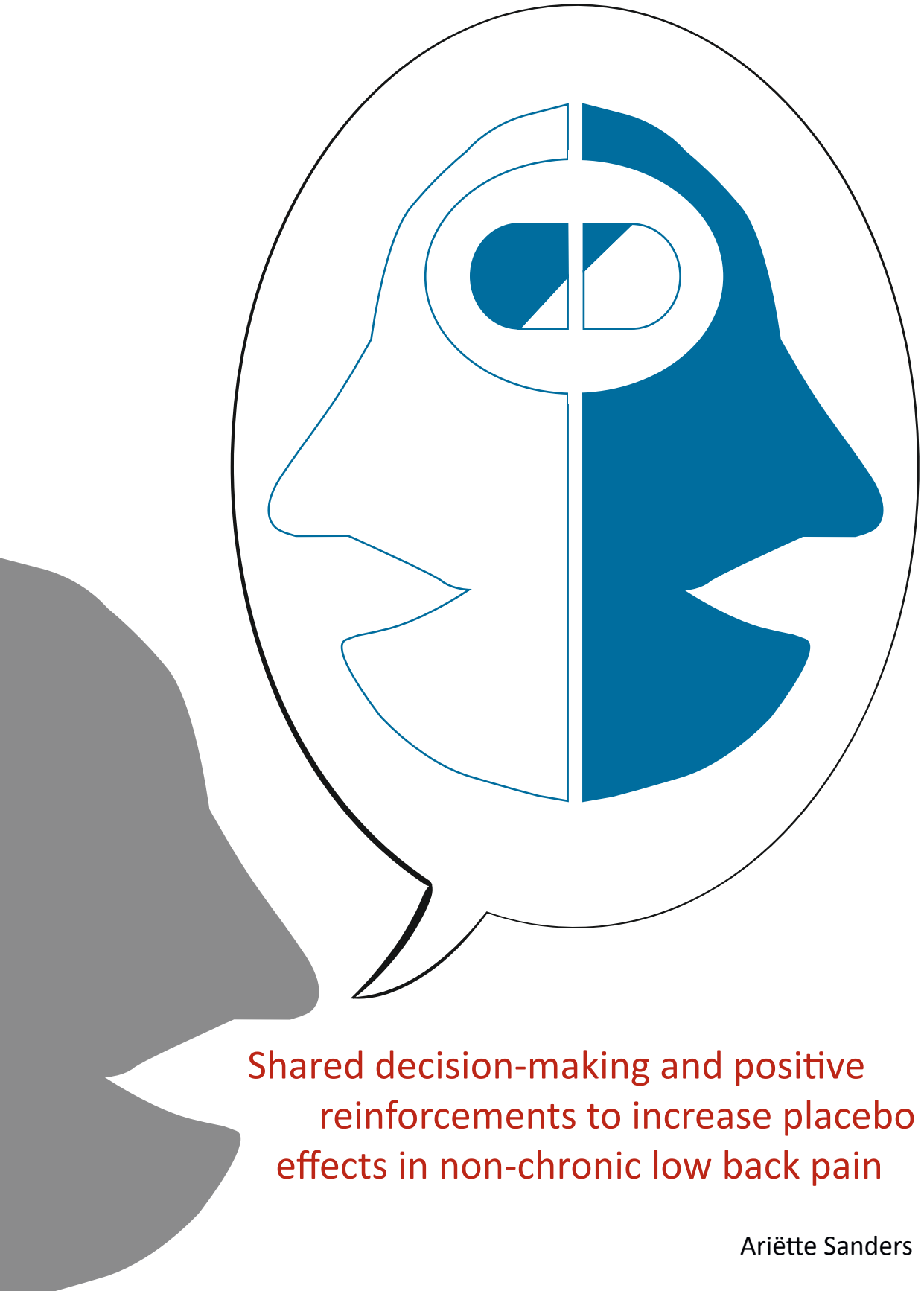


Shared decision-making and positive reinforcements to increase placebo effects in non-chronic low back pain

Ariëtte Sanders

‘Back-to-back consultations’

Shared decision-making and positive reinforcement to increase placebo effects in non-chronic low back pain

‘Ruggespraak’

Samen beslissen en positieve bekrachtiging ter versterking van placebo-effecten bij niet chronische lagerugpijn

Ariëtte Sanders

About the cover

The cover was designed by Hans (J.P.A. van Lennep), son of the author. The cover depicts the complex interplay between communication and expectations in a patient-clinician interaction. The patient shares his non-specific low back pain complaints and the clinician elaborates on patient-centered treatment options. The positive treatment effects are not only established by the pharmaceutical substance, but also by communication-derived placebo effects.

