

Patients

Recruitment and baseline characteristics

An overview of the patient enrollment procedure and follow-up assessments in the trial is provided in figure 1. 86 intervention patients (77.5%) completed the patient evaluation questionnaire at 4 months. Their mean age was 54.5 and 69% were female. The mean PHQ-15 score at baseline was 13.6 (SD=4.9), which indicates medium symptom severity. An overview of patient characteristics is provided in table 3.

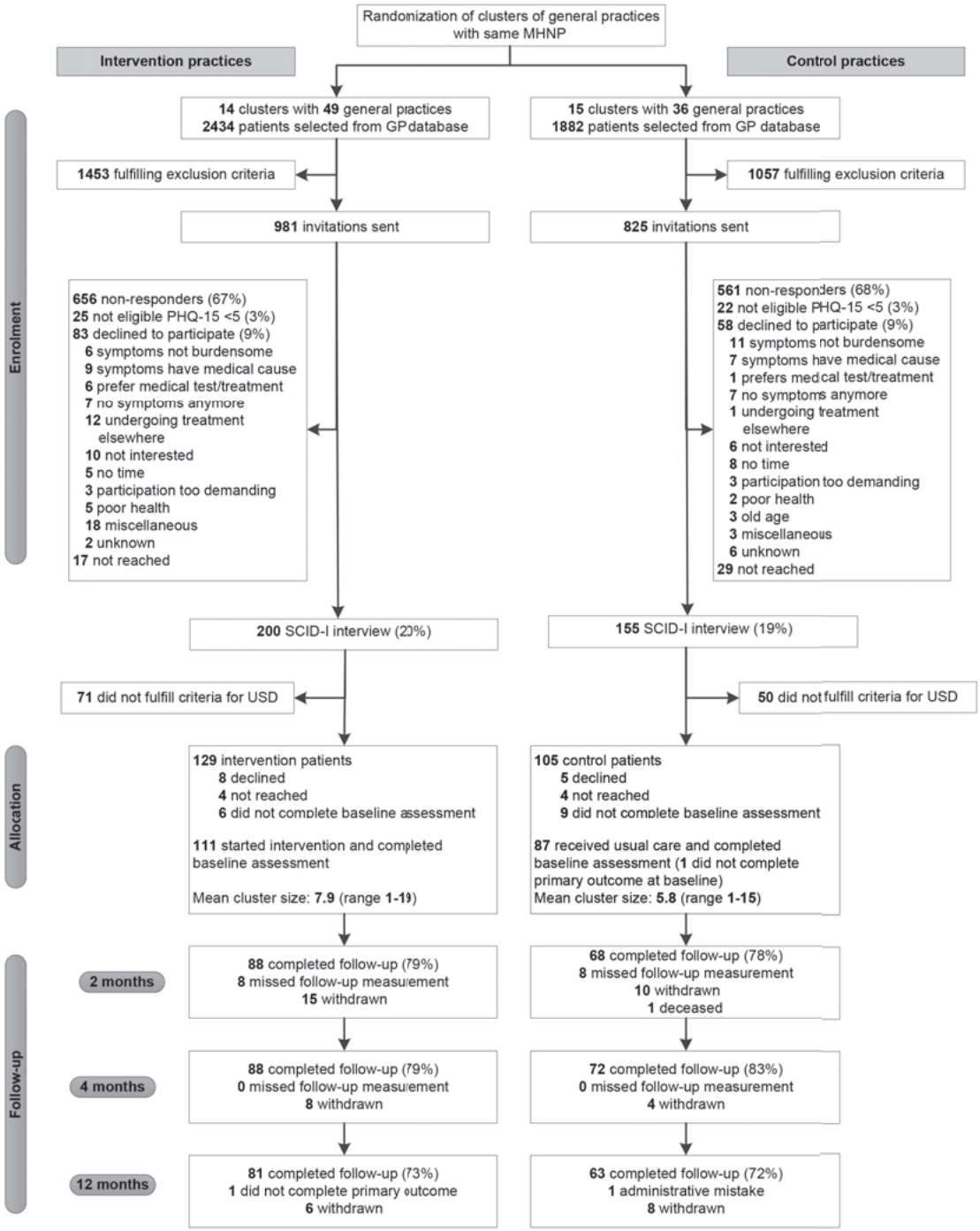


Figure 1. Flow of study participants
GP general practitioner; MHNP mental health nurse practitioner; PHQ-15 Patient Health Questionnaire 15-item somatic symptom severity scale; SCID-I Structured Clinical Interview for DSM-IV Axis I Disorders

Table 3. Patient characteristics at baseline

Characteristics	Intervention group (n=86)
Age, mean (sd)	54.5 (15.5%)
Female	59 (68.6%)
Born in the Netherlands	80 (93.0%)
Education level *	
No education	1 (1.1%)
Lower education	3 (3.4%)
Intermediate vocational	44 (51.2%)
High education	27 (31.5%)
Academic education	8 (9.3%)
Work status	
Paid job	33 (38.4%)
No paid job	33 (38.4%)
Retired	20 (23.3%)
Living situation	
With partner / children / other co-habitants	65 (75.6%)
Alone	21 (24.4%)
Somatic symptom severity (PHQ-15) (sd)	13.6 (4.9)

* n=85 due to one missing
Abbreviation: PHQ-15: Patient Health Questionnaire 15-item somatic symptom severity scale; sd: standard deviation

Motivation

The most important reason for patients to participate was because their GP advised them to do so (45%), followed by wanting to find a better way to deal with their symptoms (40%) and wanting to get rid of the symptoms (29%). These percentages add up to more than 100% as patients could choose more than one answer option. About a quarter of the patients (26%) reported that they had a need for support in dealing with their physical symptoms before participation in the study.

Effectiveness and uptake of the intervention

The majority of patients rated the quality of the intervention as good (66%), or excellent (11%), while 14% rated it as mediocre or bad (1%). For nearly half of the patients, their expectations were (almost) completely met (44%), while for 17%, only some of the expectations were met, and 6% reported none of their expectations were met. Most patients (62%) felt that the number of sessions was sufficient but 20% felt there were too few sessions. A large majority had used the patient workbook (83%). 15% found the

workbook very useful, 31% found it somewhat useful, 22% was neutral, 11% did not really find it useful and 4% found it not useful at all.

More than half of the patients reported that the intervention helped them somewhat or a lot in dealing with their physical symptoms (51%), 22% were neutral and 17% said it did not help them. Most patients reported being fairly satisfied (42%), a quarter was extremely satisfied (25%) and 20% was neutral. Three people were somewhat (4%) or very unsatisfied (1%). We additionally investigated whether patients who reported that the intervention helped them deal better with their symptoms were also the same patients whose symptoms lasted for a shorter period of time and had fewer physical comorbid disorders. Of the patients with symptoms lasting shorter than 2 year (n=16 in the intervention group), the majority (68.8%) indeed said that the intervention helped them somewhat or a lot, 12.5% was neutral and 6.3 % said it did not help. Of the patients with symptoms lasting shorter than the median duration (5.7 years) (n=58 in the intervention group), less than half (43.1%) reported the intervention helped them somewhat or a lot, 17.2% was neutral and 12.1% said it did not help. Finally, of the patients who had 0-2 comorbid physical disorders (n=58 in the intervention group), 43.1% reported the intervention helped them somewhat or a lot, 19% was neutral and 9.6% said it did not help. Symptom duration, therefore, seems to play a more important role than comorbidity.

Future use

About a quarter of the patients (23%) said they would definitely recommend the intervention to a friend or family member with similar symptoms, 31% said they likely would, 27% said they might, and 12% said they would not. If patients would have many physical symptoms in the future again, 16% would definitely undergo the intervention with their MHNP again, 15% likely would and 29% might. More than a quarter of the patients (27%) said they thought they wouldn't follow the intervention again in the future and 4% definitely would not.

Withdrawal

30 of 111 intervention participants withdrew during the study versus 24 of 87 controls. 25 of the 86 intervention patients who completed the evaluation questionnaire dropped out of the intervention before the six sessions were completed. Twenty participants reported reasons for doing so. The most common reason for withdrawal was not gaining anything from the intervention, the intervention not being suitable for their symptoms, logistic reasons and not feeling comfortable having the sessions with the MHNP. However, most reasons for withdrawal were only mentioned once or twice.

Qualitative results (MHNPs only)

Themes

KS and SDK grouped the results of the interviews held with the MHNPs into 9 themes, based on the topics 'general', 'general practice setting', 'reach', 'effect', 'implementation' and 'maintenance'. The 9 themes were 1) motivation of MHNP and GP to participate in the study, 2) CIPRUS study, 3) routine care for USD, 4) patient eligibility and drop-out, 5) effectiveness of the intervention, 6) gains for MHNP, 7) facilitators and barriers for implementation, 8) MHNP adherence, 9) future use and recommendations. An overview of the relationship between topics and themes is provided in table 4. The coding tree is provided in Appendix E.

All of the participating patients were diagnosed with USD but were referred to as 'MUPS patients' or 'patients with MUPS' when communicating with the MHNP during the interviews, and are therefore referred to in these terms in the text below.

Table 4. Overview of topics and themes from interviews with MHNPs

Topic: General
Theme 1: Motivation of MHNP and GP to participate in the trial
Theme 2: The intervention and intervention manual
2a. Training, supervision and competencies of the MHNP
2b. Use of intervention manual
2c. Logistics
Topic: General practice setting
Theme 3: Routine care for MUPS
Topic: Reach
Theme 4: Patient suitability and drop-out
Topic: Effect:
Theme 5: Perceived effectiveness of the intervention
Topic: Implementation
Theme 6: Usefulness of the manual and gains for the MHNP
Theme 7: Facilitators and barriers for implementation of the intervention
Theme 8: Protocol adherence by MHNPs
Topic: Maintenance
Theme 9: Future use: recommendations

Abbreviations: GP: general practitioner; MHNP: mental health nurse practitioner; MUPS: medically unexplained physical symptoms

General

Theme 1: Motivation of MHNP and GP to participate in the trial

The majority of the MHNPs reported that it was a combined decision of both the MHNP and GP to participate in the CIPRUS study. Five MHNPs regarded it as their own choice and one MHNP said it was the GP's choice. However, although the GP initially decided on participation, she was still given the choice to participate or not.

The most common reason for MHNPs to participate in the trial was to obtain tools in treating patients with MUPS, as they perceived this to be a challenging group of patients.

Researcher (R): " …Why did you participate in the CIPRUS study?"
MHNP: "Umm, because MUPS is a challenging problem in the practice and umm, we all frequently don't know which way to turn, like God, what do we do with it? And then it just gets passed around a bit from one to another person, and yeah, it's hard! So any tool that we got, that we get, is great."

Other common reasons for participation were enhancing knowledge and personal development, because they wanted a new challenge and wanted to learn a new therapy format.

Theme 2: The intervention and intervention manual

2a. Training, supervision and competencies of the MHNP

MHNPs were satisfied with the amount and content of the training they received before treating patients. Ten MHNPs felt competent in delivering the treatment according to the manual. Reasons for feeling competent were being familiar with PST as a method and having worked with it before, the clarity and structure of the manual, the training MHNPs received and the fact that the MHNPs noticed that patients responded well to the intervention. One MHNP felt competent because of the rationale behind the intervention, that patients would get tools to handle their problems instead of trying to get rid of the symptoms. This empowered her in carrying out the intervention.

MHNP: *"Beforehand, I didn't have the impression that, I was going to cure these people because… I saw several people, who have had these problems for years and years, or, sometimes even decades. So I just thought, you know, I'm just going to try my best to carry this out and, I'm just going to go for it, I'm excited, I'm confident, but I shouldn't*

expect that within these 5, 6 sessions their symptoms will disappear. […] I'll just provide [the patients] with tools."

Three MHNPs did not feel competent in treating patients with the protocol at the start. Reasons for not feeling competent were having difficulty setting SMART goals (step 2 of the augmented PST steps in the intervention protocol) never having worked with this method before and having too little knowledge of MUPS. The latter was described by the MHNP as:

MHNP: "So I worked with the manual for about two months, meanwhile thinking... Yeah, I really don't feel confident enough."
R: "And with regard to what…"
MHNP: "Umm... Just some knowledge about MUPS. I had just literally never worked with it before. And then it's great that there's a manual but I just missed the whole background on MUPS patients."

In order to feel more competent, one MHNP practiced the text in the manual with family at home (since MHNPs were requested to follow the text as closely as possible). Receiving supervision sessions during the intervention and talking any difficulties over with the clinical psychologist also helped MHNPs make sense of the manual and apply it in practice. The clinical psychologist provided useful suggestions, such as not taking the text in the manual too literally, not overcomplicating the method and on how to motivate patients. One MHNP wished she had received the supervision session earlier. After the supervision session took place, the MHNP felt more competent, but by that time she had already completed a number of sessions.

2b. Use of intervention manual

Motivation

Nine MHNPs felt motivated to follow the intervention manual. The main motivation was that they knew they were participating in scientific research and the researchers had stressed the importance of following the manual for research sake during the training. Another was that the manual provided good guidance for the sessions. Two MHNPs started off being motivated, but became less motivated as the intervention continued, because they noticed that some patients either did not keep up with the sessions or that the intervention did not seem suitable for all patients. One MHNP was motivated to follow the manual in general, but not to read everything out loud because she felt it stood in the way of making contact with the patient.

Experience with the use of intervention manual

MHNPs generally found the intervention manual clearly written and understood what was expected from them during the sessions. However, some MHNPs felt that there was too much text to read out loud and that the explanation was too extensive.

MHNP: *"There were some things where I thought, I don't really think this is necessary. I thought the segments of text that you had to read out loud were waaaay too long. You just end up losing the patient. That's not good. They want to be the ones doing the talking. And that is also how I usually work."*

Especially with the first couple of patients, some MHNPs had difficulty to strictly follow the manual. They were not used to working with protocols and had developed their own working style during their careers. Some, therefore, experienced following the protocol and reading it out loud as challenging or disruptive. They sometimes felt that it created distance in the contact between MHNP and patient.

Several MHNPs noticed that if the manual was followed too strictly and the patient did not provide the answer they expected, the MHNP lost track of what they were supposed to say next, according to the manual. It was not always clear for the MHNPs where they could find the next step in their manual and several MHNPs got stressed about which answer option they were supposed to provide. A better structure for answer options was desirable. One MHNP felt that the texts were too complicated for the patient, especially for patients with a lower education level. The cognitive behavioural intervention itself (combination of the consequences model with PST) was fine, but the texts were too extensive, complex, formal and the MHNP was given too much floor.

On the other hand, the manual offered guidance and structure during the sessions. One MHNP normally conducted her sessions intuitively and was surprised by solutions and positive outcomes that she and the patient got from following the intervention manual.

MHNP: "So that was one of those things where I think 'Well, I would have really never have thought of this otherwise'. Then I would have definitely focused on the food a lot, with a dietician and you name it. But that wasn't it. It was, well…"
R: "The small practical solutions really."
MHNP: "Yes!"

By using this method the patient was the one put to work instead of the MHNP. The patient got more empowered and active to solve his/her own problems.

MHNP: "And I noticed that I barely heard the answer 'I don't know'. Because somehow the questions [in the manual] are formed that way, that people start thinking 'what could help me and what can't?' Because often if you ask [them] a question like "What do you think we could do about it?", [the answer you get is] "Yeah, I don't know, that's why I'm here!"
R: "Yes, exactly. And then they throw the ball back to you."
MHNP: "Yes. And I'm prone to that. And now it wasn't like that."

MHNPs who got used to the manual and did not blindly follow the text and developed their own routine, said they did not experience the reading out loud as disturbing. Not following the text in the manual rigidly made the contact between them and the patient grow.

One MHNP did not feel comfortable following the manual literally, but did follow the PST steps. However, she felt that the way the PST steps were described in the manual was too extensive, which lost the patient. Therefore, she used a simplified version of the PST steps that she had used with other patients before.

MHNP: "Well, what I noticed is, it's ambiguous. On the one hand it's very surprising if you work with the plusses and minuses. If you list the consequences [of the symptoms] for yourself and for the other, short and long-term, if you do that, then it's surprising what can come out of it. Which solution the patient eventually gets. But at the same time, it was also what turned off the patients. Because it was just too, too intense. They had to think too much about the consequences for themselves and for others, and of course that is important. That is also how I saw it, it is very important to think about it. But it was just too much".

Content of the intervention manual

One MHNP suggested that the number of sessions could be personalized according to the patient and not be set to strictly 6 sessions. MHNPs repeatedly reported finding 30 minutes too short for the content of the sessions. For the first two sessions they felt there was too much explanation, but found 30 minutes sufficient. They were able to get through session 2 quickly. Session 3, on the other hand, was too intensive for the 30 minute window, as in this session patients really started working with augmented PST and completing the forms. One MHNP did not manage to keep the interval of two weeks between the sessions due to her own and the patients' schedule. Another MHNP on the other hand, felt that the pace of the sessions was too slow. Because the first two sessions basically explain the method and the patient does not work on solving the problem yet, this MHNP felt that working on problem solving could have started sooner.

MHNPs generally found the content of the manual was clear, practical and efficient, but also containing a lot of repetition of explaining the method. However, this also varied per patient, as some patients needed an extensive explanation to understand the concept properly, whereas others understood it immediately.

MHNPs generally found our cognitive behavioural intervention to be a suitable technique for treating MUPS, that helped patients enhance their problem-solving abilities and look at their problems differently. Most MHNPs felt it activated patients to work on their problems and think out of the box to find new solutions. They felt it broke the patients' vicious circle of focusing on their symptoms and switched the focus to quality of life instead. One MHNP said she wouldn't use our cognitive behavioural intervention with these patients again.

MHNP: "I think that it's [the intervention] too structured. I think that with depression, that patients require more structure, and to go through the steps, in order to move towards a new goal. And with MUPS patients, first, they need to be satisfied in their need to be heard and seen. Because of all the [medical] specialists and doctors, and, whoever, people just don't feel understood".

Some MHNPs noted that patients could relate to sentences in the text about MUPS and its consequences. One MHNP found the manual more casual and light-hearted. She reported:

MHNP: "What I noticed with patients is that it actually stayed very light-hearted, and that they weren't really used to that. That's really positive actually. That it's not so serious, and very, well, problem-solving of course. It's really VERY different to what we normally do. We really focus on the symptoms and 'Oh that must be so hard!' and 'Oh dear!' And here of course you do the same at the beginning, but then you immediately turn to 'OK, what CAN we do?' And that is in fact a lot less exhausting, also for us".

Some, however, had difficulty with the text describing the steps in the manual and did not always feel that the manual fitted every type of patient. Patients who got the hang of the method were able to go through the sessions more quickly compared to those who hadn't.

MHNP: "Quickly moving on to implementing the things on their own, that was not achievable for everyone".

Suitability of the intervention for patients varied depending on factors such as patients' age, level of education, severity and duration of symptoms and comorbidity with other psychological problems. One MHNP felt that the intervention was maybe just not enough for more severe cases. However, at the same time she felt that particularly for patients with severe symptoms that lasted longer, it was possibly important to offer a simple and straight-forward method, in order not to get lost in the complexity of their symptoms.

2c. Logistics

The majority of the MHNPs thought it was a pleasant and informative experience to participate in the research and they were satisfied with the support from the research team.

Eight MHNPs invited patients for sessions themselves, which was not common practice outside of the study. In three practices the GP receptionist invited the patients, in one practice the MHNP and the GP receptionist both invited the patients and in one practice patients were expected to call in themselves to make an appointment. MHNPs who planned their appointments themselves or with the aid of the GP receptionist, did not encounter any problems. In fact, some reported that they preferred to have spoken with the patients before seeing them during the first session. The MHNP in whose practice the patients were expected to plan the appointments themselves, reported that there were generally no problems with this, but that patients did not always mention that they were making the appointment for the CIPRUS study, which sometimes ended up being confusing for the patient and GP receptionist and required more input from the patient. However, because the receptionists were aware of the ongoing trial, they were often able to figure out what the appointment was for.

One MHNP reported that it cost her a lot of effort to schedule patients from the CIPRUS study in between the other regular patients.

One MHNP started seeing the first patients several months after she completed the training. She felt that this period was too long, as some of the things she had learned during the training were forgotten by the time she saw her first patient. She also reported that not all patients seemed to be well informed about the study and did not know exactly what to expect from the sessions with the MHNP.

Five randomly selected MHNPs were asked to audiotape their sessions, so that the researchers could check adherence to the protocol and identify possible problems that

arose during the sessions. Audio-recording did not generated any problems, but MHNPs were conscious of the fact that recordings were being made and this needed some getting used to. They did not feel that they acted differently because of the recorder.