Colophon

Provided by thesis specialist Ridderprint, ridderprint.nl
Printing: Ridderprint
Layout and design: Jeroen Reith, persoonlijkproefschrift.nl

ISBN: 978-94-6483-243-3

Financial support by van Lennep Huisartsenpraktijk for the printing of this thesis is gratefully acknowledged

©A.R.J. Sanders, Utrecht, the Netherlands, 2023

All rights reserved. No part of this dissertation may be transmitted or reproduced in any form or by any means, without prior permission of the author.

‘Back-to-back consultations’

Shared decision-making and positive reinforcement to increase placebo effects in non-chronic low back pain

‘Ruggespraak’

Samen beslissen en positieve bekrachtiging ter versterking van placebo-effecten bij niet-langdurige lagerugpijn

(met een samenvatting in het Nederlands)

Proefschrift

ter verkrijging van de graad van doctor aan de
Universiteit Utrecht
op gezag van de
rector magnificus, prof. dr. H.R.B.M. Kummeling,
ingevolge het besluit van het college voor promoties

in het openbaar te verdedigen op

donderdag 14 september 2023 des middags te 2.15 uur

door
Ariëtte Reina Johanna Sanders

geboren op 26 oktober 1962
te Apeldoorn

Promotoren:

Prof. dr. N.J. de Wit

Prof. dr. A.M. van Dulmen

Beoordelingscommissie:

Prof. dr. R.A.M.J. Damoiseaux

Prof. dr. J.J.M. van Delden

Prof. dr. A.W.M. Evers

Prof. dr. M.J. Schuurmans (voorzitter)

Prof. dr. G.D.E.M. van der Weijden

Dit proefschrift werd (mede) mogelijk gemaakt met financiële steun van ZonMw; project nummers 42011009 en 808330098030.

Open access-kosten voor de review werden gefinancierd door de Nederlandse Organisatie voor Wetenschappelijk Onderzoek.

“Als de wijnfles leeg is, is de appel rot”
Henk Willem B. Sanders

CONTENT

CHAPTER 1

General introduction

INTRODUCTION

Placebo effects in healthcare

In healthcare, positive outcomes can be induced not only by specific medical treatments but also by non-specific interventions, usually referred to as placebo effects. When people think of the placebo effect, they tend to visualise an inert pharmaceutical substance that induces beneficial treatment effects. However, placebo effects are not limited to medication. Scientists agree that placebo effects are better described by the term 'contextual effects' to avoid a too narrow conceptualisation of the placebo effect and also include the context (e.g. psychosocial) that accompanies any medical intervention. (1,2)

The word 'placebo' induces ambiguous feelings. It originates from Latin and literally means "I will please", with the connotation that patients will experience relief. Healthcare professionals aim to relieve patients' suffering but doing so indirectly can feel like deception, because transparency is highly appreciated in healthcare. (3)

In medical research, it proves difficult to discriminate between a true effect of a medical substance or intervention and the contextual effects. (1,4) This is the origin of placebo-controlled trials, which are nowadays considered a key instrument in providing high levels of confidence for the causal effect of an intervention. (5) Considering the important influence of contextual effects on health outcomes, research on the background of this multifactorial phenomenon emerged last century. (4,5) However a complete elucidation of how placebo effects work has not yet been achieved.

Placebo effects are largely explained by positive outcome expectations in patients, a trustworthy patient-professional relationship and an improved affective status of the patient, leaving a small rest group containing biological factors such as genomic background. (1,4,6-8)

Learning is assumed to be one of the underlying mechanisms of ***outcome expectations***, which can be classical conditioning, instructional learning, and social learning. (9) As far back as 1927, Pavlov demonstrated that dogs can become restless (effect of morphine on dogs) on hearing a bell when conditioned by the repeated administrations of morphine on the sound of the bell. Likewise, contextual clues such as a white coat or a pill can evoke conditioned responses like pain relief after pairing with an active treatment e.g. a painkiller (unconditioned stimulus). A recent example of placebo effect after instructional learning is the relief from dyspnoea after the instruction of a general practitioner (GP) to place the children suffering from laryngitis subglottica in a steamy bathroom. This was a common treatment till 2013, but later proved not to be effective. (10) Finally, social learning occurs when a person experiences pain relief from a treatment although they only observed someone else experiencing the placebo-induced response, such as for instance a video demonstration of someone reporting less pain after the application of a placebo. (11)

Some scientists attribute most or all placebo effects to learning. (9) Others regard a combination of expectations, professional-patient interaction and affective status as

the background of placebo effects. (3,8,12,13) Regardless, most scientists agree that expectations play a significant role. (2,6,12,14-16) Expectations can be developed through experience (conditioning or social learning) or explicit knowledge (verbally suggested or instructed). (11) In an experiment with 66 patients suffering from migraine, patients were given either medication (10 mg rizatriptan) or placebo. There were three different branches, either the patients were told that they had received the drug, the placebo or one of the two (but not which). The untreated migraine episode before intervention served as the control condition. Compared to no treatment, the placebo accounted for more than 50% of the drug effect under each information condition. Even if placebo prescription was openly communicated, participants benefited from the placebo and, in line with the increase of a positive message, the efficacy of both placebo and medication during migraine attacks grew. (17) In conclusion, expectations, even if receiving a placebo, can relieve patients' pain.

The ***patient-professional relationship*** also induces contextual effects. The encounter is characterised by trust and hope on the patients' side and empathy and compassion on the professionals' side. (18) Through verbal and non-verbal communication, for instance by gestures (holding a hand), eye-contact or facial expressions, professionals interact with their patients thus influencing a variety of feelings like anxiety, self-efficacy, hope and trust. (18) In an experiment among 112 depressive patients treated by 9 psychiatrists, half of them received a placebo and the others an antidepressant. The verum was more effective than the placebo but the effectiveness of the psychiatrist was even bigger than the antidepressant. (19) Obviously the patient-professional interaction itself creates placebo effects.

Patients' ***affective status***, such as feelings of anxiety, fear and stress due to the reason for encounter, can be altered positively by receiving trust, hope and support. (8) Research revealed that modulations of pain and anxiety are related to placebo effects. (8,9) They induce variations in activity in parts of the brain that are involved in pain-response regions such as the insula, dorsal anterior cingulate cortex, thalamus, amygdala and right lateral prefrontal cortex. (9,20) On the other hand, placebo effects can be reduced by the use of medications like opioid antagonists. (21)

The role of healthcare professionals in placebo effects

Placebo effects can also be induced or augmented by healthcare professionals. (4,5) Healthcare professionals consider using placebos acceptable especially in case of psychologically determined complaints and stress the importance of being transparent with patients about the use of placebos. (3,22) In 2021, an international scientific panel of the Society of Interdisciplinary Placebo Studies reached consensus on how healthcare professionals could use placebo-communicative behaviour to increase patients' well-being. (3) The recommendation is that patients should be informed based on evidence, even though the scale of effects should not be overestimated and tailored to individual situations. (3) Professionals should be trained to communicate about placebo effects including the underlying neurobiological mechanisms in order to maximise placebo

effects. (3) In this training, they should also be taught how to use their communication in order to enhance placebo effects in their patients as much as possible.

Improving professional communication to enhance placebo effects

As previously mentioned, healthcare professionals can shape context effects by both verbal and non-verbal communication. (1,6,23)

Affective communication, through showing care, concern or empathy, helps create a trusting doctor-patient relationship. (24) This trust can let patients participate in consultations, which, in a trustful context, can lead to more favourable health outcomes. (25)

Moreover, as patients generally tend to trust their healthcare professionals*, cognitive communication* like *information provision* or *a positive statement* about the expected effects of the therapeutic regime that is decided upon has been shown to benefit patients' health. (6,7,26) In 1999, it was demonstrated in UK general practice that, in patients with symptoms without a physical substrate, two thirds of those who received a positive consultation, i.e. the doctor gave the patient a clear diagnosis and informed them confidently that they will be better in a few days, got better two weeks after the consultation. On the other hand, only slightly more than one third who received a negative consultation, i.e. the doctor expressed uncertainty about the diagnosis and the recovery, got better after two weeks. (27) More recently, this effect of communication was quantified in a systematic review. Mistiaen et al. identified 19 cognitive care studies and found that a positive suggestion may reduce pain, whereas a negative suggestion may increase pain. (28) However, the effects were small. In the same review 14 studies focusing on emotional care showed no evidence of a direct effect on pain, although four studies showed a tendency for emotional care to lower patients' pain. (28)