General practice setting

Theme 3: Routine care for MUPS

In 5 of the intervention practices, GPs sometimes referred patients with MUPS to MHNPs as part of routine care. In 4 practices, MUPS patients were either not routinely recognized as such and ended up having an appointment with the MHNP due to various complaints for which the GP could not find a solution and suspected a psychological cause, or they were referred to the MHNP with other comorbid psychological complaints such as 'tension/stress complaints'. In two intervention practices MUPS patients were not referred to the MHNP as part of routine care.

MHNP: "Before the trial started, we didn't really talk about MUPS patients [in the surgery]. People came with depression, or fatigue, or physical complaints. And it was never called MUPS. And that came to mind because of the trial actually. I had also never put the two and two together, with the patients that I had been seeing. And only when I saw the lists, and the patients who were going to start the trial, I thought 'Oh yeah! These are MUPS of course."

The GPs who did not refer MUPS patients to MHNPs as part of routine care, typically saw MUPS patients themselves. MHNPs normally saw patients with light symptoms and, after consulting the GP, referred patients with more severe symptoms to a psychologist, MUPS outpatient center or medical centers.

Reach

Theme 4: Patient suitability and drop-out

Patient suitability

MHNPs felt that the selection procedure for the trial did not always yield suitable patients and that GPs should also have played a role in estimating whether the patient could work with the manual. MHNPs estimated that on average about 35-40% of patients were not suitable for the intervention, although this varied across MHNPs. Reasons for not being suitable for the intervention varied. The most common reasons were possible comorbid personality disorders, comorbidity with other psychosocial problems and focus on seeking a medical cause for the symptoms.

MHNP: "Well, I had one patient, who also had umm complex PTSD and ADHD besides MUPS, so then you already think 'OK, and now I'm going to treat this person with a very strict protocol and give her homework, that's never going to work'. "

MHNP: "And one patient just didn't accept the fact that there was no explanation for the symptoms and constantly kept seeking medical help. Even though you try not to focus on that, despite of that, [the patient] kept making all these appointments with, neurologists and whatever, but finished the sessions though. And also achieved good results. So that went well actually, but I thought [the patient] was still so much in the process, and needed to reach acceptance first".

Other reasons were for not being suitable were older age, level of intelligence (too low or too high), having sufficient problem-solving tools (from the view of the patient), externalizing, negative and unmotivated attitude towards the treatment and having other expectations of the intervention.

MHNP: "And of course you also see that with older people physical symptoms occur more often. And then the question is, of course, is it MUPS or age related? That someone has a worn out hip that's causing a lot of pain, or that there is another physical symptom that hasn't been recognized, And, well, I noticed that with a number of elderly people, who haven't learned to talk about things, [they] have a lot of respect for the doctor. And because of that [they] are less assertive in saying what's really bothering them. […]"
R: "Do you think they expect a medical solution?"
MHNP: "Umm… Well, an improvement, but then a physical one. And not by putting any effort in themselves".

Patient motivation

When asked whether patients had a demand for care, MHNPs frequently first answered no, but then said that as the sessions carried on, patients did develop a demand for care. However, this did not happen with all patients. Specifically patients who were not suitable for the intervention or who discontinued the intervention, were the ones who did not have a demand for care of their own, but participated in the study only because they were invited to do so.

R: "And did the patients have a demand for care related to MUPS?'
MHNP: "Well, the ones who dropped out didn't. They were screened and got an invitation, and they sort of started [the intervention] like, well, it doesn't hurt to try, so we'll see what happens."

R: "Did they feel a little pushed because of the study, or…?"
MHNP: "Umm… Yes, and there was a mix of curiosity and finding it nice to be able to talk about it and, well, who knows, maybe there's a solution".

On the other hand, people who did finish the intervention seemed to have a demand for care.

R: "And the people who finished the intervention? Do you feel that they had a demand for care to deal with their MUPS?"
MHNP: "Yes! Because they really had clear problems that they encountered. And that could have been something simple like, for example, I can't make my bed anymore, my husband has to do everything, and he's already so busy with work and this and that. And well, then you start with the little items in the household… and then you reach activation, then they do start doing it and ask for more help and… So they really had very specific things, which they couldn't do anymore because of their MUPS."

Drop-out

Nearly half of the patients dropped out from the intervention before the 6 sessions were over. The most common reasons for dropping out, according to the MHNPs, were similar to the reasons for not being suitable for the study, i.e. focus on medical explanation, comorbidity with personality problems/disorders and already having sufficient problem-solving skills. MHNPs also reported that patients felt the intervention was physically too burdensome, that they did not seem to understand the manual and that it was too overwhelming and too much paperwork for them.

MHNP: "There were some people who totally panicked because of all the things they had to do. They got the pile of papers, I gave it to them, homework too, to get started with. Well, and then, they actually got confused immediately."

Other reasons were having more complaints from other explained symptoms, having a decrease and an increase in symptoms. Also patients' lack of motivation and demand for care, and attitude and expectations towards the treatment played a role in dropping out.

MHNP: "And some people just expected other things. Expected more depth, expected more […] more like psychotherapy. More sessions in dealing with, a little acceptation and umm… That is why they dropped out I think, why they did not have a specific demand for care".

MHNP: " 'I already do everything', 'I don't have any problems in that area'. That's also one of those quotes, 'I already do everything'. And then they can't think of the things that they can't do."

Final reasons for dropping out were logistic problems such as, being too busy, a longer stay abroad, moving, switching to another treatment elsewhere and not willing to receive treatment from a new MHNP.

Effect

Theme 5: Perceived effectiveness of the intervention

According to the MHNPs, the patients who were suitable for the study and completed the sessions seemed to profit from the intervention.

MHNP: "Yes, I have the feeling that they did get some tools to be more problem-solving in life. And that that also helped them. They became, like I said in the beginning, they became less demoralized, and saw some more opportunities to come out of the circle of negativity and pain, and physical symptoms. So, I think it helped, yes".

Several MHNPs felt that even patients who did not complete the intervention also profited from the sessions somewhat, by becoming more aware of the fact that you can do something about the symptoms and that they don't necessarily have to impair functioning.

MHNPs also reported that the demand for help was met for patients who were suitable for the study and completed the sessions. Some MHNPs wondered whether this effect was only temporary:

MHNP: "It was effective for both [patients]. They both felt better. One had a chronic migraine and deals really well with it now. And the other one had stress-related symptoms, and started looking at life in a really different way. Yes. Whether that's still the case, that's the question. But then it definitely was effective temporarily ."

Another MHNP wondered whether the request/demand for help was met sufficiently:

R: "Was their request for help met by this manual?"
MHNP: "Well, yes and no. So, people who completed it [the intervention] and actually reached their goals, I think that their request for help was met. But, whether they've all

come to accept the fact that they can't do some things anymore, or that they're not in need of more treatment, I doubt that a little."

MHNPs reported that patients also voiced that the intervention helped them:

MHNP: "There were several people who were really very positive, several people a little less, but they felt like: "Well, you know, it did give me insight and brought me something. Maybe it's not completely the way I would do it , but it definitely gave me greater insight and brought me something".

Reasons for perceived effectiveness of the intervention were patients' cognitive skills such as being able to reflect on themselves, gathering insight into own behaviour, aspects from the intervention protocol and factors unrelated to the intervention or the manual. Reasons mentioned for the intervention not being effective were: patients being too chaotic, patients' comorbidity, psychosocial problems and low IQ, other explained physical symptoms prevailing and other expectations of patients.

Implementation

Theme 6: Usefulness of the manual and gains for the MHNP

Usefulness of manual for the MHNP

All but one MHNP said that the manual was useable. Main reasons for this were that the manual provided structure and clarity for the sessions. Two MHNPs would have liked the manual to be more flexible and to incorporate space for their own input during the sessions. One MHNP said the manual was fairly usable. She felt that the cognitive behavioural intervention was a good technique, but the texts describing the sessions were not suitable for the patient she treated.

Specific gains for MHNPS

When asked whether the manual provided specific gains for the MHNP to deal with patients with MUPS, all but one MHNP answered positively. One of the gains that MHNPs gathered was a better proficiency in PST.

MHNP: "Because what I experienced is that usually with MUPS patients I think 'I'm going to refer them, because I don't really know what to do with them'. That's the feeling I get. I'm particularly good in discovering peoples' behaviour together with them and […] and changing it step by step, and playing around with that a little. But this, provided met with a tool now. To see how are we going to do this, how are we going to make sure that you

can start doing other things? […] So yes, it brought me new things. And also not to look, not to pause at 'Where did it originate?' That's not important at all. Yes, that, for me, I am a nurse, so I'm very much into here and now. So I liked that".

Another gain MHNPs reported was that they became more problem-solving themselves:

MHNP: "My way of thinking actually changed, much more towards problem-solving. Like, oh, that's terrible, but how shall we move on?"

A third gain reported by the MHNPs was that the manual was practical, which also shifted the focus from possible explanations for the physical symptoms and negative emotions, to pragmatically solving the problems that the symptoms caused.

Other gains reported by the MHNPs were that the manual helped them and the patients set realistic and achievable goals and helped MHNPs provide psychoeducation. Finally MHNPs reported that the manual helped them stay structured during the sessions and also allowed the patients to have greater insight into what the sessions and the intervention were built up of. MHNPs reported that this gave patients a clear guideline of what they were doing and it was easier to guide the sessions.

Theme 7: Facilitators and barriers for implementation of the intervention

Facilitators for implementation

The most commonly mentioned facilitator for implementation within the practice was that MHNPs had sufficient time to see patients and prepare the sessions, without feeling pressured by the GP to see as many patients as possible during the day. Other commonly mentioned facilitators were that MHNPs were either able to plan their own appointments and manage their own agenda, or that all practice staff members took the CIPRUS study patients into consideration when planning appointments for the MHNP.

Other facilitators were the working space of the MHNP, the fact that the GPs were motivated to participate in the CIPRUS study and referred eligible patients to the MHNP, that the surgery was a familiar place close to the patients' homes, and that patients were familiar with the employees and trusted them. In one practice patients were informed by a poster in the waiting room that the surgery participates in various scientific projects, making the patients somewhat prepared for this.

Barriers for implementation

Most MHNPs reported there were no barriers for implementation within the surgery where they worked. However, a few did mention some barriers. The most frequently mentioned barrier was the MHNPs' limited time. MHNPs generally had busy schedules and therefore did not always succeed in scheduling appointments every two weeks as was recommended in the manual. Also MHNPs happened to work in multiple surgeries, and therefore only worked part-time in a certain surgery. This contributed to a busier schedule due to the CIPRUS study patients. For one MHNP barriers were that she couldn't veer from the protocol and that she didn't see enough patients to master the protocol.

MHNP: "I eventually saw 3 or 4 [patients], so you just don't get into the rhythm of implementing the protocol. Because it's spread out over a longer period of time too. You know the main outlines, but sometimes you have to dive into it again, and then if you're in a session, on a day with 10-12 sessions, then you also sometimes have to do some work for the other patients as well. And then you don't take all the time for it".

Factors deemed necessary for successful implementation in the future

Most surgeries already had everything they needed for implementation at their disposal. However, some MHNPs mentioned that they needed a form of supervision in the surgery and interaction and communication with MHNPs from other surgeries delivering this intervention to share their experiences. Other factors mentioned were adjustments to the manual to make it more accessible for the MHNP and patient, and availability of the manual online or as an e-health module, since paper is being used less and less in surgeries, informing the GP thoroughly about the selection process of eligible patients for this intervention, and the GP making the first step in informing the patients about the intervention were also deemed necessary for successful implementation. A final factor mentioned was scientific evidence for effectiveness of the intervention.

Theme 8: Protocol adherence by MHNPs

MHNPs felt it was important to adhere to the manual due to the fact that they were participating in scientific research. Four MHNPs reported having completely adhered to the manual, whereas the remaining nine said that they deviated from the manual somewhat. All MHNPs reported having followed the main outline of the manual. Two MHNPs reported having difficulty following the manual, because they did not feel the freedom to explore other topics they deemed important. MHNPs especially had difficulty adhering to the manual with patients who were chaotic, talked a lot or had a lower IQ.

MHNP: "Yes, she has ADHD. She talks a lot, she keeps getting sidetracked. And that umm, well of course you have to put restrictions on that, but it would have been nice if people could have been given some space for their own individuality."

MHNPs typically deviated from the manual by adjusting the number of sessions or adjusting the pace of the content of sessions, for instance by splitting the content of a session into two sessions, or giving a part of the session as homework. They also had somewhat longer sessions with the patients. This made them feel less rushed and have better contact with the patients. They also used their own words. One MHNP started using the text from the manual, but felt that patients did not follow her, so she started using her own words and used a less extensive version of PST steps that she was familiar with from a previous study. This less extensive version does not investigate the short-term and long-term consequences for the patient and for others. She also used the first session to explore a patients', what she called, 'biographic sketch', which included one's demographic background, but also questions about one's past, relationships and history of the complaints. According to the MHNP this made her patients feel taken seriously and understood.

Maintenance

Theme 9: Future use: recommendations

Recommendations for future use

Nine MHNPs would use the manual, or elements of it, in the future. Three MHNPs would not. Those who would use the manual again, said they would not like to follow it as strictly as during the study and that they would like to make their own adjustments. One MHNP said she would like to add some more theory and psychoeducation about MUPS.

MHNP: "There are good elements in there. And I still give the forms to the people. The questions that are asked [in the forms] are good. And, I think it's useful to invite people to look at their own behaviour. The sort of function analysis that is in there. What's the behaviour, what you do, and, what your tendency is, and what long-term effect does it have, for your environment and for yourself. And advantages, disadvantages. It's just useful. No, I still use it".

MHNP: "It's just a really clear manual. As long as I can just give my own twist to it, I would definitely use it. And definitely the questions that can lead to a clear goal, to a SMART goal, are very useful".

Three MHNPs reported having used elements of the manual with patients with explained symptoms and anxiety outside of the study.

The reasons of the three MHNPs who would not use the manual again were 1) preference for other treatment techniques such as Acceptance and Commitment Therapy (ACT), 2) that following the manual was too strict, and 3) because one MHNP was already sufficiently familiar with PST and used the short version when deemed necessary.

Recommendations for future use of the manual were personalizing the number of sessions and pace of sessions to the patient. MHNPs were happy to learn the cognitive behavioural intervention and thought it was a sound method for treating MUPS, but felt that it was just a part of treatment and would like to use other methods in combination or in some cases instead of our intervention. Also, MHNPs recommended not necessarily using the entire extensive version of the PST steps, but that sometimes a more simplified version would do.

MHNP: "I'm on board with it a little, [the intervention] but [I'm] much more into talking. And also saying 'Well, think about it, go explore things', but not really very structured: short-term consequences, long-term consequences, consequences for you, consequences for the others. Sometimes people found it useful, but sometimes childish as well, the plusses and minuses, to work it out like that. I'd skip that".

MHNPs suggested combining session 1 and 2, and giving the rest of the sessions 45 minutes each due to their content.

One MHNP recommended to have an intake before the sessions, for the MHNP and patient to get acquainted. Two other MHNPs suggested providing more medical education to the patients about MUPS and one suggested doing activating exercises like going outside and walking or running with the patient. Another MHNP suggested placing the manual in the GPs' electronic e-health database to offer it as an e-health program.

Ideal treatment for MUPS

An almost unanimous opinion was that all MUPS patients are different and that treatment for MUPS should therefore be personalized, rather than applying the cognitive behavioural intervention blindly to all types of patients. MHNPs felt it was important to discover more about one's symptoms, for instance, whether there is unprocessed trauma or that someone has negative cognitions. Treatment could then be personalized depending on these outcomes.

MHNPs mentioned that ideally a combination of treatment methods should be used. Frequently mentioned useful therapies, to combine our intervention with, were psychoeducation and Acceptance and Commitment Therapy (ACT) in order to accept the fact that the symptoms are unexplained and that functioning has changed. Two MHNPs suggested a multidisciplinary approach involving a collaboration between the MHNP, a psychologist, a (psychosomatic) physiotherapist and a social worker in the living area of the patient.

MHNP: "So it's the combination of factors, that makes me think yes, that is something that I think will eventually…, if you eventually succeed in completing the whole picture, then you can offer people proper guidance".

Another MHNP mentioned that the treatment should also take the patients' ethnic background into account, as people from non-Western cultures may manifest their (psychological) symptoms differently. Again physical activation and e-health were mentioned as parts of an ideal treatment.

MHNPs also suggested that patients with more severe MUPS or comorbid personality disorders should be referred to other primary or secondary health care settings and that they would prefer to see the less severe patients themselves. One MHNP preferred not having any MUPS patients referred to herself as she did not really have an affinity with this group of patients.

DISCUSSION

Main findings

General

This study evaluated the MHNPs' and patients experiences with participation in the intervention arm of the CIPRUS study. MHNPs thought that the intervention was a sound method for treating patients with MUPS. Some would have liked to personalize the intervention and incorporate other treatment techniques and spend more time on building a relationship with the patient when deemed necessary.

According to the MHNPs, patients were not always sufficiently informed about the steps in the study by the researchers, despite having signed an informed consent. In the future, more attention should be paid to informing patients about what participation in the study and the intervention entail and preparing them thoroughly for every step, so they would know what to expect and what is expected of them.

Reach

Although patients had at least mild symptoms according to the PHQ-15, fulfilled the DSM-IV criteria for undifferentiated somatoform disorder and their GPs confirmed that their symptoms were medically unexplained, MHNPs reported that at least a third of the patients was not suitable for the intervention. They thought that the intervention was not suitable for patients with comorbidity (especially personality disorders and trauma), psychosocial problems and too low or too high intelligence. Furthermore, a substantial proportion (nearly half) of the patients stopped the sessions before the end either because they had reached their goals or did not wish to participate anymore. Patients who were not suitable for participation according to the MHNPs were also the ones who had dropped out. In the MHNPs' opinion, excluding patients with comorbid personality disorders, older age (above 70 or 80), low IQ and patients with sufficient problem-solving skills would have contributed to a more successful treatment. Furthermore, MHNPs could play a role in selecting the patients for such a treatment alongside the GPs. An intake with an MHNP could possibly be useful for the MHNP to estimate the patient's eligibility.

Nearly half the patients dropped out before the end of the intervention. Reasons reported by patients were not gaining anything from the intervention, the intervention not being suitable for their symptoms, logistic inconveniences and not being able to get along with the MHNP. MHNPs, however, also reported comorbidity and psychosocial problems as reasons for drop-out among patients. The drop-out rate was higher than we expected beforehand. Although drop-out rates are usually high in similar trials for medically unexplained symptoms (6), we hoped the drop-out rate would be limited due to the familiar general practice setting. This did not seem to be the case. An explanation for this could be that participating patients were recruited by a case-finding procedure. Only about a quarter of the patients indicated having had the need for help with their symptoms when they were approached for participation. The rest of the patients may therefore have been less motivated.

Effect

MHNPs felt that the intervention was effective for the people whom they deemed suitable. MHNPs got the impression that patients acquired problem-solving tools and helped them deal with their symptoms better.

A little more than half of the participating patients felt that the intervention helped them at least somewhat. An interesting finding was that patients with a shorter symptom duration (shorter than 2 years) were also more likely to report that the intervention helped

them. This is in line with the findings from chapter 5, and shows that our intervention is more suitable for this group of patients.

Implementation

The majority of the MHNPs reported that the intervention provided them with specific tools for treating patients with MUPS such as becoming more proficient in PST and acquiring a more problem-solving and practical attitude themselves.

MHNPs generally adhered to the manual by following the steps, but deviated from the text. They described the texts as too lengthy, formal, repetitive and sometimes complicated. Also reading out loud was sometimes an obstacle. The sessions of the intervention and the interval between sessions were too short, therefore many MHNPs adjusted the duration and pace of the sessions. There were few barriers to implementation, the main barrier mentioned was the time of the MHNP. Facilitators were the cooperation of the GPs, proper coordination within the surgery and the working space of the MHNPs.

Maintenance

In the end, most MHNPs (69%) said they would use (elements of) the intervention again in the future. They recommend e-health and more training for MHNPs to gather knowledge about MUPS in order to implement the intervention. Less than half the patients would follow the intervention again in the future and nearly half would recommend it to someone else.

Link with trial results

These findings fit in with the results of our trial that showed that patients with symptoms that had lasted shorter and/or with only few comorbid physical diseases benefitted from the intervention, whereas patients with symptoms that lasted longer and with a greater number of comorbid diseases did not (13). However, participating MHNPs particularly stressed psychological or psychiatric comorbidity as barriers, whereas the trial showed that a higher degree of somatic comorbidity was related to a smaller or absent effect of the intervention. In analyses on the effectiveness of the trial, no differences were found for patients with and without comorbid psychiatric disorders. A possible explanation for this discrepancy could be that patients were asked to report on psychiatric comorbidity such as personality disorders, to which MHNPs referred, in self-report items. Patients were simply asked whether or not they had a particular (personality) disorder. We did not use a validated scale to measure this type of psychiatric comorbidity. Therefore, it could be possible that patients were not aware or did not report their own psychiatric comorbidity on these self-report items. On the other hand, it could be that MHNPs

referred to psychiatric comorbidity when they actually meant certain personality traits that stood out to them. Personality traits were not assessed in the trial.

Strengths and limitations

A strength of this study is that we invited all available MHNPS who carried out the intervention to participate in the interviews and that nearly all (13/15) actually participated. We also incorporated the views of the patients to provide an additional perspective. Another strength is the use of theoretical frameworks such as RE-AIM in setting up interviews for MHNPs and questionnaires for patients. We therefore conducted a thorough evaluation of the intervention based on various relevant dimensions.

A limitation is that we did not conduct additional interviews with the patients to try to integrate the data sources. One of the researchers conducting the interviews with the MHNPs was also the main researcher of the CIPRUS-study. Therefore, it is possible that this has led to socially desirable responding from the MHNPs, or a biased interpretation of the information from interviews by the researcher. However, before MHNPs completed the evaluation questionnaire and participated in the interview, the researchers emphasized the importance of sincere feedback, in order to help the researchers improve the intervention. During the interviews MHNPs provided researchers with what seemed to be open and critical feedback. Additionally, another researcher coded and interpreted all the interview data as well, in order to minimize bias.

CONCLUSIONS

Our new intervention is a useful addition for MHNPs who help MUPS patients. It provided MHNPs with tangible tools for treating this patient group and was an effective method according to the majority of MHNPs and patients. The intervention was particularly easy to implement with patients with mild symptoms, with no or little comorbidity or other complicating factors. The manual should perhaps not be too structured and should give MHNPs space for their own interpretation and personalization. The intervention can be optimized by a more personalized approach with room for incorporating other techniques according to the MHNPs.

ETHICAL APPROVAL AND CONSENT TO PARTICIPATE

This study was approved by the VU Medical Centre Ethics Committee on 9 July 2014, reference number 2014.305. It is conducted according to the principles of the Declaration of Helsinki (version 2013). Informed consent was obtained from all participants.

REFERENCES

1. Verhaak PF, Meijer SA, Visser AP, Wolters G. Persistent presentation of medically unexplained symptoms in general practice. Fam Pract. 2006;23(4):414-20.
2. De Waal MWM, Arnold IA, Eekhof JAH, Van Hemert AM. Somatoform disorders in general practice. Br J Psychiatry. 2004;184:470-6.
3. van der Weijden T, van Velsen M, Dinant G, van Hasselt CM, Grol R. Unexplained complaints in general practice: prevalence, patients' expectations, and professionals' test-ordering behavior. Med Decis Making. 2003;23:226-31.
4. Haller H, Cramer H, Lauche R, Dobos G. Somatoform disorders and medically unexplained symptoms in primary care. Dtsch Arztebl Int. 2015;112(16):279-87.
5. APA. Diagnostic and statistical manual of mental disorders, 4th ed. (DSM-IV). Washington DC: American Psychiatric Association; 1994.
6. van Dessel N, den Boeft M, van der Wouden JC, Kleinstäuber M, Leone SS, Terluin B, et al. Non-pharmacological interventions for somatoform disorders and medically unexplained physical symptoms (MUPS) in adults (Review). Cochrane Database Syst Rev. 2014;11(CD011142).
7. Mack S, Jacobi F, Gerschler A, Strehle J, Hofler M, Busch MA, et al. Self-reported utilization of mental health services in the adult German population--evidence for unmet needs? Results of the DEGS1-Mental Health Module (DEGS1-MH). Int J Methods Psychiatr Res. 2014;23(3):289-303.
8. Murray AM, Toussaint A, Althaus A, Lowe B. The challenge of diagnosing non-specific, functional, and somatoform disorders: A systematic review of barriers to diagnosis in primary care. J Psychosom Res. 2016;80:1-10.
9. Salmon P, Dowrick CF, Ring A, Humphris GM. Voiced but unheard agendas: qualitative analysis of the psychosocial cues that patients with unexplained symptoms present to general practitioners. Br J Gen Pract. 2014;54:171-6.
10. Crimlisk HL, Bhatia KP, Cope H, David AS, Marsden D, Ron MA. Patterns of referral in patients with medically unexplained motor symptoms. J Psychosom Res. 2000;49:217-2019.
11. Magnee T, de Beurs DP, de Bakker DH, Verhaak PF. Consultations in general practices with and without mental health nurses: an observational study from 2010 to 2014. BMJ Open. 2016;6(7):e011579.
12. Verhaak PFM, Nielen M, de Beurs D. Factsheet 1: [Demand for care - aid regarding psychological and social problems in the general practice 2011-2016] (in Dutch). Utrecht: Netherlands Institute of Health Services Research (NIVEL); 2017.
13. Sitnikova K, Leone SS, van Marwijk HWJ, Twisk J, van der Horst HE, van der Wouden JC. Effectiveness of a cognitive behavioural intervention for patients with undifferentiated somatoform disorder: Results from the CIPRUS cluster randomized controlled trial in primary care. J Psychosom Res. 2019:109745.
14. Sitnikova K, Leone SS, Zonneveld LNL, van Marwijk HWJ, Bosmans JE, van der Wouden JC, et al. The CIPRUS study, a nurse-led psychological treatment for patients with undifferentiated somatoform disorder in primary care: study protocol for a randomised controlled trial. Trials. 2017;18(1):206.
15. Sitnikova K, Finch AP, Leone SS, Bosmans JE, van Marwijk HWJ, van der Horst HE, et al. A brief cognitive behavioural intervention is cost-effective for primary care patients with medically unexplained physical symptoms compared to usual care. Manuscript submitted for publication. 2019.
16. Kroenke K, Spitzer RL, Williams JW. The PHQ-15: validity of a new measure for evaluating the severity of somatic symptoms. Psychosom Med. 2002;64:258-66.
17. First M, Spitzer R, Gibbon M, Williams J. Structured clinical interview for DSM-IV-TR axis I disorders, research version, patient edition (SCID-I/P). New York: Biometrics Research, New York State Psychiatric Institute; 2002.

18. Keijsers G, van Minnen A, Hoogduin K. [Protocollized interventions for adults with psychological symtpoms part 2] (in Dutch). Amsterdam, the Netherlands: Boon; 2011.
19. Zonneveld LN, van Rood YR, Timman R, Kooiman CG, Van't Spijker A, van Busschbach JJ. Effective group training for patients with unexplained physical symptoms: a randomized controlled trial with a non-randomized one-year follow-up. PLoS One. 2012;7(8):e42629.
20. Mynors-Wallis L. Problem-solving treatment for anxiety and depression: a practical guide. Oxford and New York: Oxford University Press; 2005.
21. Glasgow RE, Vogt TM, Boles SM. Evaluating the public health impact of health promotion interventions: the RE-AIM framework. Am J Public Health. 1999;89:1322-7.
22. Grant A, Treweek S, Dreischulte T, Foy R, Guthrie B. Process evaluations for cluster-randomised trials of complex interventions: a proposed framework for design and reporting. Trials. 2013;14(15).
23. Saunders RP, Evans MH, Joshi P. Developing a process-evaluation plan for assessing health promotion program implementation: a how-to guide. Health Promot Pract. 2005;6(2):134-47.
24. Pope C, Ziebland S, Mays N. Qualitative research in health care: analysing qualitative data. BMJ. 2000;320:114-6.
25. Tong A, Sainsbury P, Craig J. Consolidated criteria for reporting qualitative research (COREQ): a 32-item checklist for interviews and focus groups. Int J Qual Health Care. 2007;19(6):349-57.

Appendix A. Evaluation questions for MHNPs

To critically evaluate the manual, we would like to ask you some additional questions.
Please complete the following information:
Background:
Psychologist / Social worker / Social psychiatric nurse / Other:
How many years have you been working as a MHNP?

_ _
How many years of other relevant working experience in mental healthcare do you have?

_ _
Please select a number that represents your opinion the best for each question below.

1. How satisfied are you with the CIPRUS study manual?

Not satisfied at all					*Very satisfied*
0	1	2	3	4	5

2. To what extent were the patients that you saw for the CIPRUS study suitable for this study?

Not at all					*Completely*
0	1	2	3	4	5

3 .Did the manual help patients to deal with their MUPS better?

Not at all					*Yes, very much*
0	1	2	3	4	5

4. Did the manual provide you with tools for dealing with MUPS patients?

Not at all					*Yes, very much*
0	1	2	3	4	5

5. Did you adhere to the manual?

Not at all					*Yes, very much*
0	1	2	3	4	5

6. Would you use the manual again with patients with MUPS in the future?

No, definitely not					*Yes, definitely*
0	1	2	3	4	5

Appendix B. Topic list

Reach:

- MHNPs' vision on suitability of patients for participating in the intervention
- MHNPs' vision on the patients' demand for care

Efficacy:

- How useful was the intervention for the MHNP?
- Effectiveness of the intervention concerning the patients' demand for care, according to the MHNP

Implementation:

- MHNPs' adherence to the manual
- Did patients follow all 6 sessions
- Barriers and facilitators in following the manual
- Unexpected events during implementation

Maintenance:

- Usefulness for the future and barriers and facilitators for implementation in the future

Appendix C. Interview guide for MHNPs

General:

- Why did you participate in the CIPRUS study? Was it your or the GP's decision?
- What did you think of the CIPRUS study program and manual?
- How did you feel about using the manual?
- How did you feel about the content of the manual?
- Do you think that this cognitive behavioural intervention is a useful technique for treating MUPS?
- What did you think about the duration of the sessions? Were they too short/too long? Did you manage within the given time?
- Did you feel competent enough to deliver the treatment? Why?/Why not?
- Were you motivated to work with the manual? Why?/Why not?

Context (general practice setting):

- Does the GP normally refer MUPS patients (outside of the CIPRUS study) to you?
- Who planned appointments for sessions with the patients for the CIPRUS study? (MHNP/GP receptionist/GP/did patients have to contact the practice themselves?) How did that go?

Reach:

- Did you find the participating patients suitable for such a treatment? How many were suitable and how many were not?
  - Symptom severity
  - Comorbidity
  - Level of functioning and quality of life before starting treatment
  - Insight into psychological factors/fixation on medical explanation
- Did the patients have a demand for care concerning dealing with their MUPS (or other symptoms)?
- Did patients quit the intervention before the 6 sessions were over? What was the reason for this?

Effect:

- Was the patient's request for help answered by the CIPRUS study?
- Do you feel that the patients in the CIPRUS study learned to deal better with the consequences of their MUPS? Why?/Why not? How many did and how many did not?

Implementation:

- What factors in your practice benefited the implementation of the CIPRUS study?

- What are the most important barriers for implementation of the CIPRUS study in your practice?
- Was the manual useful to you as a healthcare provider? If no, why not?
- Did you deviate from the manual? What were the reasons for this (e.g. choosing for another treatment type, referral, frequency of appointments, duration of sessions)? How did you deviate?
- Did the manual offer you tools for dealing with MUPS patients? If yes, which tools?
- Only if applicable: How did you feel about the audio recordings? Did it bother you?

Maintenance:

- Would you use (parts of) the manual in the future? Why?/Why not?
- What is needed to implement the program in your practice?
- How would you describe the ideal treatment for MUPS patients? Referral?

Appendix D. Evaluation questions for patients

1. Why did you decide to participate in the CIPRUS study?
    a) I wanted to deal with my physical complaints
    b) I wanted to get rid of my physical complaints
    c) I felt a need for support in dealing with my physical complaints
    d) It was recommended to me (e.g. by the GP)
    e) I felt it was important to participate in scientific research
    f) Other reason: ……

2. To what extent did your physical complaints influence your life negatively beforehand?
    Not at all / A little bit / Somewhat / A lot / Extremely

3. To what extent did you feel the need for support in dealing with your physical complaints before the CIPRUS study?
    Not at all / A little bit / Somewhat / A lot / Extremely

4. Did you follow al 6 sessions of the individual training?
    Yes/No

    4.1 If no, what was the reason for this? (open question)

5. What did you think of the amount of sessions that you had with the MHNP?
    Way too many / A little too many / Sufficient / A little too few / Way too few

6. What did you think of the quality of the individual training that you followed?
    Excellent / Good / Medium / Bad / Very bad

7. Did the individual training meet your expectations?
    All my expectations were met / Most of my expectations were met / Some of my expectations were met / Only a few of my expectations were met / None of my expectations were met

8. Did the individual training help you deal better with your physical complaints?
    Yes, it helped me a lot / Yes, it helped me somewhat / Neutral / No, it did not help me / No, it aggravated my complaints

9. Did you use the workbook during the individual training?
    Yes/No

    9.1 If no, why not? (open question)

10. Did you find the workbook useful?

Extremely / A little / Neutral / Not really / Not at all

11. How satisfied are you in general with the individual training you received?

Very satisfied / Somewhat satisfied / Not satisfied, not unsatisfied / Somewhat unsatisfied /Very unsatisfied

12. Imagine that someone you know happens to have unexplained physical complaints, would you recommend this individual training?

Yes, definitely / Yes, I think so / Maybe / No, I don't think so / No, definitely not

13. Imagine that you encounter unexplained physical symptoms again in the future, would you follow this training again with your MHNP?

Yes, definitely / Yes, I think so / Maybe / No, I don't think so / No, definitely not

Appendix E. Coding tree

1. Motivation of MHNP and GP to participate in the trial

a. Decision regarding participation GP or MHNP
b. Motivation for participation GP
c. Motivation for participation GP and MHNP
d. Motivation for participation MHNP

2. The intervention and intervention manual

a. Training, supervision and competencies of the MHNP

i. Contribution of training
ii. Previous experience with PST
iii. Feeling competent while carrying out the intervention
iv. Dealing with not feeling competent
v. Reason for not feeling competent while carrying out the intervention
vi. Reason for not feeling competent
vii. Supervision session
viii. Training MHNP

b. Use of intervention manual

i. Experience with using the protocol
ii. Experience with using the protocol: adherence to structure
iii. Experience with using the protocol: complicated
iv. Experience with reading the protocol out loud
v. Flexibility of the protocol
vi. View of manual in general
vii. View of manual: too much explanation
viii. View of manual: too much paper
ix. View of manual: a lot of text/MHNP talks too much
x. Motivation for use of the manual
xi. Reason no motivation to work with the manual
xii. Suitability content protocol and patient
xiii. Number of sessions
xiv. Frequency of sessions
xv. Contact MHNP & patient
xvi. Suitability protocol for patients
xvii. Suitability cognitive behavioural intervention for MUPS
xviii. Length of sessions

xix. View content protocol
xx. View content protocol: clarity
xxi. View content protocol: repetition
xxii. View content protocol: light-hearted
xxiii. View content protocol: personalization
xxiv. View content protocol: texts from manual
xxv. View on cognitive behavioural intervention
xxvi. Session 1
xxvii. Session 2
xxviii. Session 3
xxix. Session 4
xxx. Session 5
xxxi. Session 6
xxxii. Setting SMART goals
xxxiii. Pace of the sessions
xxxiv. Proportion content to length sessions

c. *Logistics*
i. Number of patients seen
ii. Support from the research team
iii. Communication about 2nd EPD search
iv. Communication between GP and MHNP about participating
v. Opinion about audiotaping the sessions
vi. MHNPs' opinion regarding participation in the trial
vii. General opinion on the CIPRUS study
viii. Audiotaping sessions
ix. Time between training and first session with patient
x. Time invested by MHNPs
xi. Second EPD search
xii. Inviting patients to sessions with MHNP
xiii. Process of inviting patients to sessions with MHNP
xiv. Informing patients about the intervention

3. Routine care for MUPS
a. Referral of MUPS patients by GP to MHNP
b. Routine care for MUPS within the surgery
c. Number of MUPS patients during consultation hour
d. Regular length of sessions with MHNP in the surgery

e. Regular process of inviting patients to a session with MHNP
f. Reason for not referring MUPS patients to MHNP

4. Patient suitability, demand for care and patient drop-out
a. Number of suitable patients
b. Number of not suitable patients
c. Number of patients with a demand for care
d. Number of patients who dropped out
e. Presence of demand for care within the patient
f. Opinion regarding presence of demand for care
g. Exclusion criteria
h. Exclusion criterion: low IQ
i. Exclusion criterion: insufficient Dutch language skills
j. Exclusion criterion: sufficient problem-solving skills
k. Suitability of patients for participation
l. Feeling pressured by participation in the study
m. Inclusion criteria
n. Inclusion criterion: social contact?
o. Patient motivation
p. Reason for not being suitable
q. Reason for not being suitable: comorbidity
r. Reason for not being suitable: patient feels she already deals with symptoms well
s. Reason for not being suitable: externalization
t. Reason for not being suitable: focus on medical cause
u. Reason for not being suitable: lack of motivation
v. Reason for not being suitable: not MUPS
w. Reason for not being suitable: symptoms present for too long to reach beneficial result
x. Reason for not being suitable: high level of intelligence
y. Reason for not being suitable: low level of intelligence
z. Reason for not being suitable: age
aa. Reason for not being suitable: multiproblem
bb. Reason for not being suitable: negative attitude
cc. Reason for not being suitable: insufficient Dutch language skills
dd. Reason for not being suitable: symptoms of personality disorder
ee. Reason for not being suitable: patient too chaotic
ff. Reason for not being suitable: too confronting

gg. Reason for not being suitable: trauma
hh. Reason for not being suitable: sufficient problem-solving skills
ii. Reason for not being suitable: female
jj. Reasons for patient drop-out
kk. Reason drop-out: decrease in symptoms
ll. Reason drop-out: other medical explanation
mm. Reason drop-out: different expectations
nn. Reason drop-out: treatment elsewhere
oo. Reason drop-out: focus on physical symptoms
pp. Reason drop-out: lack of demand for care
qq. Reason drop-out: lack of motivation
rr. Reason drop-out: not MUPS
ss. Reason drop-out: MHNP's attitude
tt. Reason drop-out: not open to new solutions
uu. Reason drop-out: symptoms of personality disorder
vv. Reason drop-out: protocol too complex
ww. Reason drop-out: poor patient selection
xx. Reason drop-out: too demanding
yy. Reason drop-out: too confronting
zz. Reason drop-out: too busy
aaa. Reason drop-out: too much paperwork
bbb. Reason drop-out: long-term stay in a foreign country
ccc. Reason drop-out: exacerbation of symptoms
ddd. Reason drop-out: patient moved
eee. Reason drop-out: sufficient problem-solving skills
fff. Reason drop-out: did not want to switch to another MHNP
ggg. Patient selection
hhh. Doubts about suitability of participant
iii. Patient drop-out

5. Perceived effectiveness of the intervention

a. Treatment effect lasting
b. Number of patients who dealt better with MUPS after the intervention
c. Number of patients whose demand for care was met by the intervention
d. Demand for care met by the intervention
e. Effectiveness: dealing with MUPS better
f. Patient experience with regard to the intervention
g. MHNP's experience with regard to meeting demand for care

h. Reason for effectiveness of the intervention
i. Reason for effectiveness: patient's cognitive ability
j. Reason for effectiveness: factors besides the intervention
k. Reason for effectiveness: patient's insight
l. Reason for effectiveness: cognitive behavioural technique
m. Reason for intervention not being effective
n. Reason no effectiveness: different expectations
o. Reason no effectiveness: low IQ
p. Reason no effectiveness: multiproblem

6. Usefulness of the manual and gains for the MHNP
a. Usefulness of manual for MHNP
b. Gains for MHNP
c. Gain: realistic goals
d. Gain: more problem-solving
e. Gain: patient doing the work
f. Gain: practical
g. Gain: PST skills
h. Gain: psychoeducation
i. Gain: structure

7. Facilitators and barriers for implementation of the intervention
a. Barriers for carrying out the CIPRUS study
b. Barrier: not being able to deviate from protocol
c. Barrier: too few patients to get into the rhythm
d. Barrier: MHNP's time
e. Barrier: stairs in the surgery
f. Factors needed for implementation
g. Factors needed for implementation: adjustment protocol
h. Factors needed for implementation: availability as an e-health program
i. Factors needed for implementation: availability of manual
j. Factors needed for implementation: feedback for MHNPs
k. Factors needed for implementation: information about MUPS for MHNPs
l. Factors needed for implementation: informing the GP about training and patient selection
m. Factors needed for implementation: introduction by GP
n. Factors needed for implementation: intervision MHNPs
o. Factors needed for implementation: online use

p. Factors needed for implementation: training MHNPs
q. Factors needed for implementation: time
r. Factors needed for implementation: scientific evidence
s. Facilitators for carrying out the CIPRUS study in the surgery
t. Facilitator: focus on MUPS in the surgery
u. Facilitator: presence of the MHNP
v. Facilitator: announcement for patients that the surgery participates in scientific research in general
w. Facilitator: location
x. Facilitator: GP as a traditional concept
y. Facilitator: planning the sessions
z. Facilitator: patients' intelligence
aa. Facilitator: MHNPs' interest in the subject
bb. Facilitator: small-scale of the surgery
cc. Facilitator: GPs' motivation
dd. Facilitator: personally knowing the researcher
ee. Facilitator: well-organized schedule
ff. Facilitator: space
gg. Facilitator: collaboration within the surgery
hh. Facilitator: MHNP's time
ii. Facilitator: trust between GP and patients
jj. Facilitator: findability of the MHNP
kk. Solution to a barrier