A recent international consensus statement acknowledges substantial placebo effects caused by professional behaviour but also emphasises that placebo effects can merely relieve symptoms such as pain and cannot provide a cure. (3)

Besides a reinforcement of the professional, patients themselves also have *expectations* about the natural course or therapeutic options and these expectations influence recovery. (29-32) They consider themselves to be sufficiently capable of judgement against their own values considering their health or healthcare. (25) However, the incorporation of patients' values, preferences or expectations in healthcare decisions is not yet optimal. (25,33-35) To increase patient participation, patients need to have the feeling of being heard or taken seriously. (5) This feeling can be empowered by a professional *communication style with listening as a central focus*. (5) Exploring individual preferences for a therapeutic approach, for instance through active listening and using open-ended questions, allows these preferences to be incorporated in the treatment decision-making. Thus patients' positive expectations about the therapeutic plan might be increased and patients' recovery may be augmented.

Patient participation has been legally embedded since 1995 in the Medical treatment act (in Dutch: Wet op de Geneeskundige Behandelings Overeenkomst or WGBO). The

WGBO dictates that healthcare professionals should encourage patient participation in decision-making. In a WGBO addendum in 2020, healthcare professionals were encouraged to challenge patients to ask questions and express concerns during the medical visit. (34,36)

However, leaving full responsibility to patients could elicit decisional conflict and produce a feeling of abandonment and unintentionally induce nocebo effects. (33,37) This demonstrates the challenge of the concept of patient participation in decision-making. From a professional point of view, the concept focuses on narrowing the information gap between professional and patient, and handing over decisional influence to the patient, even if there is no consensus. (38) Patients, on the other hand, emphasise the reciprocal relationship and advocate that shared decision-making should at least include sharing information, an open-minded and a respectful attitude from both the professional and the patient, patient self-advocacy and a personalised recommendation by the physician. (38,39)

Shared decision-making

The process in which patient and professional together decide upon management plans is known as shared decision-making (SDM). (6,34) It includes four key steps: (1) the professional informs the patient that a decision is to be made and that patients' opinions are important; (2) the professional explains the options and their pros and cons; (3) the professional and the patient discuss the patients' preferences and the professional supports the patient in deliberation; (4) the professional and patient discuss the patient's wish to make the decision, they make or defer the decision, and discuss follow-up. (40) Figure 1.

SDM has been shown to be effective in decreasing decisional conflict while increasing patient satisfaction even if patients say they do not prefer active involvement. (41-44) To make it work, the professional and patient need to assess and discuss the patient's preferred level of participation in decision-making. (37,45,46) Patients' 'expert role' as reflected in their preferences, concerns and self-efficacy, needs to be addressed. (34) When patients' own viewpoints and expectations are incorporated in the chosen therapy, it increases their trust in the management plan, thus potentially increasing placebo effects. (41,47,48) However, implementing this behaviour appropriately among healthcare professionals has proved challenging. It requires a patient-oriented attitude, sufficient knowledge and skills of SDM process steps and potentially more time. (33, 49, 50) Besides that the effects of SDM on patients' health or costs have not yet been sufficiently well demonstrated. (51)