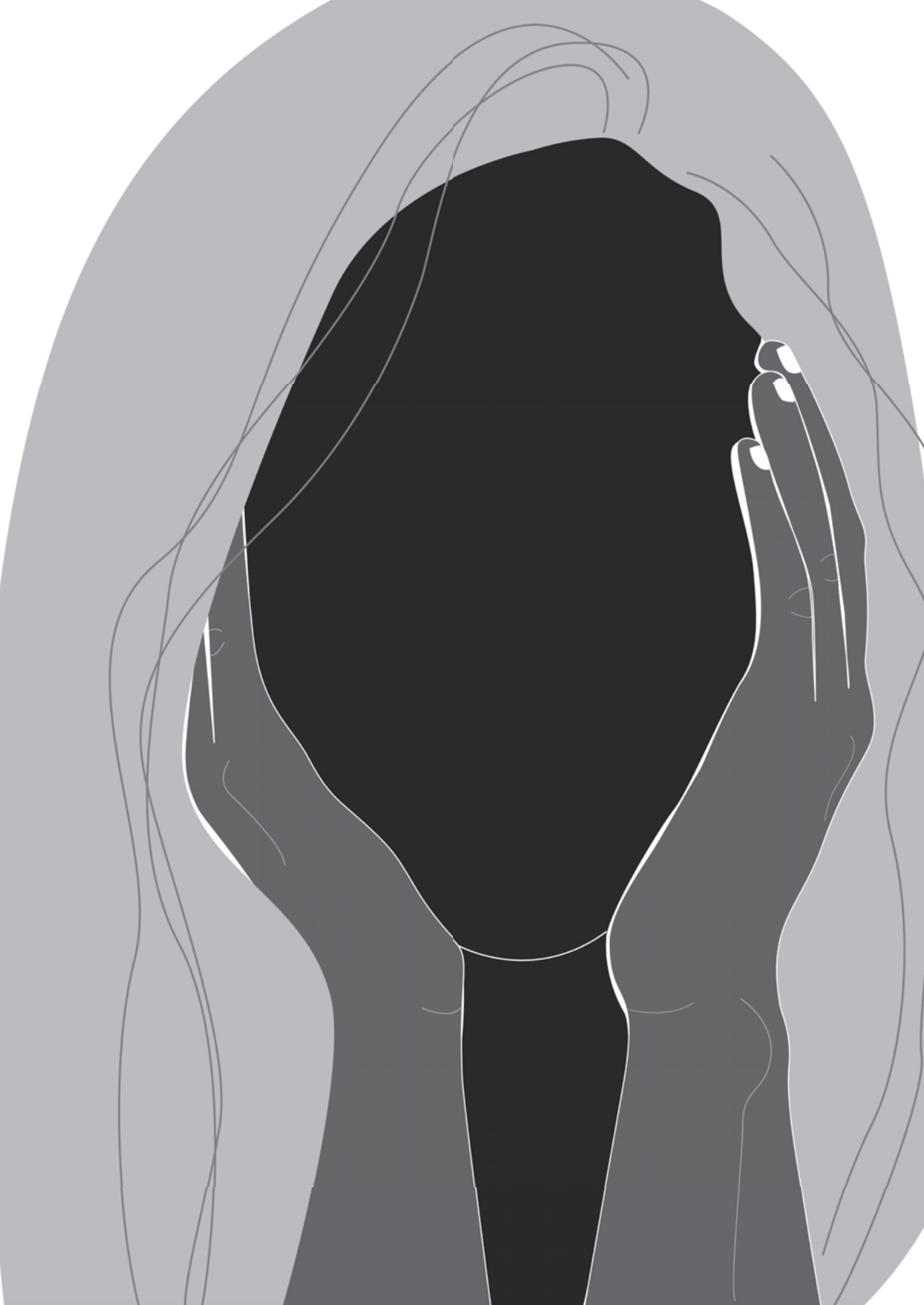
8. Protocol adherence by MHNPs
a. Own adjustments to the manual
b. Experience with adhering to the manual
c. Adhering to reading the texts out loud such as in the manual
d. Reason not adhering to the manual

9. Future use: recommendations
a. Availability of MUPS treatment in secondary care
b. Use of manual outside of the trial/in the future
c. Use of cognitive behavioural intervention outside of the trial
d. Ideal treatment for MUPS
e. Ideal treatment for MUPS: Acceptance and Commitment Therapy (ACT)
f. Ideal treatment for MUPS: CBT
g. Ideal treatment for MUPS: combination of treatment methods

h. Ideal treatment for MUPS: referral
i. Ideal treatment for MUPS: e-health
j. Ideal treatment for MUPS: education
k. Ideal treatment for MUPS: physical activation
l. Ideal treatment for MUPS: use of apps
m. Ideal treatment for MUPS: personalized treatment
n. Ideal treatment for MUPS: moderate patients with MHNP, referral of severe patients
o. Ideal treatment for MUPS: mindfulness
p. Ideal treatment for MUPS: multi-cultural
q. Ideal treatment for MUPS: multidisciplinary
r. Ideal treatment for MUPS: PST
s. Ideal treatment for MUPS: psychomotor physiotherapy
t. (Ideal) own interpretation of manual for future use
u. Own interpretation of manual for future use: e-health
v. Own interpretation of manual for future use: psycho-education
w. Incorporating the manual into usual GP care
x. Reason for (not) using the manual in the future

10. Miscellaneous

a. MHNP's interest in MUPS
b. Other diseases/symptoms
c. Treatment of participants after the CIPRUS study
d. MUPS awareness among MHNP/GP
e. Use of manual with patients with symptoms other than MUPS
f. Use of cognitive behavioural intervention with patients with symptoms other than MUPS
g. MHNP's interest in other MHNPs' experiences
h. Miscellaneous comments by the MHNP
i. Patient's expectation with regard to the intervention
j. Patient's expectation with regard to the role of the MHNP



Chapter 8

General discussion

Aims and design of the study

We wanted to investigate several issues relevant for patients with MUPS in primary care. First, we wanted to establish which currently available self-report questionnaire has the best measurement properties to measure MUPS. As questionnaires directly measuring MUPS are lacking, questionnaires measuring somatization are often used for this purpose. Therefore, we wanted to know which questionnaires measuring somatization had the best clinimetric properties. Secondly, we investigated the current management of MUPS patients in primary care in the Netherlands. In 2013 a MUPS guideline for general practitioners (GPs) was published in the Netherlands. The guideline contains recommendations for diagnostic and treatment strategies. We wanted to know to which extent GPs manage MUPS according to the guideline. We also designed a cognitive behavioural intervention for patients with persistent MUPS, that can be delivered by mental health nurse practitioners (MHNPs) in general practice. We tested its effectiveness and cost-effectiveness over a period of one year. In order to operationalize persistent MUPS, we included patients who met the DSM-IV classification criteria for undifferentiated somatoform disorder (USD). Finally, we wanted to know patients' and MHNPs' evaluation of our intervention, and whether MHNPs thought it is acceptable and feasible to implement the intervention on a larger scale.

Outline

In this final chapter we first provide a summary of the main findings per research question presented in chapter 1, and interpret the results within the context of previous research. We then explore the methodological considerations of our research. Finally, based on our findings, we discuss clinical implications and provide recommendations for future research.

Main findings

Research questions

1) *What is the best self-report measurement instrument to measure somatization in primary care?*

To answer this question, we conducted a systematic review of studies investigating the clinimetric characteristics of self-report questionnaires that measure somatization in primary care. Twenty-four publications describing nine different questionnaires were eventually included. We investigated the methodological quality of these nine questionnaires using the COnsensus-based Standards for the selection of health Measurement Instruments (COSMIN) checklist. For each questionnaire the overall level of evidence on each separate measurement property was synthesized using data on

measurement properties from all included studies. We adjusted the levels of evidence for the methodological quality of each study according to predefined criteria. The somatization subscale of the 4-Dimensional Symptom Questionnaire (4DSQ) and the Patient Health Questionnaire 15-item somatic symptom severity scale (PHQ-15) were studied in the largest number of articles and covered the broadest range of measurement properties. These two questionnaires also had the best results regarding the quality of measurement properties. Both questionnaires are similar in length and had the best internal consistency, test-retest reliability, structural and construct validity. The 4DSQ is validated in different languages, so could possibly be the instrument of choice for use in a population speaking one of those languages. The PHQ-15, on the other hand, scored better on criterion validity, meaning that its cut-off score correlates best with the gold standard, a DSM-IV somatoform disorder. Additionally, the PHQ-15 includes two items on menstruation and sexual intercourse, so a choice of questionnaire can be made depending on whether information on these health aspects is of interest. Several other questionnaires, the Bodily Distress Syndrome (BDS) checklist, the Physical Symptom Checklist (PSC-51) and Symptom-Check-List (SCL-90-R) somatization subscale also showed promising results on some measurement properties. However, these questionnaires were investigated in only one or two studies, and the various measurement properties were not studied exhaustively.

2) What does the current management of MUPS patients in Dutch primary care entail and to what extent is it in line with the national guideline for persistent MUPS?

We conducted an observational study of adult primary care patients with MUPS in 30 general practices. We extracted data routinely recorded by GPs from electronic medical records of 77 patients with MUPS, over a 5-year period prior. In total, data from 1035 consultations were collected. We discerned diagnostic and therapeutic management strategies. The most frequently recorded diagnostic strategies were physical examination (24.5% of all consultations) and additional investigation within the surgery (14.6%). The most frequently recorded therapeutic strategies were prescribing medication (24.6% of all consultations) and vitamins and supplements (11.7%). GPs also frequently gave education and explanation about the symptoms (11.2%), gave different kinds of (lifestyle) advice (10.8%) and discussed medication use (9.4%) and progress (16.2%). GPs rarely recorded fully exploring other dimensions of the symptoms than the somatic dimension (i.e. cognitive, emotional, behavioural and social dimensions) (3.5%), referring to a mental health nurse practitioner within their own surgery (1.4%) or a psychologist outside the surgery (0.5%). A small number of consultations (2.8%)

covered consultations with a MHNP. We concluded that the management of MUPS by these participating GPs was only partly in line with the Dutch guideline.

3) What is the effectiveness of a cognitive behavioural intervention for patients with undifferentiated somatoform disorder carried out by mental health nurse practitioners in Dutch primary care?

We designed and conducted a cluster-randomized controlled trial evaluating the effectiveness of a cognitive behavioural intervention among adult primary care patients with undifferentiated somatoform disorder (USD), compared to usual care. Our primary outcome of interest was physical functioning as measured by the physical component summary score (PCS) of the RAND-36. We also looked at the remaining subscales of the RAND-36 as well as anxiety, depression and somatic symptom severity, as secondary outcomes measures. The intervention group consisted of 111 participants and the usual care group consisted of 87. Over a 12-month period, with five time points including baseline, the intervention group showed an improvement in physical functioning (mean difference 2.24 [95% CI 0.51; 3.97]; p=0.011), a decrease in limitations due to physical problems (mean difference 10.82 [95% CI 2.14; 19.49]; p=0.015) and in pain (mean difference 5.08 [95% CI 0.58; 9.57]; p=0.027). There were no statistically significant differences for anxiety, depression and somatic symptom severity. Effects were larger and clinically relevant for patients with a shorter symptom duration and patients with fewer comorbid physical diseases. Patients with longer symptom duration, on the other hand, showed a deterioration in general health perceptions after our intervention (mean difference -5.98 [95% CI -11.09; -0.86]; p=0.022).

4) Is the new cognitive behavioural intervention cost-effective compared to current usual care?

We performed an economic evaluation from a societal perspective alongside the clinical trial over a 12-month period. We evaluated differences in costs, quality adjusted life years (QALYs), physical functioning (RAND-36 Physical Component Summary score (PCS)), anxiety and depression (Hospital Anxiety and Depression Scale (HADS)) and somatic symptom severity (Patient Health Questionnaire 15-item somatic symptom severity subscale (PHQ-15)). Mean total costs in the intervention group were significantly lower than in the usual care group (mean difference -2300€, 95% CI -3257 to -134), over the 12-month period. The mean difference in QALYs was 0.01 (95% CI -0.01 to 0.04), the mean difference in the PCS was 2.46 (95% CI 1.44 to 3.47), -0.26 (95% CI -0.81 to 0.28) in the PHQ-15, and -0.07 (-0.81 to 0.67) in the HADS. Although the health related quality

of life outcomes showed no significant changes, the costs decreased significantly and substantially. Therefore, the intervention was more cost-effective than usual care. At a willingness to pay of 0 € per additional unit of effect, the probability of the intervention being cost-effective was 0.93 for QALYs and 0.92 for PCS, PHQ-15 and HADS scores.

5) How did patients and MHNPs evaluate the cognitive behavioural intervention for undifferentiated somatoform disorder?

For the process evaluation of our trial, patients and MHNPs completed written questionnaires. Additionally, all participating MHNPs were interviewed and a large part of our findings originate from qualitative data from these interviews. Both patients and MHNPs were generally satisfied with most aspects of the cognitive behavioural intervention as provided in our trial. As MHNPs were requested to adhere to the text in the protocol as much as possible during their sessions with the patients, they reported that they generally did so, but did not feel comfortable following the text very strictly. They would have liked more flexibility. MHNPs generally perceived the intervention to be an effective one and reported that it seemed to improve patients' problem-solving skills. They also reported that the intervention provided specific tools for treating patients with USD. At the same time they expressly reported that the intervention was not suitable for at least a third of the patients due to patients' psychological/psychiatric comorbidity, to psychosocial problems, to patients not being open for change, to lower education level and older age. MHNPs also recommended that the GP should explain better to patients why they are referred to the MHNP with physical symptoms. MHNPs felt that this would have made the intervention go smoother. Half of the patients reported that the intervention helped them at least somewhat in dealing with their symptoms and a third would undergo it again in the future, if needed. However, nearly half of the patients did not complete the 6 sessions for various reasons.

Interpretation of our findings

1) What is the best self-report measurement instrument to measure somatization in primary care?

We were interested in investigating which self-report measurement instruments are most suitable for measuring MUPS, but MUPS is not commonly directly measured by questionnaires. In research, frequently used questionnaires for this purpose are the ones measuring the number of self-reported somatic symptoms as an operationalization of somatization. We therefore investigated the clinimetric properties of questionnaires measuring somatization as a proxy for measuring MUPS.

The results of the systematic review may be somewhat unsurprising, as the questionnaires that scored best, the PHQ-15 somatic symptom severity scale and the somatization subscale of the 4DSQ, were also the ones investigated the most. It may be slightly unfair to draw these conclusions, as other promising questionnaires could possibly have shown positive results on a broader range of measurement properties if they would have been studied more extensively. However, as the best scoring questionnaires seem to measure somatization well, new questionnaires may not be needed. The choice for either one of the questionnaires can be based on preferences such as the studied population and the time period for which information is gathered (4DSQ: previous week, PHQ-15: previous 4 weeks). Whereas the former may be somewhat more accurate due to a smaller chance of recall bias, the latter may provide more complete information and may not be prone to bias from transient symptoms. Moreover, the 4DSQ somatization subscale is part of the 4DSQ questionnaire, which also includes subscales on depression, anxiety and distress (all of which are associated with somatization). All of these subscales are presented conveniently on two sheets of paper and can be administered and scored quickly in the GP surgery (e.g. the waiting room). So if the practitioner or researcher is interested in a more complete overview of the patient's complaints, and wants to know whether depression, anxiety and distress symptoms are present, the 4DSQ could be the questionnaire of choice. In our own trial we chose for the PHQ-15, partly because of it being a widely used measurement instrument in international scientific research, and partly because of better comparability with other similar trials.

A point of consideration is that none of the included questionnaires measure various psychological and behavioural aspects associated with somatization such as unhelpful or maladaptive cognitions, emotions and behaviours. Several questionnaires, particularly focusing on measuring the criteria for DSM-5 somatic symptom disorder, such as the Screening for Somatoform Symptoms-2 (SOMS-2) questionnaire (1), do measure such psychological and behavioural aspects. However, these questionnaires are more broadly used in secondary care, as the DSM-5 is a classification system for psychiatric disorders, mainly used in secondary care. It would be insightful to validate or develop a similar questionnaire for primary care and find out whether detection of psychological aspects of somatization would be feasible and useful in an earlier stage of the stepped-care process.

Another point of consideration regarding the term 'somatization' is that it lacks a gold standard. This causes somewhat of a problem when evaluating criterion validity, which is the degree to which the measurement instrument is an adequate reflection of a gold standard. In our review we chose the DSM-IV somatoform disorder as the gold standard,

as at the time of our review no studies evaluating clinimetric properties of questionnaires with the DSM-5 criteria, that fit our inclusion criteria, had been published yet. However, since the DSM-IV has been replaced by the DSM-5 in 2013, new and possibly different results for criterion validity may become available. It would be particularly interesting to validate all of the included questionnaires against the DSM-5 category somatic symptom disorder because these criteria list the cognitive, emotional and behavioural aspects of persistent and distressing somatic symptoms, whereas those of the DSM-IV somatoform disorder do not. A new validation against the DSM-5 criteria of somatic symptom disorder could, therefore, provide us with the new insights and most up to date information on the criterion validity of measurement instruments.

2) What does the current management of MUPS patients in Dutch primary care entail and to what extent is it in line with the national guideline for persistent MUPS?

GPs in our sample tended to turn to "physical" strategies during both diagnostic and treatment phase. During the diagnostic phase, GPs mainly carried out physical investigations and additional investigations such as laboratory tests, electrocardiograms and x-rays, presumably to rule out physical disease. As their GPs had confirmed MUPS for all patients in our sample, we expected GPs to turn to physical and additional investigations less frequently, unless patients presented new symptoms. During the treatment phase, prescribing medication was the most frequently applied treatment strategy. Although the Dutch MUPS guideline for GPs does not completely discourage prescribing medication, this treatment strategy is not recommended as the main one, and is supposed to be applied temporarily, to alleviate strong, more acute symptoms. It seems that the GPs in our sample prescribed medication more regularly. This could reflect various issues. A possible reason could be that GPs were not yet sufficiently informed about the recommendations in the Dutch MUPS guideline at the point in time the study took place. The Dutch MUPS guideline appeared in 2013 and data for our observational study were collected in 2016 and 2017 with a time window starting 5 years earlier. Therefore, part of the collected data took place before the guideline was published. We investigated whether there was a trend in time, and found that over time, GPs tended to adjust medication more frequently, discuss progress more often, schedule more follow-up appointments and encourage patients more to contact the practice if necessary. Another reason could be that patients expect or request GPs to prescribe medication and that GPs comply to do so, or at least that GPs perceive this to be what patients want. GPs report that they sometimes feel pressured by the patients to provide a biomedical intervention (2). However, evidence for patients' insistence on medical interventions is lacking. On the contrary, previous studies of GP consultations with MUPS

patients found that patients expect somatic investigations, interventions and referrals less frequently than GPs propose or execute these (3, 4). This suggests that "physical" interventions are generally instigated by the GPs, rather than the patients. This could be the case in our study as well.

We were somewhat surprised to find that GPs rarely referred patients to the MHNP and barely made use of the 4DSQ questionnaire, measuring the level of somatization and frequently used for measuring MUPS (alongside anxiety, depression and distress). Since the MHNP has a prominent role in Dutch general practice and all of the participating surgeries had a MHNP, it is an interesting finding that MUPS patients are not commonly referred to the MHNP. However, the MHNP only received a prominent role in general practice starting from 2014, so it is a relatively recent development. GPs may have needed time to make optimal use of this collaboration. Possible other reasons for lack of referral to the MHNP may be that GPs felt that it is their own task to treat patients with physical symptoms, or that the MHNPs did not feel comfortable treating these patients. As the results of the process evaluation of our randomized trial showed (chapter 7), MHNPs were not always happy to have MUPS patients referred to them, because they did not perceive to have 'tools' for managing or treating these patients and therefore would have preferred to refer these patients elsewhere. However, we saw that referrals to other primary or secondary care psychologists did not happen frequently either. The burden of managing MUPS patients therefore remained largely with the GP.

When we compared the results of the intervention provided by the MHNPs in our trial to usual care, we saw that patients in the intervention group improved in physical functioning (primary care outcome) compared to the usual care group. Interestingly, part of this effect could be contributed to a deterioration of physical functioning in the usual care group. This deterioration remained stable 12 months after baseline. Evidently, the usual care provided by the GPs was not effective. As we saw, GPs mainly opted for medical strategies, whereas more 'conversation-like' strategies that focused more on listening to the patient and involving patients in their own diagnostic and therapeutic process, and decision making, were lacking. Assuming that the data gathered from our observational study are representative for the usual care that patients actually received, we might conclude that engaging in these 'conversation-like' strategies would be an improvement in GPs' management of MUPS patients. Our findings are supported by results from a previous systematic review which found that improved doctor-patient communication, perceiving MUPS patients' expectations correctly and explaining the nature of MUPS and the meaning of normal test results, reduced patients' symptoms and anxiety, and improved patient satisfaction (5).

A side note must be made that due to the observational nature of this study, we do not know what actually went on during the consultations and we were dependent on the GPs' recordings. As mentioned in chapter 4, it would have been interesting if we had asked the GPs what took place during the consultation, whether they possibly applied the more 'conversation-like' techniques, but failed to record them. Video recording consultations might also provide more insight into what is going on.

3) What is the effectiveness of a cognitive behavioural intervention for patients with an undifferentiated somatoform disorder carried out by mental health nurse practitioners in Dutch primary care?

We were happy to find a positive effect on our primary outcome physical functioning, and on some of our secondary outcomes, i.e. limitations due to physical symptoms and bodily pain. Although we found a statistically significant difference on the primary outcome (PCS of the RAND-36) in the intervention group compared to the control group (2.24), the difference was just not large enough to count as a clinically relevant difference (3-5). This effect therefore has to be interpreted with caution. Nevertheless, it is a promising finding, particularly considering our sample size.

In the group of patients with a shorter duration of symptoms, general health perceptions were positively influenced as well. A striking finding is that all outcomes that showed an improvement were physical ones. None of the mental outcomes were affected significantly by our intervention. At first glance this might be somewhat surprising, as the intervention offered in our trial was a psychological one. However, our intervention focused on the consequences of the physical symptoms, which were then "solved" using adapted Problem-Solving Treatment (PST). Patients were free to choose any type of consequences of their MUPS. Presumably, patients chose physical consequences or limitations that arose due to their MUPS first, as patients frequently report being limited in their daily physical functioning (e.g. doing groceries, cleaning, picking children up from school). This was also one of the reasons we chose this outcome as the primary one. Unfortunately, we do not have documentation of what consequences patients identified and worked on during the sessions. We can therefore only assume that as physical consequences were more urgent for patients, they worked through these, and mental consequences received less attention.

Although generally speaking our intervention was CBT-based, it did not focus on identifying, challenging and altering cognitions, as is done in classic CBT. The process of

altering cognitions in order to alter (negative) emotions was not part of our intervention. It is therefore not surprising that emotional outcomes were unaffected.

Our intervention also did not actively focus on symptoms of depression and anxiety. However, these symptom levels were not very high in our population to start with (chapter 5, table 1). Scores between 0-7 on each Hospital Anxiety and Depression subscale (HADS anxiety and HADS depression) are considered to be normal, scores of 8-11 indicate mild symptoms, scores between 11 and 14 moderate and scores between 14 and 21 indicate severe symptoms. Patients in our sample had a mean HADS anxiety score of 7.8 and a mean HADS depression score of 7.2 at baseline. These scores lay between the normal and mild symptom categories and therefore there may have been too little room for improvement, although we did expect that by improving physical functioning, depressive and anxiety symptoms would indirectly perhaps decrease as well.

A previous randomized controlled trial by Zonneveld and colleagues, investigating a more intensive group psychological intervention for MUPS patients, provided by psychologists within general practice did find effects on mental health outcomes (6). The design of our intervention was partly based on this trial. It showed significant improvements on role limitations due to emotional problems, vitality and social functioning alongside the physical functioning outcomes. Possible reasons for these positive findings could be that their intervention protocol covered the entire range of consequences (physical, cognitive, emotional, social and behavioural) separately. Their protocol consisted of 13 sessions, with a whole session dedicated to each single consequence. Therefore, patients may have become more aware of and paid more attention to emotional, social and behavioural consequences of their symptoms and worked at solving these. In our study, patients were free to work on any consequences they chose, and areas they did not choose to work on were just ignored. It is possible that our patients therefore did not work on the emotional and social consequences, and hence did not improve on any of these outcomes. Anxiety and depression did not significantly improve in the previous trial by Zonneveld et al. either.

4) Is the new cognitive behavioural intervention cost-effective compared to current usual care?

We were pleased that the intervention proved to be cost-effective over usual care as well. Even though the effects on the clinical outcomes were not very large and for some outcomes not statistically significant, the costs were so much lower in the intervention

group compared to usual care (-2300€ from the societal perspective and -754€ from the healthcare perspective) that the intervention was cost-effective anyway.

As we saw in chapter 4, GPs currently do not tend to make use of the MHNPs' services as part of their usual care. Our findings from the cost-effectiveness study show that economically it would be beneficial to let MHNPs treat MUPS patients within the general practice setting. However, MHNPs would, of course, first have to be trained to do so. This would mean that resources would have to be allocated towards the MHNPs' time and training. We hope our project helps convince GPs and practice managers on spending resources on an intervention that initially may cost some money and requires reorganization of resources. However, our results show that just within a period of 12 months, the intervention reduces costs, and particularly the costs in primary care, substantially (by nearly a third, see table 2 in chapter 6). Therefore, it seems worthwhile to implement an intervention like ours, particularly after having improved and focused it as suggested by MHNPs as described below.

Our cost-effectiveness study also has a drawback; the design of our cost-effectiveness analysis was a so-called "piggyback" design, in which the economic and additional health outcomes were added to those of the randomized controlled clinical trial. Although this is an efficient design that obtains timely cost-effectiveness data, it also usually means that the power calculation is not based on costs, but on clinical outcomes. Despite this limitation, cost-effectiveness was still shown in our trial and this is in line with previous research investigating cost-effectiveness for MUPS, which show that CBT interventions are generally cost-effective compared to various other pharmacological and non-pharmacological interventions and waiting-list control groups (7). These studies found that group interventions were generally more cost-effective than individual interventions. Although our intervention already proved to be cost-effective, it could be interesting to investigate the effect of a group intervention based on our design.

5) How did patients and MHNPs evaluate the cognitive behavioural intervention for undifferentiated somatoform disorder?

Reflecting on the interviews with MHNPs and questionnaires with patients, several points stood out. Generally speaking MHNPs and patients thought our intervention was helpful in dealing with USD and provided MHNPs with useful tools for treatment and provided patients with practical problem-solving skills. However, for an even more successful treatment MHNPs suggested several essential changes. First of all, GPs could have played a larger role in the patient selection and information process and MHNPs could

have spoken to patients or have had an intake session before starting the intervention. This involvement of the GPs and MHNPs before starting the intervention could have contributed to a better understanding and motivation of patients and a better relationship between MHNPs and patients. Furthermore, the selection of patients could have been stricter according to the MHNPs, regarding patients with (psychiatric) comorbidity, severe psychosocial problems, too high or too low IQ, older age and sufficient problem-solving skills, as MHNPs viewed these patients as being unsuitable for our intervention. MHNPs for example reported that the intervention was too simplistic for some people with a higher IQ. The suggestion to exclude patients with these type of comorbid factors is a challenging one, as MUPS patients in general practice typically do have psychiatric comorbidity, psychosocial problems and so on. It seems as if our patient sample did reflect the actual MUPS patients in primary care. However, it is understandable that it was hard for MHNPs to deal with these patients. Possibly, to tackle this problem in the future, MHNPS could collaborate with other primary care workers when treating patients with too many comorbid psychiatric or psychosocial problems. If this is insufficient, these patients could be referred to secondary care settings. Regarding patients with low IQ or older age, our intervention could possibly be adjusted and simplified in order to cater to this group of patients, while retaining the consequences model and PST working elements. Another common point reported in the interviews was the formal, rigid structure of the treatment protocol. As we developed the treatment protocol as part of scientific research, it was imperative that the manual was structured and adhered to in a similar way as much as possible by all participating MHNPs, to further uniformity. However, the booklet in which the manual was provided could clearly have been designed and organized better. MHNPs complained about the texts being too long and having to flip over the pages to find the right piece of text they were supposed to read out loud next. Also following the text so strictly formed a problem and was an annoyance for MHNPs. Having a more flexible protocol and leaving more space for the patient and MHNP to make genuine contact could have facilitated the relationship between MHNPs and patients and MHNPs' motivation to continue using the treatment manual. Finally, although we were mainly investigating a combination of the consequences model and PST as the intervention technique in our study, it would be desirable to have the flexibility to switch between treatment techniques (such as other forms of CBT or Acceptance and Commitment Therapy (ACT)) depending on the needs and the type of patient. In conclusion, in order to implement our intervention in the future, the above points would first have to be considered or adjusted in the manual.

Embedding the results of our randomized trial in what we already know

In a Cochrane review on non-pharmacological interventions for MUPS, the effect size of CBT on functional disability and quality of life within a year after treatment was 0.22 [95% CI -0.08 to 0.53], compared to usual care or waiting list (8). We found a corresponding total effect size of 0.23 on the physical component summary score and of 0.33 on limitations due to physical health problems, within a year after treatment. When investigating this effect in groups of patients with a shorter symptom duration and patients with fewer comorbid physical diseases, we found that the effect sizes increased and even exceeded the effect size found in the Cochrane review. The effect sizes found a year after our treatment were 0.39 on the physical component summary score for patients with a symptom duration below 5.7 years and 0.36 for patients with 0-2 comorbid physical diseases.

Comparing our intervention to similar trials, for instance, once again, the 21 studies included in the Cochrane review described above, we can conclude that our intervention was less intensive than most studies included in the review. Our intervention had fewer sessions than three quarters of the included studies. In one study, not part of the three quarters, patients received 5 sessions, but of 50 minutes each, totaling to a larger number of minutes spent with the patient than in our intervention (6 x 30 minutes) (9). It is therefore promising that with a shorter amount of time than in most studies, we were still able to achieve effects of small to moderate size, in some cases the same size as in trials with more intensive interventions.

In the Cochrane review the pooled effect size on somatic symptom severity within a year was -0.29 [95% CI -0.49 to -0.09]. Contrary to our expectations we did not find a statistically significant effect on somatic symptom severity measured by the PHQ-15 in our trial (effect size -0.14). However, we did find a reduction in bodily pain with an effect size of 0.23. The effect size of the included studies in the Cochrane review on participant-rated severity of anxiety symptoms was 0.0 [95% CI -0.59 to 0.59] and of depressive symptoms 0.21 [95% CI -0.07 to 0.50] (both not statistically significant), within a year after treatment. In our study we found no effects on either anxiety or depressive disorders. Our findings therefore correspond to the anxiety aspect, but not the depressive aspect.

Finally, none of the interventions reported on in the Cochrane review were provided by a mental health nurse practitioner, most were provided by a physician, psychologist or psychotherapist. Therefore, it is promising that we achieved such results with professionals with less intensive training who were not specifically trained in

management of people with MUPS (other than the training they followed for the CIPRUS study). This suggests that our intervention is easy to learn and can be implemented by professionals such as mental health nurses and does not necessarily have to be carried out by a (more costly) psychologist or physician.

Methodological considerations concerning the RCT

Number of included patients

We were faced with several challenges in recruiting and maintaining a sufficient number of participants throughout the entire process of our study. To start at the beginning, after possible eligible patients were selected from the GPs' electronic databases, GPs were asked to confirm whether these patients had MUPS and exclude patients who fulfilled the exclusion criteria. Moreover, GPs had the freedom to exclude patients based on their own judgement of patient suitability. GPs were therefore free to (but absolutely not obligated to), for instance, exclude patients who had just experienced a difficult life event, such as death of a relative, or a divorce, or were simultaneously dealing with another somatic disease. As figure 1 in chapter 5 shows, GPs excluded more than half of the patients (58%) selected from the GPs' electronic databases. It is possible that a part of these patients were actually eligible but still excluded by the GP at this point in time, because the GP judged them not suitable. These patients were therefore not invited to participate.

After GPs selected patients who fulfilled the inclusion criteria, the eligible patients were invited by mail to participate. Most patients had not talked to their GP about the ongoing trial or had not heard of the trial through other channels before receiving the letter. Although the letter came from the general practice and explained why the patient was invited to participate in the study, it is still possible that patients were surprised by being invited to sessions with a MHNP for their physical symptoms. In fact, the researchers were telephoned by a few patients, who did not understand why they were being approached. As figure 1 in chapter 5 shows, over two thirds (67%) of invited patients did not respond at all. Perhaps they did not feel they would benefit from receiving psychological treatment for physical symptoms, due to lack of information and explanation beforehand. Previous research found that patients with MUPS tend to be ambivalent in their motivation for psychological treatment due to their largely somatically focused health beliefs (10). MUPS patients frequently tend to believe there is a medical cause for their symptoms and therefore seek a medical treatment, which can in turn negatively impact their motivation for participating in psychological interventions (11). Involvement of and information provided by the GP could have been beneficial for the recruitment process.

If patients were interested in participation, they were interviewed using the SCID-I, in order to confirm that they fulfilled the diagnostic criteria for USD. Figure 1 in chapter 5 shows that over a third (34%) of the interviewed patients were excluded at this point. Eventually, after passing the steps described above, 213 patients signed an informed consent. However, not all of these patients completed the baseline measurement. So although we had 213 informed consent forms, only 198 patients provided baseline data on the primary outcome.

Finally, we had a higher drop-out rate than we expected (27% versus 20% after 12 months). The drop-out rate was similar in the intervention (27%) and usual care group (28%). Patients in the usual care group mainly dropped out because they found it too time consuming, burdensome and did not have any personal gain from completing the questionnaires. Patients in the intervention group reported various reasons, such as wanting to find a medical reason and treatment for their symptoms, and not being satisfied with the treatment manual and/or MHNP. It is always hard to keep patients in the usual care group interested in participation if there is no obvious personal gain. All participants received small incentives in the form of gift cards after completing the second and the final questionnaire. However, this was apparently not sufficient in motivating patients to continue to participate. Regarding patients in the intervention group, the reasons for drop-out suggest that some patients were not ready to accept mental help for their physical symptoms or did not understand the benefits of mental help. Once again, this confirms that greater involvement of the GP in explaining the symptoms and the reason for being referred to a MHNP for MUPS, might heighten patients' motivation for participation. Also, as mentioned earlier, the manual could have been tailored more to the patients' needs. Although we had a higher drop-out rate than we expected, it is not exceptionally different when compared to similar trials. More than half of the trials (11 out of 21) included in the Cochrane review on non-pharmacological interventions for somatoform disorders and MUPS, also had a drop-out rate of >20% (8).

Primary outcome measure

We used the physical component summary score (PCS) of the RAND-36 as our primary outcome measure. There has been some discussion on and diversity in the use and scoring of the RAND-36 and its component summary scores. First of all the RAND-36 is very similar to another questionnaire, the SF-36. The questions are basically identical, but the two questionnaires have different algorithms for scoring (12). Also the RAND-36 is freely available, whereas the SF-36 is not. Nevertheless, these questionnaires are frequently used interchangeably and the PCS and mental component summary score (MCS) are scored differently among different researchers and distributors (13).

The authors of the Dutch version of the RAND-36 discourage the use of component summary scores and recommend using the 8 separate subscales only, scoring these according to the 0-100 algorithm (14), rather than using norm-based scores (mean=50, sd=10) as is recommended by the authors of the SF-36 (15). Possible arguments for not using the component summary scores could be their proneness to ceiling effects and sensitivity to patients' symptom attribution (15, 16). The authors of the SF-36, on the other hand, do recommend the use of the component summary scores and recommend the use of norm-based scoring for the 8 subscales as well as the component summary scores (17). It is not always clear from previous studies which methods have been used for scoring, which makes interpretation and comparison of the RAND-36 component summary scores complicated and less reliable. Furthermore, when calculating the PCS and MCS, we used the American population as the norm. Obviously Dutch and American populations may differ, so data from a Dutch norm population would be more informative for our sample. Unfortunately, these data are not available yet.

Despite these shortcomings, we chose to use the RAND-36 after all because it is validated and widely used in similar trials (6, 18-22). We also investigated the separate subscales of the RAND-36 as secondary outcomes, to gather information from the separate dimensions of quality of life and tackle the possible shortcomings of the component summary scores.

Mediation

When designing the intervention protocol we aimed to incorporate the identification and modification of psychological features, such as unhelpful cognitions, attributions and behaviour regarding MUPS. We therefore hypothesized that changes in these psychological and behavioural features would mediate the final effect of the intervention on our primary (and secondary) outcome(s). The potential mediators we investigated were problem-solving skills (Social Problem-Solving Inventory (SPSI-R)) (23), health anxiety (Whitely Index) (24), cognitive and emotional representations of illness (brief version of the Illness Perception Questionnaire (IPQ)) (25), cognitive and behavioural responses to illness (Cognitive and Behavioural Responses Questionnaire (CBRQ)) (26, 27), and level of perceived control (Pearlin Master Scale) (28). After patients had identified the consequences of their symptoms in daily life, these consequences were then solved using problem-solving treatment. We therefore expected a change in their problem-solving skills (SPSI-R). Also, if patients chose cognitive or behavioural consequences to work through in the sessions, we expected change in perceptions (brief IPQ) and cognitive and behavioural responses (CBRQ) to MUPS. However, the mediation analyses showed that none of the suggested mediators, not even problem-solving skills,

significantly mediated the effect. The highest percentages of mediation of the primary outcome were 11% and 10% and were due to the catastrophizing subscale of the CBRQ questionnaire and the item on how strongly the symptoms are experienced on the brief IPQ, respectively.

As we found that the intervention was particularly effective for patients with a shorter symptom duration and for patients with a smaller number of comorbid physical diseases, we were also interested in finding out whether the effect in these subgroups was possibly mediated by any of our proposed mediators. Therefore, we investigated mediation further and repeated our analyses for these subgroups of patients in exploratory analyses. Similar results to the original mediation analyses were found for subgroups 'symptom duration shorter than the median' and '0-2 comorbid physical diseases'. We then repeated the analyses for a subgroup of patients with a symptom duration of 2 years or less. The results showed that for this group 59% of the effect on the primary outcome was mediated by catastrophizing (catastrophizing subscale of the CBRQ). Other aspects of illness cognitions such as the perceived amount of influence of the symptoms, how strongly they are experienced and how worried the patient is (all measured by various items of the brief IPQ), seemed to mediate the effect somewhat (between 17% and 18%), but these mediation effects were not statistically significant.

Although these were exploratory analyses, they may shed light on processes that might play a role in persisting MUPS. Previous research found that psychological features such as catastrophizing (mis)interpretations are related to and even predict somatoform-related symptoms and new onsets of somatoform disorder (29, 30). Various models of somatoform symptoms incorporate the notion of patients overestimating the association between physical symptoms and negative outcomes at their core (31). Particularly catastrophizing interpretations of physical symptoms and organic causal explanations seem to be crucial cognitive processes that may lead to help-seeking, distress, avoidance, disability, emotional arousal and more physiological symptoms (31). While there is strong evidence for elevated catastrophizing interpretations of physical symptoms in patients with health anxiety (previously described as hypochondriasis), less research has been done on this association in patients with MUPS (32-34). Research conducted on this association found similar results for patients with MUPS as for patients with health anxiety (29). Results of our exploratory analyses seem to confirm these findings. New trials on interventions targeting unhelpful psychological features such as catastrophizing misinterpretations would be insightful.

Pragmatic trial

The way our study was set up highly resembles clinical practice in primary care in the Netherlands. MHNPs already work in Dutch surgeries and are supposed to have some of the patients with MUPS referred to them by the GP. Furthermore, we chose to recruit the MHNPs actually working in the participating practices, rather than train a dedicated group of MHNPs to deliver sessions in the surgeries. Therefore, we did not interfere with the way primary healthcare is being delivered in these practices. Also, sessions with the MHNP usually last between 30 to 45 minutes, and we designed the sessions in our trial to fit into this time. The only things we added were two training sessions for the MHNPs that were planned before carrying out the intervention, a structured intervention protocol that MHNPs could use during sessions with patients, and the offer of supervision. Because of the highly pragmatic nature of our trial it would be easy to implement, which contributes to generalizability in Dutch primary care.

We also did not withhold patients in the usual care group from any treatment. They received all care they would usually receive from their physician. This also meant that they were free to have sessions with their MHNP if they, or the GP, deemed this necessary. This could have diminished the contrast in effect between the two groups. However, when investigating management of patients with MUPS in the usual care group, as described in chapter 4, we saw that only 7 of the 77 control group patients (9%) from whom data was collected from their electronic medical records had actually attended sessions with a MHNP. Therefore, it's not probable that this has substantially diluted the contrast. Moreover, we still found statistically significant, and in some subgroups also clinically relevant, effects on our primary and some secondary outcomes when comparing the intervention to the usual care group.

Implications for practice

One of the most important findings from our process evaluation was that MHNPs would have liked more flexibility when applying the protocol. Although the structured outline of the manual was evaluated positively, MHNPs did not always feel comfortable following the protocol as strictly as was requested. MHNPs would have liked to get to know patients and their symptoms more extensively during the first session(s). They would have then liked to use the treatment manual, but also to enhance it with other psychological techniques, tailoring it to the patients' needs. The intervention protocol could in that case function as a framework, which MHNPs could personalize depending on the patients' knowledge, education level, skills, previous treatment and personality. Also the pace of the intervention could be adjusted depending on how fast the patient picked up the method.

When designing our trial we assumed that MHNPs had 30 minutes with a patient on average. To be on the safe side we took this (rather short) time frame into account when developing the intervention, so that, hypothetically, MHNPs in practices all over the Netherlands would be able to carry out the intervention. It turned out that in the majority of the participating surgeries, MHNPs' sessions lasted between 40 and 45 minutes, and sometimes even 60 minutes. MHNPs reported that, when following our 30 minute instructions, they usually had time to complete the first two sessions in time. However, from the third session on, and in particular during the third session, they were faced with time constraints and had to rush to complete the entire content of the session. As explained in more depth in chapter 7, the content of session 3, as described in the intervention manual, was experienced as too much for 30 minutes. Therefore, the time for our intervention could be redistributed, allowing more time for sessions 3-6.

Based on the feedback we received from the MHNPs, GPs and patients, we concluded that if such an intervention is to be implemented, the GP should have a crucial role in informing and guiding the patient towards the intervention and coordinating this process. It is important for the GP to extensively explain why a psychological intervention could be beneficial for the patient and to stress that the GP supports this type of intervention. Previous studies showed that doctors' explanations of MUPS can create common ground for psychosocial and physical aspects of the symptoms, avoiding unnecessary somatic interventions (5, 35, 36). This does not only make the MHNPs' job easier, but also reassures the patient and creates an atmosphere of trust and takes away possible doubts about the intervention or the feeling of not being taken seriously and sent to see someone for 'psychological' complaints. In our trial the GP played a minimal role, the trial was set up in such a way that the GP would not have to invest much of his/her time for practical reasons. This appealed to many GPs and probably helped them to consent to participate in our study. However, several GPs insisted they inform their own patients, which resulted in better understanding and larger percentage of participants in those surgeries. Therefore, we believe that involvement and guidance towards the intervention by the GP, in the form of one or more consultations with the patient, including information, education and advice, would increase the uptake and possibly the effectiveness of the intervention.

As involvement of the GP would probably improve the uptake and the effects of the intervention, the question could arise whether the entire intervention could be carried out by the GP. Psychological interventions carried out by GPs have been previously investigated, but results are mixed on the effectiveness of these (37-40). Also GPs' time is more scarce and costly compared to that of the MHNP. Our recommendation is therefore

to involve the GP in the process of providing sufficient explanation and motivation to the patient for seeing the MHNP, but letting the MHNP deliver the intervention. Patients may have different expectations from a GP than from a MHNP and may persist in trying to find a medical cause when faced with a medical doctor. If sessions are delivered by a MHNP, such expectations may not be present, making it easier to focus on the consequences of and learning to deal with the symptoms in a more helpful way.

If the intervention manual were to be adjusted more towards the wishes of the MHNP (made more flexible and tailored somewhat more to individual patients), we would recommend implementing it more broadly. The intervention should be applied to patients with a shorter symptom duration and fewer comorbid physical diseases. Our intervention could be the first step in the stepped care for MUPS within primary care and could be incorporated in the Dutch Guideline for MUPS for general practitioners. Furthermore, in the Netherlands MHNPs can follow a post-graduate training to receive an official MHNP certificate. Currently, the curriculum of this training does not entail much about MUPS. Information about MUPS and (parts of) the intervention investigated during the CIPRUS study could be incorporated into the MHNP training.

Recommendations for future research

First of all, it is important to note that the DSM-IV category of somatoform disorders (we used USD from the DSM-IV to operationalize persistent MUPS) has been replaced by the DSM-5 diagnosis somatic symptom disorder. The most important changes are that for a diagnosis of a somatic symptom disorder, the lack of a medical explanation for the symptoms is no longer required. Furthermore, the new classification includes at least one of the following psychological features: health anxiety, disproportionate and persistent thoughts about the symptoms, and excessive time and energy devoted to the symptoms (41). A comparison of the DSM-IV somatoform disorder and the DSM-5 somatic symptom disorder categories showed that the somatic symptom disorder criteria are more restrictive and are associated with higher symptom severity and lower physical functioning (42). Regarding the questionnaires that measure somatization the best, it would therefore be useful to investigate which questionnaires take cognitions, emotions and behaviour in account and approximate the somatic symptom disorder the most.

Secondly, although already mentioned above, it cannot be stressed enough that the role of the GP in the selection of patients is very important. Participating patients in the CIPRUS study were often unaware of why they were invited for participation and what the intervention entailed. For future studies into the effectiveness of interventions in primary

care, we therefore recommend a first meeting with the GP, followed by reassurance and guidance throughout participation in the trial. A previous study among patients with chronic fatigue showed that being in a partnership with the GP and the GPs' understanding, were factors that facilitated patient satisfaction (43) .

Regarding the content of the intervention, various psychological features, other than problem-solving skills investigated in our study, seem to be associated with MUPS (30). This finding was also evident from the interviews with the MHNPs in our study, who reported that the 'one size fits all' intervention does not seem to be suitable for all MUPS patients and that it is preferable to target the specific problems within a specific patient's life. For a large part we tackled this by using the consequences model as the base for our intervention. Patients were free to choose to work on any consequences of MUPS in their life that applied to their specific situation. Therefore, the intervention was personalized to each patient's own care needs and working through all possible psychological features associated with MUPS (regardless of what individual patients experienced) was not part of our intervention. However, in future research, it would be interesting to expand an intervention such as ours by adding one or more specific elements focusing on other psychological features. Cognitions such as catastrophizing have already been described above, but other psychological features also seem to play a role in MUPS. For instance, previous research showed that MUPS patients seem to have difficulty with emotional processing and expression (44, 45). Particularly, emotion regulation strategies in order to increase symptom tolerance seemed to be lacking. It would be interesting to incorporate this aspect of emotion regulation into a trial like ours, to see whether results would improve. In a similar ongoing trial, researchers added emotion regulation training to CBT (46). Its results are currently being awaited.

Furthermore, as we were not successful in altering mental functioning measured by the mental component summary score (MCS) of the RAND-36 and its separate mental functioning subscales, it would be useful to test alternative treatment strategies. It seems that enriching our intervention with mindfulness techniques, could contribute to more positive results. A recent randomized controlled trial investigating mindfulness-based CBT in MUPS patients showed promising results, i.e. after 9 months, mental functioning was significantly improved (22).

Based on our findings that there were no beneficial effects for patients with a longer symptom duration and higher number of comorbid physical diseases, we can emphasize that it is of great importance to treat symptoms that have not lasted for a long time yet. Furthermore, it is important to continue looking for successful ways to help these

patients. Presumably patients with a longer symptom duration and higher number of comorbid physical diseases suffer more and are the ones viewed as 'difficult' by the GP and MHNP. It is therefore essential for both patients and physicians to find what interventions are effective in treating these patients in primary and secondary care. Interventions focused on acceptance, rather than problem solving skills, could be beneficial.

The role of acceptance was stressed by several MHNPs during the process evaluation of our trial and also in the process evaluation of the mindfulness-based CBT for MUPS described above (47). Increasing awareness of the present moment, learning to accept rather than resist painful symptoms and increasing self-care and self-compassion, could improve future interventions for MUPS and their effects.

Finally, in our process evaluation we now only conducted interviews with MHNPs. Patients completed written questionnaires with multiple choice questions. Before implementing our intervention, it would be useful to conduct qualitative research based on interviews with patients who have participated in the CIPRUS study. Patients could provide us with insights on how to improve the intervention.

CONCLUSION

This thesis contributes to the body of knowledge on the effectiveness of low-intensity cognitive behavioural interventions for MUPS in primary care. In particular, we provide the first evidence for the effectiveness of an intervention over usual care delivered by a MHNP, a relatively new and important healthcare provider in Dutch general practice. We found that a low intensity cognitive behavioural intervention provided by a MHNP improved functioning and decreased limitations and bodily pain over usual care. The intervention was effective in patients with symptoms that lasted a shorter period of time and with few other comorbid physical diseases. Patients with a longer duration of symptoms and a higher number of comorbid diseases did not benefit from the intervention. We suggest to implement our intervention nationally and see scope for further improvement using the suggestions received from MHNPs and participating patients. Effective healthcare for these people can be provided with few barriers within the familiar and nearby setting of the general practice. From the economic perspective, implementation of this intervention would mean that resources within surgeries would have to be allocated differently. MHNPs would have less time for other patients and the training and mastering the intervention would cost time and money, so an investment would have to be made. However, the long-term costs would substantially decrease, as

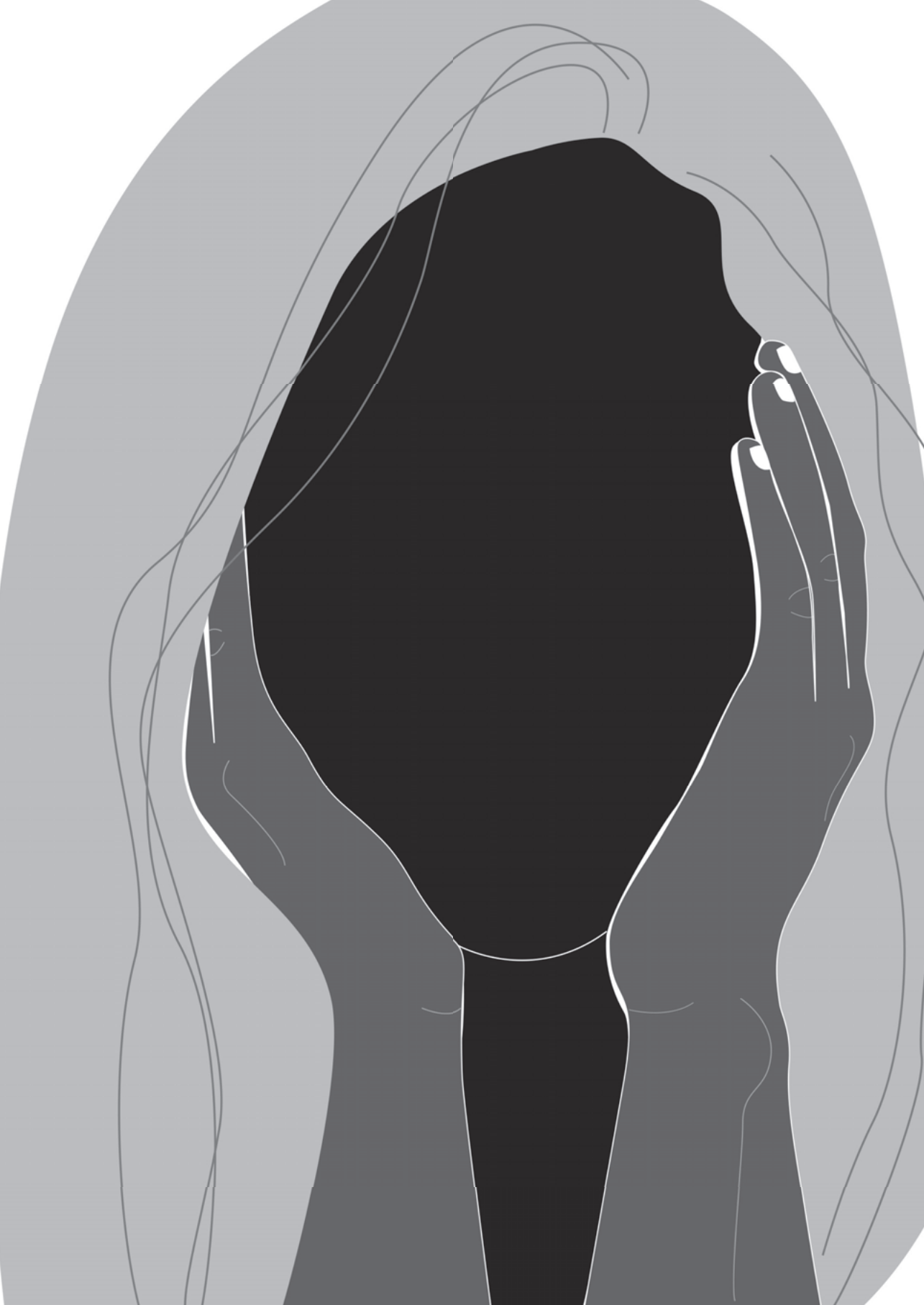
patients would presumably make less use of primary healthcare (among others). The intervention would therefore be beneficial for a broad range of stakeholders. Regarding patients with a longer symptom duration and a larger number of comorbid physical diseases, more research has to be done to discover what type of intervention would be (cost-)effective for them.

REFERENCES

1. Fabiao C, Silva MC, Barbosa A, Fleming M, Rief W. Assessing medically unexplained symptoms: Evaluation of a shortened version of the SOMS for use in primary care. BMC Psychiatry. 2010;10.
2. Reid S, Whooley D, Crayford T, Hotopf M. Medically unexplained symptoms—GPs' attitudes towards their cause and management. Family Practice. 2001;18(5):519-23.
3. Ring A, Dowrick CF, Humphris GM, Davies J, Salmon P. The somatising effect of clinical consultation: what patients and doctors say and do not say when patients present medically unexplained physical symptoms. Soc Sci Med. 2005;61(7):1505-15.
4. Ring A, Dowrick C, Humphris G, Salmon P. Do patients with unexplained physical symptoms pressurise general practitioners for somatic treatment? A qualitative study. BMJ. 2004;328(7447):1057.
5. Weiland A, Van de Kraats RE, Blankenstein AH, Van Saase JL, Van der Molen HT, Bramer WM, et al. Encounters between medical specialists and patients with medically unexplained physical symptoms; influences of communication on patient outcomes and use of health care: a literature overview. Perspect Med Educ. 2012;1(4):192-206.
6. Zonneveld LN, van Rood YR, Timman R, Kooiman CG, Van't Spijker A, Busschbach JJ. Effective group training for patients with unexplained physical symptoms: a randomized controlled trial with a non-randomized one-year follow-up. PLoS One. 2012;7(8):e42629.
7. Wortman MSH, Lokkerbol J, van der Wouden JC, Visser B, van der Horst HE, Olde Hartman TC. Cost-effectiveness of interventions for medically unexplained symptoms: A systematic review. PLoS One. 2018;13(10):e0205278.
8. van Dessel N, den Boeft M, van der Wouden JC, Kleinstäuber M, Leone SS, Terluin B, et al. Non-pharmacological interventions for somatoform disorders and medically unexplained physical symptoms (MUPS) in adults Cochrane Database Syst Rev. 2014;11:CD01114.
9. Schweickhardt A, Larisch A, Wirsching M, Fritzsche K. Short-term psychotherapeutic interventions for somatizing patients in the general hospital: a randomized controlled study. Psychother Psychosom. 2007;76(6):339-46.
10. Kleinstauber M, Witthoft M, Hiller W. Efficacy of short-term psychotherapy for multiple medically unexplained physical symptoms: a meta-analysis. Clin Psychol Rev. 2011;31(1):146-60.
11. Heider J, Kock K, Sehlbrede M, Schroder A. Readiness to change as a moderator of therapy outcome in patients with somatoform disorders. Psychother Res. 2018;28(5):722-33.
12. Farivar SS, Cunningham WE, Hays RD. Correlated physical and mental health summary scores for the SF-36 and SF-12 Health Survey, V.I. Health Qual Life Outcomes. 2007;5(54).
13. Hays RD, Morales LS. The RAND-36 measure of health-related quality of life. Ann Med. 2001;33(5):350-7.
14. van der Zee KI, Sanderman R. [Measuring general health with the RAND-36, a guideline] (in Dutch). 2nd ed. Groningen: Noordelijk Centrum voor Gezondheidsvraagstukken, NCG; 2012.
15. Schroder A, Oernboel E, Licht RW, Sharpe M, Fink P. Outcome measurement in functional somatic syndromes: SF-36 summary scores and some scales were not valid. J Clin Epidemiol. 2012;65(1):30-41.
16. Taillefer SS, Kirmayer LJ, Robbins JM, Lasry J. Psychological correlates of functional status in chronic fatigue syndrome. J Psychosom Res. 2002;53:1097–106.
17. Ware R. User's manual for SF-36v2 health survey. 2nd edition. Lincoln: Quality Metric Inc.; 2002.
18. Fjorback LO, Arendt M, Ornbol E, Walach H, Rehfeld E, Schroder A, et al. Mindfulness therapy for somatization disorder and functional somatic syndromes: randomized trial with one-year follow-up. J Psychosom Res. 2013;74(1):31-40.
19. Sattel H, Lahmann C, Gundel H, Guthrie E, Kruse J, Noll-Hussong M, et al.

Brief psychodynamic interpersonal psychotherapy for patients with multisomatoform disorder: randomised controlled trial. Br J Psychiatry. 2012;200(1):60-7.
20. Schaefert R, Kaufmann C, Wild B, Schellberg D, Boelter R, Faber R, et al. Specific collaborative group intervention for patients with medically unexplained symptoms in general practice: a cluster randomized controlled trial. Psychother Psychosom. 2013;82(2):106-19.
21. Schroder A, Rehfeld E, Ornbol E, Sharpe M, Licht RW, Fink P. Cognitive-behavioural group treatment for a range of functional somatic syndromes: randomised trial. Br J Psychiatry. 2012;200(6):499-507.
22. van Ravesteijn H, Lucassen P, Bor H, van Weel C, Speckens A. Mindfulness-based cognitive therapy for patients with medically unexplained symptoms: a randomized controlled trial. Psychother Psychosom. 2013;82(5):299-310.
23. Dreer LE, Berry J, Rivera P, Snow M, Elliott TR, Miller D, et al. Efficient assessment of social problem-solving abilities in medical and rehabilitation settings: a Rasch analysis of the Social Problem-Solving Inventory-Revised. J Clin Psychol. 2009;65(7):653-69.
24. Pilowsky I. Dimensions of hypochondriasis. Br J Psychiatry. 1967;113(494):89-93.
25. Broadbent E, Petrie KJ, Main J, Weinman J. The brief illness perception questionnaire. J Psychosom Res. 2006;60(6):631-7.
26. Moss-Morris RE, Chalder T. Illness representations: where to from here? 16th Conference of the European Health Psychology Society; Kos, 2003.
27. Ryan EG, Vitoratou S, Goldsmith KA, Chalder T. Psychometric properties and factor structure of a long and shortened version of the cognitive and behavioural responses questionnaire. Psychosom Med. 2018;80(2):230-7.
28. Pearlin L, Schooler C. The structure of coping. J Health Soc Behav. 1978;19(1):2-21.
29. Woud ML, Zhang XC, Becker ES, Zlomuzica A, Margraf J. Catastrophizing misinterpretations predict somatoform-related symptoms and new onsets of somatoform disorders. J Psychosom Res. 2016;81:31-7.
30. Rief W, Mewes R, Martin A, Glaesmer H, Braehler E. Are psychological features useful in classifying patients with somatic symptoms? Psychosom Med. 2010;72(7):648-55.
31. Rief W, Broadbent E. Explaining medically unexplained symptoms-models and mechanisms. Clin Psychol Rev. 2007;27(7):821-41.
32. Barsky AJ, Ahern DK, Bailey ED, Saintfort R, Liu EB, Peekna HM. Hypochondriacal patients' appraisal of health and physical risks. Am J Psychiatry. 2001;158:783-7.
33. Marcus DK, Church SE. Are dysfunctional beliefs about illness unique to hypochondriasis? Journal of Psychosomatic Research. 2003;54(6):543-7.
34. Smeets G, de Jong PJ, Mayer B. If you suffer from a headache, then you have a brain tumour: domain-specifc reasoning `bias' and hypochondriasis. Behaviour Research and Therapy. 2000;38:763-76.
35. Dowrick CF, Ring A, Humphris G, Salmon P. Normalisation of unexplained symptoms by general practitioners: a functional typology. British Journal of General Practice. 2004;54:165-70.
36. Salmon P, Peters S, Stanley I. Patients' perceptions of medical explanations for somatisation disorders: qualitative analysis. BMJ. 1999;318:372-6.
37. Sumathipala A, Siribaddana S, Abeysingha MR, De Silva P, Dewey M, Prince M, et al. Cognitive-behavioural therapy v. structured care for medically unexplained symptoms: randomised controlled trial. Br J Psychiatry. 2008;193(1):51-9.
38. Arnold IA, de Waal MWM, Eekhof JAH, Assendelft WJJ, Spinhoven P, van Hemert AM. Medically unexplained physical symptoms in primary care: a controlled study on the effectiveness of cognitive-behavioral treatment by the family physician. Psychosomatics. 2009;50(5):515-24.
39. Burton C, Weller D, Marsden W, Worth A, Sharpe M. A primary care Symptoms Clinic for patients with medically unexplained symptoms: pilot randomised trial. BMJ Open. 2012;2:e000513.
40. Rosendal M, Blankenstein AH, Morriss

R, Fink P, Sharpe M, Burton C. Enhanced care by generalists for functional somatic symptoms and disorders in primary care. Cochrane Database Syst Rev. 2013;10, CD008142.
41. APA. Diagnostic and statistical manual of mental disorders, 5th ed. (DSM-5). Arlington, VA: American Psychiatric Publishing; 2013.
42. Claassen-van Dessel N, van der Wouden JC, Dekker J, van der Horst HE. Clinical value of DSM IV and DSM 5 criteria for diagnosing the most prevalent somatoform disorders in patients with medically unexplained physical symptoms (MUPS). J Psychosom Res. 2016;82:4-10.
43. Lian OS, Hansen AH. Factors facilitating patient satisfaction among women with medically unexplained long-term fatigue: A relational perspective. Health (London). 2016;20(3):308-26.
44. De Gucht V, Heiser W. Alexithymia and somatisation. Journal of Psychosomatic Research. 2003;54(5):425-34.
45. Pedrosa Gil F, Ridout N, Kessler H, Neuffer M, Schoechlin C, Traue HC, et al. Facial emotion recognition and alexithymia in adults with somatoform disorders. Depress Anxiety. 2009;26(1):E26-33.
46. Kleinstauber M, Gottschalk J, Berking M, Rau J, Rief W. Enriching cognitive behavior therapy with emotion regulation training for patients with multiple medically unexplained symptoms (ENCERT): design and implementation of a multicenter, randomized, active-controlled trial. Contemp Clin Trials. 2016;47:54-63.
47. van Ravesteijn HJ, Suijkerbuijk YB, Langbroek JA, Muskens E, Lucassen PL, van Weel C, et al. Mindfulness-based cognitive therapy (MBCT) for patients with medically unexplained symptoms: process of change. J Psychosom Res. 2014;77(1):27-33.



SUMMARY

Chapter 1 introduces the concept of Medically Unexplained Physical Symptoms (MUPS), its measurement, current management in primary care and the associated costs. Experiencing MUPS is common for all people and does not necessarily lead to problems. However, experiencing many MUPS from various organ systems may imply 'somatization'. Somatization is a tendency to experience and communicate somatic distress and symptoms unaccounted for by pathological findings, to attribute them to physical illness, and to seek medical help for them. Previous research found that the number of symptoms predicts the course of MUPS. Also the rationale for our intervention and implementation considerations are provided in the first chapter. Finally the research questions investigated in this thesis are presented. The research questions are:

1) What is the best self-report measurement instrument to measure somatization in primary care?
2) What does the current management of MUPS patients in Dutch primary care entail?
3) What is the effectiveness of a cognitive behavioural therapy (CBT) intervention for patients with undifferentiated somatoform disorder carried out by mental health nurse practitioners in Dutch primary care?
4) Is the new CBT intervention cost-effective compared to current usual care?
5) How did patients and MHNPs feel about and evaluate the CBT intervention for undifferentiated somatoform disorder?

Chapter 2 presents results of a systematic review of clinimetric properties of self-report questionnaires that measure somatization among adult primary care patients. We assessed the methodological quality of the included studies using the COnsensus-based Standards for the selection of health Measurement Instruments (COSMIN) checklist. We included 24 articles describing 9 questionnaires. Most studies we found investigated the Patient Health Questionnaire 15-item somatic symptom severity scale (PHQ-15) and the somatization subscale of the 4-Dimensional Symptom Questionnaire (4DSQ). These two questionnaires had the highest quality considering various measurement properties (internal consistency, test-retest reliability, structural validity and construct validity). The PHQ-15 had good criterion validity, whereas the 4DSQ was validated in several languages. Several other questionnaires (Bodily Distress Syndrome checklist, Physical Symptom Checklist and the somatization subscale of the Symptom Checklist) showed promising results but were studied in a small number of studies, making

robust conclusions impossible. We therefore recommend the use of either the PHQ-15 somatic symptom severity scale or the 4DSQ somatization subscale for measurement of somatization in primary care.

Chapter 3 contains the detailed description of the study design of the CIPRUS study, a cluster randomized controlled trial investigating the effectiveness of a short-term cognitive behavioural intervention for undifferentiated somatoform disorder, provided by a mental health nurse practitioner, compared to usual care. The intervention consisted of 6 sessions of 30 minutes each and was based on the 'consequences model' and Problem-Solving Treatment (PST). The 'consequences model' shifts the focuses of treatment from the cause of the symptoms to their consequences in the patients' daily life. Using PST, the negative consequences are addressed one by one and patients acquire general problem-solving skills. Furthermore, this chapter provides a description of the primary and secondary outcomes, the choice for potential moderating factors, and the (rationale behind the) various statistical analyses. It describes that we chose for assessments at 0, 2, 4, 8 and 12 months and how we planned to select 212 adult patients from Dutch surgeries.

In **chapter 4** we present results from an observational study on the current management of MUPS patients within Dutch general practices. The Dutch College of General Practitioners published a guideline for the management of MUPS in 2013. We were interested to know whether general practitioners (GPs) adhere to the guideline and whether there were changes over time. In the observational study we screened routinely recorded health care data from electronic medical records of 77 patients participating in the 30 general practices included in the control group of the CIPRUS study. Data on GPs' management strategies were collected over the past five years for each patient and were categorized into diagnostic and therapeutic management strategies. Results showed that the most common diagnostic strategies used by GPs were physical examination (24.5%) and additional investigations by the GP (11.1%). Most common therapeutic strategies were prescribing medication (24.6%) and providing explanations (11.2%). GPs tended to adjust medication, discuss progress and schedule follow-up appointments as symptoms persisted. Surprisingly, exploring the symptoms according to all complaint dimensions (not only somatic but also cognitive, emotional, behavioural and social) as recommended by the guideline, and referrals to a psychologist or psychiatrist were among the least frequently reported strategies (3.5%, 0.5% and 0.1% respectively). Therefore, our results suggest that management of MUPS patients by the GPs was partly in line with the guideline, but some core elements were missing.

In **chapter 5** we present the clinical results of our randomized controlled trial, the CIPRUS study, where we investigated the effectiveness of a cognitive behavioural intervention for patients with undifferentiated somatoform disorder, delivered by mental health nurse practitioners (MHNPs). Practices were randomly assigned to the intervention or usual care group. The intervention consisted of six sessions with the MHNP. The usual care group received care that they normally would for their MUPS from the GP or potential other health care providers the GP might have referred them to. The primary outcome was physical functioning. Secondary outcomes were various aspects of quality of life, mental functioning, anxiety, depression and somatic symptom severity. There was a one-year follow-up. There were 111 participants in the intervention group and 87 in the usual care group. Compared to usual care, participants in the intervention group showed improvement in physical functioning, less limitations due to physical problems and less pain over 12 months. Although significant, these effects were not very large. Effects were larger and more relevant for patients whose symptoms developed more recently and who had fewer physical diseases. No significant effects were found for the remaining outcomes. Therefore, we concluded that our cognitive behavioural intervention was effective in improving physical functioning and decreasing pain, and it was particularly suitable for patients with symptoms that had been present for a limited number of years and with few comorbid physical diseases.

Chapter 6 presents the results of the cost-effectiveness of the CIPRUS study. We performed an economic evaluation from a societal and healthcare perspective with a 12 months follow-up. The primary outcomes were quality-adjusted life-years (QALYs), physical functioning, somatic symptom severity, anxiety and depression. Over a period of 12 months, the mean total healthcare costs in the intervention group were significantly lower than in the usual care group. At a willingness to pay of 0 € per additional unit of effect, the probability of the intervention being cost-effective was 0.93 for QALYs and 0.92 for physical functioning, somatic symptom severity, anxiety and depression. Therefore, the intervention was cost-effective compared to usual care. This implies that implementation of such an intervention on a larger scale would result in a decline in healthcare costs, but would obviously require increased MHNP capacity.

In **chapter 7** we provide the process evaluation of the CIPRUS study, where the MHNPs' and patients' experiences with the intervention are described. MHNPs provided information for the process evaluation by participating in semi-structured interviews with the researcher. Patients in the intervention group of the CIPRUS study provided information by completing written evaluation questionnaires. Overall, MHNPs reported that the intervention manual gave them useful tools for working with patients with

undifferentiated somatoform disorder. MHNPs also reported that the intervention seemed effective for the patients, especially for those with less comorbidity and psychosocial problems and those open to change. MHNPs reported that they generally adhered to the intervention manual, but adjusted the length of the sessions, considering the given 30 minutes too short. They would have felt more comfortable having more flexibility when using the treatment manual. They also indicated that the intervention did not seem suitable for at least a third of the participating patients due to psychiatric comorbidity, psychosocial problems, lower IQ and older age. In the future, MHNPs and GPs could play a larger role in patient selection. Regarding the patients, half of them reported the intervention helped them at least somewhat and a third were positive about following the intervention in the future, if needed. Patients with a shorter symptom duration were more likely to report that the intervention was helpful.

In the general discussion (**chapter 8**) we provide a reflection on the main findings of this thesis. We embed our results in what is already known on the topic and discuss methodological considerations, clinical implications and provide recommendations for future research.

In conclusion, this thesis contributes to the body of knowledge on MUPS and particularly on its measurement and management, and on the effectiveness and cost-effectiveness of a cognitive behavioural intervention in primary care. We investigated which two questionnaires are best suited for measuring MUPS and learned that Dutch GPs partly follow the Dutch guideline for MUPS, but seem to neglect several essential elements. We also provided the first evidence for the effectiveness of an intervention for undifferentiated somatoform disorder delivered by MHNPs, and found that a cognitive behavioural intervention with a short duration improved physical functioning and decreased limitations and bodily pain, compared to current usual care. The intervention was particularly effective in patients with a shorter symptom duration and few comorbid physical diseases. The intervention was also cost-effective compared to usual care, so it seems that its implementation on larger scale would decrease costs. Resources would have to be allocated differently for this purpose. Finally the intervention did not seem to be effective for patients with a longer symptom duration and higher number of comorbid physical diseases. Future research should focus on discovering what type of interventions would be effective for these groups of patients.

DUTCH SUMMARY/NEDERLANDSE SAMENVATTING

In **hoofdstuk 1** worden somatisch onvoldoende verklaarde lichamelijke klachten (SOLK) uitgelegd, het meten ervan, het huidige beleid in de eerste lijn en de kosten ervan. SOLK komt veel voor en hoeft niet tot problemen te leiden. Echter, last hebben van SOLK in diverse orgaansystemen kan wijzen op 'somatisatie'. Somatisatie is een neiging om stress lichamelijk te ervaren en te uiten, en onverklaarde klachten toe te schrijven aan een lichamelijke ziekte en hier medische hulp voor te zoeken. Eerder onderzoek heeft aangetoond dat het aantal klachten het beloop van SOLK voorspelt. Daarnaast beschrijven we in het eerste hoofdstuk de achtergronden van de onderzochte interventie en overwegingen voor de implementatie ervan. Tot slot beschrijven we de onderzoeksvragen die in dit proefschrift worden beantwoord. De onderzoeksvragen zijn:

1) Wat is het beste self-report meetinstrument om somatisatie in de eerste lijn de meten?
2) Wat houdt het huidige beleid voor patiënten met somatisch onvoldoende verklaarde lichamelijke klachten (SOLK) in de Nederlandse huisartsenpraktijken in?
3) Wat is de effectiviteit van een cognitief gedragsmatige interventie voor patiënten met een ongedifferentieerde somatoforme stoornis uitgevoerd door de praktijkondersteuner van de huisarts voor de geestelijke gezondheidszorg (POH-GGZ)?
4) Is de nieuwe cognitief gedragsmatige interventie kosteneffectief vergeleken met gebruikelijke zorg?
5) Hoe hebben de POH's-GGZ en patiënten de cognitief gedragsmatige interventie voor ongedifferentieerde somatoforme stoornis ervaren en hoe beoordelen ze die?

Hoofdstuk 2 beschrijft de resultaten van een systematische review van klinimetrische eigenschappen van zelf gerapporteerde vragenlijsten die somatisatie meten in volwassen eerstelijns patiënten. Wij onderzochten de methodologische kwaliteit van de geïncludeerde studies aan de hand van de COnsensus-based Standards for the selection of health Measurement Instruments (COSMIN) checklist. Wij includeerden 24 artikelen over 9 vragenlijsten. De meeste studies gingen over de somatische symptoomernst schaal van de Patient Health Questionnaire 15-item (PHQ-15) en de somatisatie-subschaal van de 4-dimensionele klachtenlijst (4DKL). Deze twee vragenlijsten scoorden het best op diverse meeteigenschappen (interne consistentie, test-hertest betrouwbaarheid, structurele validiteit en constructvaliditeit). De PHQ-15 had goede

criteriumvaliditeit, en de 4DKL is gevalideerd in verschillende talen. Een aantal andere vragenlijst (Bodily Distress Syndrome checklist, Physical Symptom Checklist en de somatisatiesubschaal van de Symptom Check-list) lieten veelbelovende resultaten zien, maar waren slechts in een klein aantal studies onderzocht, waardoor er geen robuuste conclusies getrokken konden worden. Voor het meten van somatisatie in de eerste lijn, raden wij daarom het gebruik van de somatische symptoomernst schaal van de PHQ-15 aan of de somatisatiesubschaal van de 4DKL.

Hoofdstuk 3 geeft een gedetailleerde beschrijving van de opzet van de CIPRUS studie, een cluster gerandomiseerd gecontroleerde trial naar de effectiviteit en kosteneffectiviteit van een kortdurende cognitieve gedragsinterventie voor ongedifferentieerde somatoforme stoornis, uitgevoerd door de POH-GGZ, vergeleken met de gebruikelijke zorg. De interventie bestond uit 6 sessies van 30 minuten en was gebaseerd op het 'gevolgenmodel' en Problem-Solving Treatment (PST). In het 'gevolgenmodel' wordt in behandeling de nadruk niet op de oorzaak, maar op de gevolgen van de klachten voor het dagelijks leven van de patiënten gelegd. Middels PST worden de negatieve gevolgen stuk voor stuk aangepakt en worden patiënten over het algemeen vaardiger in het oplossen van problemen. In dit hoofdstuk worden verder de primaire en secundaire uitkomstmaten, de keuze van potentieel modererende factoren en de verschillende statistische analyses gepresenteerd. Het beschrijft waarom we voor 0, 2, 4, 8 en 12 maanden follow-up hebben gekozen en beoogd hebben om 212 volwassen patiënten in Nederlandse huisartsenpraktijken te werven.

In **hoofdstuk 4** geven we de resultaten weer van een observationele studie over het huidige beleid voor SOLK binnen Nederlandse huisartsenpraktijken. In 2013 publiceerde het Nederlandse Huisartsengenootschap (NHG) een nieuwe richtlijn voor SOLK. Wij waren benieuwd in hoeverre de Nederlandse huisartsen zich aan de richtlijn houden en of er veranderingen zijn in de tijd. In deze observationele studie zijn gegevens over beleid uit de huisartsinformatiesystemen (HIS) van 77 patiënten in 30 huisartsenpraktijken gehaald, die deelnamen aan de controlegroep van de CIPRUS studie. Gegevens over het beleid van huisartsen waren verzameld over de afgelopen 5 jaar en ingedeeld in diagnostisch en behandelbeleid. Resultaten lieten zien dat het vaakst gebruikte diagnostisch beleid lichamelijk onderzoek (24,5%) en aanvullend onderzoek binnen de praktijk (11,1%) was. Het meest voorkomend behandelbeleid was: medicatie voorschrijven (24,6%) en uitleg geven (11,2%). Naarmate de klachten langer duurden, hadden de huisartsen de neiging om medicatie aan te passen, de voortgang vaker te bespreken en vaker controleafspraken in te plannen. Tegen onze verwachting in, waren het exploreren van klachten aan de hand van diverse klachtdimensies (niet alleen somatisch, maar ook cognitief, emotioneel,

gedragsmatig en sociaal), zoals geadviseerd in de richtlijn, en verwijzingen naar een psycholoog of psychiater, de minst genoteerde beleidsstrategieën (respectievelijk 3,5%, 0,5% en 0,1%). Concluderend geven onze resultaten aan dat het beleid voor SOLK door de huisartsen deels overeenkwam met de richtlijn van het NHG, maar een aantal belangrijke onderdelen ontbraken.

In **hoofdstuk 5** presenteren we de klinische resultaten van onze gerandomiseerde gecontroleerde trial, de CIPRUS studie, waarin de effectiviteit werd onderzocht van een cognitief gedragsmatige interventie voor patiënten met een ongedifferentieerde somatoforme stoornis, uitgevoerd door de POH-GGZ. Deelnemers waren at random toegedeeld aan de interventie- of gebruikelijke-zorggroep. De interventie bestond uit zes sessies met de POH-GGZ. De gebruikelijke-zorggroep kreeg zorg die ze normaal zouden krijgen voor SOLK van de huisarts of andere potentiele zorgaanbieders waar de huisarts mogelijk naar verwees. De primaire uitkomstmaat was fysiek functioneren. Secundaire uitkomstmaten waren diverse onderdelen van kwaliteit van leven, mentaal functioneren, angst, depressie en ernst van lichamelijke klachten. De deelnemers werden een jaar gevolgd. Er deden 111 mensen mee in de interventiegroep en 87 in de gebruikelijke-zorggroep. Vergeleken met de gebruikelijke-zorggroep, hadden de patiënten uit de interventiegroep een hogere mate van fysiek functioneren, minder beperkingen door fysieke problemen en minder pijn over de follow-upperiode van 12 maanden. De effecten waren statistisch significant, maar niet heel groot. De effecten waren groter en relevanter voor patiënten met een kortere duur van klachten en met minder fysieke comorbiditeit. Er waren geen significante effecten op de andere uitkomstmaten. We concluderen daarom dat onze cognitief gedragsmatige interventie effectief was in het verbeteren van fysiek functioneren en verminderen van pijn, met name bij patiënten met een kortere klachtenduur en kleiner aantal comorbide fysieke aandoeningen.

Hoofdstuk 6 geeft de resultaten weer van de kosteneffectiviteit van de CIPRUS studie. We hebben een economische evaluatie uitgevoerd vanuit het maatschappelijk en gezondheidszorgperspectief met 12 maanden follow-up. De uitkomstmaten waren quality-adjusted life-years (QALYs), fysiek functioneren, ernst van de lichamelijke klachten, angst en depressie. Over een periode van 12 maanden, waren de gemiddelde gezondheidszorgkosten in de interventiegroep significant lager dan in de gebruikelijke-zorggroep. Bij een bereidheid om 0 € voor een extra eenheid van effect te betalen, was de kans dat de interventie kosteneffectief was 0,93 voor de QALYs, en 0,92 voor fysiek functioneren, ernst van de lichamelijke klachten, angst en depressie. De interventie was dus kosteneffectief vergeleken met de gebruikelijke zorg. Dit betekent dat het

implementeren van deze interventie op grotere schaal zou zorgen voor een afname in gezondheidszorgkosten, maar een grotere inzet van POH-GGZ zou nodig zijn.

In **hoofdstuk 7** beschrijven we de procesevaluatie van de CIPRUS studie, waarin de ervaringen van POH's-GGZ en patiënten met de interventie worden beschreven. POH's-GGZ namen deel aan een semigestructureerd interview met de onderzoeker. Patiënten uit de interventiegroep van de CIPRUS studie vulden schriftelijk een evaluatievragenlijst in. Over het algemeen gaven POH's-GGZ aan dat de interventiehandleiding hen bruikbare handvatten gaf om met patiënten met een ongedifferentieerde somatoforme stoornis te werken. POH's-GGZ gaven ook aan dat de interventie effectief leek te zijn voor de patiënten, vooral voor degenen met minder comorbiditeit en psychosociale problemen en patiënten die open waren voor verandering. POH's-GGZ gaven aan dat ze zich over het algemeen hebben gehouden aan de interventiehandleiding, maar dat sommigen af en toe de duur van de sessies aanpasten, aangezien ze 30 minuten te kort vonden. Ze gaven aan behoefte te hebben aan meer flexibiliteit tijdens het gebruik van de handleiding. Verder gaven ze aan dat de interventie niet geschikt leek te zijn voor minstens een derde van de deelnemende patiënten. Redenen waren onder andere psychiatrische comorbiditeit, psychosociale problemen, lager IQ en oudere leeftijd. In het vervolg zouden POH's-GGZ en huisartsen een grotere rol kunnen spelen bij patiëntenselectie.

Van de patiënten gaf de helft aan dat de interventie hen minstens enigszins hielp en een derde was positief over het nogmaals volgen van de interventie in de toekomst (indien nodig). Patiënten met een kortere klachtenduur gaven vaker aan dat ze de interventie als helpend hebben ervaren.

In de algemene discussie (**hoofdstuk 8**) presenteren we een terugblik op de hoofdresultaten van het proefschrift. Wij bespreken onze resultaten tegen de achtergrond van de reeds bekende literatuur en discussiëren over methodologische overwegingen, klinische toepassingen en doen aanbevelingen voor toekomstig onderzoek.

Samenvattend draagt dit proefschrift bij aan onze kennis over SOLK en specifiek over het meten ervan, het beleid en effectiviteit en kosteneffectiviteit van een cognitief gedragsmatige interventie in de huisartsenpraktijk. Wij hebben onderzocht welke vragenlijsten het meest geschikt zijn voor het meten van SOLK en zijn erachter gekomen dat Nederlandse huisartsen zich deels houden aan de SOLK richtlijn, maar dat er ook belangrijke onderdelen in hun beleid missen. Wij beschrijven het eerste bewijs voor effectiviteit van een interventie voor ongedifferentieerde somatoforme stoornis uitgevoerd door de POH-GGZ, en hebben gevonden dat een kortdurende cognitieve

gedragsinterventie fysiek functioneren verbeterde, en beperkingen door lichamelijke problemen en pijn verminderde, vergeleken met gebruikelijke zorg. Deze interventie was vooral effectief bij patiënten met een kortere duur van klachten en weinig comorbide fysieke aandoeningen. De interventie was ook kosteneffectief vergeleken met de gebruikelijke zorg, dus het lijkt erop dat een grootschalige implementatie ervan tot kostenvermindering zou leiden. Financiële middelen zouden echter anders verdeeld moeten worden. Tot slot was de interventie niet effectief bij patiënten met een langere duur van klachten en een hoger aantal fysieke aandoeningen. Toekomstig onderzoek zou kunnen focussen op uitvinden welke typen interventies wel effectief zijn voor deze patiënten.

DANKWOORD

Allereerst wil ik alle deelnemende **patiënten**, **huisartsen**, **POH's-GGZ en huisartsassistenten** enorm bedanken voor hun bijdrage aan de CIPRUS-studie. Wetenschap zonder deelnemers is onmogelijk, tegelijkertijd is het moeilijk om deelnemers te vinden, zeker als er niet veel persoonlijke winst tegenover lijkt te staan. Dus alle gemotiveerde deelnemers, in welke vorm dan ook, ontzettend bedankt voor jullie (altruïstische) inzet voor mijn project en de wetenschap, jullie tijd, feedback en hulp. Zonder jullie was dit proefschrift en de kennis die we ermee hebben vergaard onmogelijk tot stand gekomen!

Dan de CIPRUS-studie **projectgroep**, de mensen met wie ik jarenlang nauw samen heb gewerkt en die mij hebben gesteund, geholpen, op sommige momenten erdoorheen hebben getrokken en er voor me waren toen het lastig werd. Eerst een groepsdank, want ik begrijp vaak dat promotie-trajecten kunnen vastlopen door o.a. (een tekort aan of juists te strakke) begeleiding. Ik heb daarentegen mijn begeleiding zo goed als PERFECT ervaren. De combinatie van jullie kennis en skills en vooral de balans tussen vrijheid (ik kon alles opperen en doen wat ik wilde) en sturing (ik voelde me nooit in het diepe gegooid) hebben jullie als team voortreffelijk gevonden. Jullie hebben de afgelopen jaren voor mij soepel en makkelijk laten verlopen, iets wat ik vanzelfsprekend ben gaan vinden, maar wat eigenlijk heel bijzonder is!

Lieve **Henriëtte**, hartelijk dank voor je intellectuele bijdrage aan de CIPRUS-studie. Ik heb veel van jou geleerd wat betreft zakelijke, wetenschappelijke code of conduct en het uitvoeren van gedegen wetenschappelijk onderzoek. Jouw oog voor een gelijke verdeling tussen mannen en vrouwen binnen de wetenschap en ook op andere manieren feministische blik bewonder en waardeer ik enorm. Dank voor zowel je inhoudelijke ondersteuning als ook de warme en persoonlijke support die ik van jou heb gekregen.

Lieve **Hans**! 1000 x maal dank voor alles. Jij verdient eigenlijk je eigen pagina, zo dankbaar ben ik je voor je bestaan. Jij bent echt een uniek persoon en begeleider. Je mode (met name t-shirt met '*p<0.005: because effect size matters*'), briljante brein, buitengewone vaardigheden (en dwangmatige belangstelling) in het bekijken van referenties, altijd voor mij klaarstaan, je oprechte interesse en altijd checken hoe het met me gaat, je positieve stemming, de laagdrempeligheid die je combineert met professionaliteit en het geven van vertrouwen zijn allemaal een medaille waard! Ik heb medelijden met alle promovendi die in de toekomst niet de kans zullen krijgen om door jou te worden begeleid. Ik heb

echt geluk gehad met jou, mijn promotietraject is zo veel aangenamer geweest door jouw betrokkenheid!

Lieve **Harm**, dank voor je creatieve ingevingen en bredere blik, ook vanuit het buitenland. Ik kon het altijd heel goed met je vinden en als ik wilde filosoferen of out-of-the-box brainstormen, was ik bij jou op het juiste adres. Hartelijk dank voor het maken van tijd voor mij vanuit Engeland, het goed gehumeurd en gastvrij zijn!

Lieve **Stephanie**, zonder jou was dit hele project überhaupt niet tot stand gekomen! Bedankt voor het aanvragen van de subsidie voor dit project en dan je 'kindje' los durven laten en overdragen aan een (toen nog) onervaren en jou onbekende nieuwe collega (aan mij dus). Ik had het gevoel dat je vanaf dag 1 vertrouwen in mij en mijn kunnen had en dit bleef eigenlijk alleen maar groeien naarmate het project verliep. Ook wanneer mijn eigen vertrouwen in mezelf m.b.t. het onderzoek ver te zoeken was, had jij juist de meest steunende woorden. Daarnaast dank voor je (bijna intimiderende) intelligentie waarmee je altijd de puntjes op de i in het project zette, kritische vragen stelde waar dat nodig was, maar ook mij geruststelde en steunde op persoonlijk vlak. Ik vond het altijd fijn dat we regelmatig belden en dat je ook checkte of ik niet te veel aan het werk was en of de rest van de projectgroep wel lief genoeg tegen mij was. Nou had ik toevallig geluk, maar heb jou altijd als een onafhankelijke steun op afstand ervaren die het beste met me voor had.

Beste leden van de **leescommissie**, dr. Nettie Blankenstein, dr. Joran Lokkerbol, prof. dr. Patricia van Oppen, prof. dr. Judith Rosmalen en prof. dr. Niek de Wit. Hartelijk dank voor het zo nauwkeurig bestuderen en goedkeuren van mijn proefschrift.

Onderzoeksassistenten Sandra en Danielle H.: Het was soms even slikken dat mijn onderzoeksassistenten slimmer zijn dan ik (jullie zijn allebei ook niet voor niks aan een promotieonderzoek begonnen, Sandra is zelfs nog eerder klaar dan ik), maar daar heb ik mee leren leven ;) Bedankt, toppers, voor jullie hulp, rotklusjes, scherpheid, mij wijzen op de (door mezelf gestelde) deadlines, mijn eigen slordigheidsfouten, etc. Jullie zijn beide buitengewoon intelligente mensen en ik heb geboft met jullie op ons project!

DAN DE OVERIGE VUMC COLLEGA'S:

Allereerst mijn roomie **Danielle v/d L**.! Ons natuurlijk selectieproces voor elkaar op de afdeling is heel soepel verlopen en we waren al gauw moeilijk van elkaar te scheiden. We regelden dan ook gelijk dat we bij elkaar op de kamer konden werken. Met jou heb ik jarenlang meer tijd per week doorgebracht dan met mijn eigen partner en meer actuele

gebeurtenissen uit mijn leven gedeeld dan met wie dan ook. Ik vind het heerlijk hoe veel ik over je relatie wist zonder je vriend ooit te hebben ontmoet (Hoi Dion!) We hebben samen menig kamergenoot weggejaagd omdat ze er niet tegen konden dat we zo veel met elkaar kletsten en lachten. Je hebt mijn tijd op de afdeling verrijkt met je aanwezigheid, humor, directheid, steun en last but not least je aantrekkelijke hoofd dat net boven je bureau vandaan kwam achter het computerscherm. Je bent een fantastische collega, vriendin, partner en moeder!

Alle andere **H&O (ex)-collega's**, bedankt voor de gezellige tijd samen op de afdeling! In het bijzonder: Emiel, Floor, Hanneke, Hieke, Karolien, Lidy, Maarten, Madelon, Marlous, Mette, Mieke, Nikki, Tjarda en Vincent. Ik vond het FANTASTISCH met jullie! Door jullie aanwezigheid was naar het werk gaan elke dag een feest (nou ja, soort van dan)! Dank jullie wel voor de mooie tijden samen, voor alle lunches, koffies, *roddels* (my favorite), borrels, feestjes, reizen, congressen en bij-elkaar-op-de-kamer-hang-momenten die het werk zo leuk hebben gemaakt. Jullie zijn te gek!

Loes & Brahim: Sowieso ontzettend bedankt voor alles wat jullie doen voor de afdeling, maar vooral voor jullie vrolijke, vriendelijke gezichten en administratieve/facilitaire hulp bij de CIPRUS-studie. Voor iemand die niet zo enorm gestructureerd is, als ik, is het een verademing om met jullie samen te werken. Jullie hebben alle logistieke processen binnen mijn promotietraject enorm vergemakkelijkt!

EN ANDERE MENSEN MET WIE IK HEB SAMENGEWERKT:

Dank aan alle **co-auteurs** voor hun bijdragen aan onze artikelen! Berend Terluin: ontzettend bedankt voor je klinimetrische hulp en kennis, met name m.b.t. de 4DKL. Wieneke Mokkink: bijzonder dank voor je hulp bij de klinimetrische review. Judith Bosmans en Aureliano Finch: veel dank voor de tijd en energie gestopt in de CEA analyses en artikel. Jos Twisk: veel dank voor de statistische ondersteuning! Lyonne Zonneveld: ontzettend bedankt voor het meeschrijven aan het protocol, het trainen van de POH's-GGZ en deelnemen aan de projectgroepoverleggen. Rinske Pret-Oskam: bedankt voor je hulp als stagiaire en het verzamelen van de gegevens voor het artikel over gebruikelijke zorg.

Daarnaast mensen die op andere manieren hebben bijgedragen aan de CIPRUS-studie, Nienke Veldhuijzen: dank voor het randomiseren van onze patienten en Caroline Terwee: dank voor de klinimetrische expertise voor onze klinimetrische review. Carine den Boer: hartelijk dank voor het helpen trainen van de POH's-GGZ in de controlegroep.

Marga van Rhoon en Caroline de Jong: dankzij jullie en jullie netwerk in Zuid-Holland hebben we toegang gehad tot een groot deel van de deelnemende POH's-GGZ en patienten. Ontzettend bedankt voor het geloven in mij en het onderzoek! Marga, dank voor de vanzelfsprekendheid waarmee je mij 'jouw' psychologen liet benaderen en voor het faciliteren van de training van de deelnemende POH's-GGZ binnen jouw praktijk.

Dirk Franssens en Gerwin Wildeboer, bij jullie heb ik voor het eerst de functie POH-GGZ bekleed en heb ik kennis mogen maken met een huisartsenpraktijk van binnenuit. Veel dank voor de mooie en gezellige tijd, ik heb met veel plezier met jullie samengewerkt. Ook veel dank voor het beschikbaar stellen van jullie praktijk voor het uitvoeren van mijn onderzoek!

OCA collega's Guus, Tom, Marlene, Iris en Renée (en uiteraard alle andere collega's die er later bij kwamen), bedankt voor het eindeloos aanhoren van mijn promotie-perikelen, jullie blijvende belangstelling en steun!

MENSEN UIT PRIVÉ KRING:

Mama: Jouw toewijding om mij te onderwijzen begon al heel vroeg, toen ik nog veel te jong was om te lezen, maar jij me al Russisch en Engels leerde. Daarna kwam de basisschool waarop we veel meer qua kennis en taal leerden dan waar dan ook op dat moment in Moskou. Dan uiteraard nog de IGCSE & IB, zodat ik overal ter wereld gelijke kansen zou krijgen. Jouw doel was om mij het best mogelijke onderwijs, of in iedere geval een kans daartoe te geven. Ik ben blij om door deze PhD te hebben gehaald jou hierin trots te hebben gemaakt. Al jouw inspanningen waren het waard. Dank je wel voor het vechten voor en creëren van deze kansen voor mij. Dit kwam ons niet aanwaaien en was zeker niet vanzelfsprekend zoals voor mensen die geboren zijn in een Westerse cultuur. Dank je wel voor het geven van een mooie toekomst waarin ik **vrij** ben, kan denken, zeggen en doen wat ik wil, waar hard werken gewoon beloond wordt, er een eerlijk rechtssysteem is en ik hele andere typen zorgen heb dan waar wij oorspronkelijk vandaan komen. Dank je wel voor het feit dat deze levensstijl 'normaal' voor mij is geworden. Ik hou van jou!

Bob: Dank je wel voor het altijd hebben gesteund van mijn onderwijs en meer recent voor het elke keer vragen wanneer ik eens een keer 'klaar ben met studeren'. Zonder jou waren veel kansen in mijn leven niet mogelijk geweest, ik waardeer enorm de vanzelfsprekendheid waarmee je die aan mij geboden hebt. Hou van jou, dochter Kate xxx

BFF's Carolien, Carlijn, Elisabeth, Hana, Julia, Karin, Meike, Pim, Soffi & Sophie: jullie hebben niks met mijn proefschrift te maken gehad, maar jullie zijn mijn favoriete mensen in de wereld en 1 voor 1 de meest bijzondere, inspirerende, sterke, krachtige ÉN prachtige vrouwen die ik ooit heb ontmoet. Ik bewonder jullie enorm en hou ontzettend veel van jullie. En dat mag gewoon in elke context gezegd worden!

Julia (Blondie): Bestest friend, OH MY GOD LOOK AT US!! Having PhDs and stuff! We are so awesome!!! LOVE YOU!

Elis, in het bijzonder bedankt voor het mega hippe en niet saaie standaard ontwerp van mijn proefschrift!

ABOUT THE AUTHOR

Kate (Ekaterina Viacheslavovna) Sitnikova was born on May 2 1987 in Moscow, Russia, where she went to the primary school that was founded by her mother. At eleven years old she moved to the Netherlands and got her IGSCE and IB diplomas at St. Maartens International School in Haren.

She went on to study Psychology with a Master in Clinical Psychology at the Rijksuniversiteit Groningen and followed a semester abroad at Northern-Arizona University (NAU) in Flagstaff, AZ in USA. She wrote her master thesis on cognitive behavioural treatment (CBT) for patients with inflammatory bowel disease and did her clinical internship at the department of Medical Psychology at the Academic Medical Center in Amsterdam.

After graduating in 2011 she worked as a research assistant, psychologist in a specialized mental health setting for children and adolescents, and mental health nurse practitioner at a general practice. In 2014 she started her PhD project, the CIPRUS study, at the department of General Practice and Elderly Care Medicine at the VU medical center in Amsterdam. The project was successfully completed in 2019. Simultaneously, Kate worked as a part-time psychologist in a multidisciplinary team at OCA, a rehabilitation center for chronic pain and medically unexplained physical symptoms of the musculoskeletal system. During her PhD program, Kate also followed epidemiology courses and registered as an epidemiologist.

In 2019 Kate started a two-year postdoctoral psychology healthcare training (GZ-opleiding) at Arkin in Amsterdam. In her first year she worked in a specialized setting for people with moderate and severe mood and anxiety disorders. Currently, in her second year, she works in a FACT team (Flexible Assertive Community Treatment), with patients with severe psychiatric illnesses.

Kate has a particular interest in psychodynamic psychotherapy, personality psychology and sexology and is keen on developing her knowledge in these areas in the future.

LIST OF PUBLICATIONS

International publications

1. Jacobs M, Macefield RC, Elbers RG, Sitnikova K, Korfage IJ, Smets EM, et al. Meta-analysis shows clinically relevant and long-lasting deterioration in health-related quality of life after esophageal cancer surgery. Qual Life Res. 2014;23(4):1155-76.
2. Dijkstra-Kersten SM, Sitnikova K, van Marwijk HW, Gerrits MM, van der Wouden JC, Penninx BW, et al. Somatisation as a risk factor for incident depression and anxiety. J Psychosom Res. 2015;79(6):614-9.
3. Lo-Fo-Wong DN, Sitnikova K, Sprangers MA, de Haes HC. Predictors of health care use of women with breast cancer: A systematic review. Breast J. 2015;21(5):508-13.
4. Bennebroek Evertsz F, Sprangers MAG, Sitnikova K, Stokkers PCF, Ponsioen CY, Bartelsman J, et al. Effectiveness of cognitive-behavioral therapy on quality of life, anxiety, and depressive symptoms among patients with inflammatory bowel disease: A multicenter randomized controlled trial. J Consult Clin Psychol. 2017;85(9):918-25.
5. Dijkstra-Kersten SMA, Sitnikova K, Terluin B, Penninx B, Twisk JWR, van Marwijk HWJ, et al. Longitudinal associations of multiple physical symptoms with recurrence of depressive and anxiety disorders. J Psychosom Res. 2017;97:96-101.
6. Sitnikova K, Leone SS, Zonneveld LNL, van Marwijk HWJ, Bosmans JE, van der Wouden JC, et al. The CIPRUS study, a nurse-led psychological treatment for patients with undifferentiated somatoform disorder in primary care: study protocol for a randomised controlled trial. Trials. 2017;18(1):206.
7. Sitnikova K, Dijkstra-Kersten SMA, Mokkink LB, Terluin B, van Marwijk HWJ, Leone SS, et al. Systematic review of measurement properties of questionnaires measuring somatization in primary care patients. J Psychosom Res. 2017;103:42-62.
8. Sitnikova K, Pret-Oskam R, Dijkstra-Kersten SMA, Leone SS, van Marwijk HWJ, van der Horst HE, et al. Management of patients with persistent medically unexplained symptoms: a descriptive study. BMC Fam Pract. 2018;19(1):88.
9. Sitnikova K, Leone SS, van Marwijk HWJ, Twisk J, van der Horst HE, van der Wouden JC. Effectiveness of a cognitive behavioural intervention for patients with undifferentiated somatoform disorder: Results from the CIPRUS cluster randomized controlled trial in primary care. J Psychosom Res. 2019;127:109745.
10. Sitnikova K, Finch AP, Leone SS, Bosmans JE, van Marwijk HWJ, van der Horst HE, et al. A brief cognitive behavioural intervention is cost-effective for primary care patients with medically unexplained physical symptoms compared to usual care. Submitted.
11. Sitnikova K, van Marwijk HWJ, Leone SS, van der Horst HE, van der Wouden JC. Process evaluation of a cognitive behavioural intervention for undifferentiated somatoform disorder in Dutch primary care: mental health nurse practitioners' and patients' experiences. Published as chapter in PhD thesis.

National publications

1. Sitnikova K, Dijkstra-Kersten SMA. Online zelfhulp voor de preventie van depressie. Ned Tijdschr Geneeskund 2017;161:D1541.
2. Sitnikova K, Leone SS, van Marwijk HWJ, Twisk J, van der Horst HE, van der Wouden JC. Cognitieve gedragsinterventie door de poh-ggz bij SOLK. Huisarts & Wetenschap 2020;63(3):31-8.
3. Sitnikova K, van der Wouden JC. Gedragsinterventie bij SOLK door de poh-ggz. TPO De Praktijk, februari 2020.