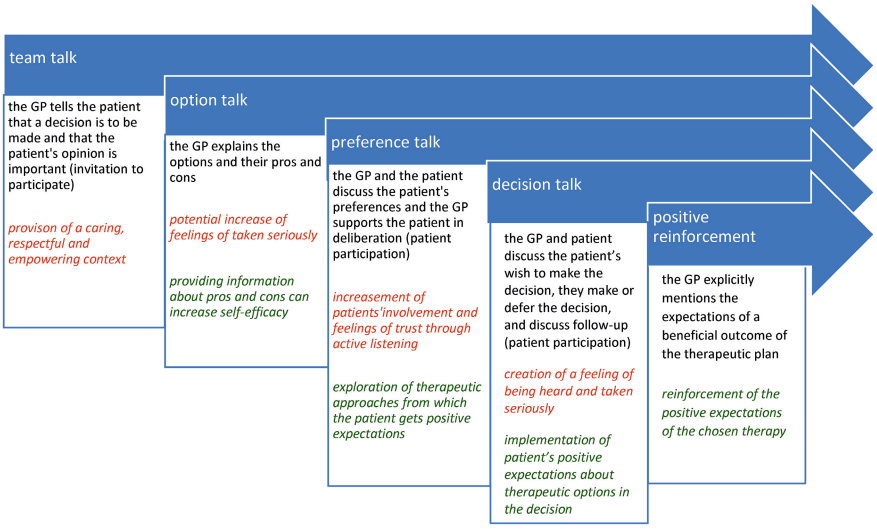


Figure 1. SDM process steps according to Glyn Elwyn and positive reinforcement and how placebo effects might be increased due to affective (red) and cognitive (green) communication. GP=general practitioner.

Strengthening SDM by positive reinforcement

General practice guidelines advise exploring patients' former experiences with recovery of illnesses and treatment effects. (52) Previous positive experiences of themselves (or even others), might result in preferences and positive outcome expectations of these therapeutic options. Patients' preferences and experiences co-determine their perspective in SDM and therapeutic options of which patients have positive outcome expectations might be preferred and chosen. This might especially improve recovery from illnesses where it is known from the literature that recovery is associated with patients' positive outcome expectations like low back pain. (53) Moreover, if patients' positive expectations about the therapeutic plan are not only taken into account but also positively reinforced by the general practitioner (GP), it could be assumed that this might increase recovery rates further. (26,27)

Finally, if this positive reinforcement is done by a *trusted professional,* this might further increase placebo effects. Research demonstrated that recovery of low back pain is associated not only with the quality of the doctor-patient relationship but also with the attitude and expectations of the professional. (54-56) When professionals have higher recovery expectations, they might convey these positive expectations to their patients.

Optimum use of placebo effects in symptom-based illness

In primary care, the majority of complaints presented are diagnosed at a symptom level. They are based on pattern recognition, the absence of reasons for detailed

diagnostic work-up, and the benefit of a positive, 'symptom-based' approach given the favourable prognosis. (57) These symptom-based diagnoses occur in the whole bodily spectrum, varying from abdominal and neurological to musculoskeletal symptoms. This pragmatic approach is reflected in the GPs' recording system, where the first part of each chapter of the International Classification of Primary Care (ICPC) consists of 'Symptoms and complaints'. (58) In symptom-based illnesses (58), the effects of therapeutic interventions rarely exceed placebo effects, and the effects of the placebo often exceed those of the intervention. Optimum use of placebo effects therefore is vital in the management of symptom-based diagnoses.

The example of aspecific low back pain: a symptom-based illness with a high burden

Low back pain is one of the complaints people most frequently present with in primary care. In the Netherlands, GPs registered an incidence of 43.7/1000 people in 2021. (59) Most of these patients suffer from non-chronic complaints (29.7/1000 people). (59) The recorded prevalence in 2021 was 25.9/1000 people suffering from non-chronic aspecific low back pain with some overrepresentation of women (28.0 women/1000 people) in contrast to men (23.8/1000 men). (59) In most patients, no pathophysiological cause is identified and the majority of patients are diagnosed with aspecific low back pain. In this thesis, for reasons of convenience, we will leave out the additional word 'aspecific' to describe low back pain, defined as pain limited to the region between the lower costal margins and the gluteal folds, with or without leg pain. (60-62) Although it subsides in the majority of patients within two weeks, it can become chronic (>3 months) or frequently recurring (≥3 episodes a year). (60) Patients not only suffer from moderate to severe pain and disabilities due to the pain in daily life but are also concerned about the origin of the complaints and the risk of persistence of the complaints. (60,63) Low back pain can therefore have considerable economic effects (63) through losses in productivity (93% of total costs). (60,64,65) As low back pain is one of the leading causes of disability worldwide, many interventions have been studied in order to reduce the burden of low back pain. (66,67)

So far, only advice about activity has proved superior to placebo, whereas therapeutic options like physiotherapy, acupuncture and spinal manipulation did not give effects superior to placebo treatment. (68-70) Besides, painkillers reduce pain (at the expense of side effects) but do not reduce disability due to low back pain. (71,72) To save on costs, guidelines and health insurance policies encourage GPs not to use ineffective (defined as less effective than a placebo intervention) therapeutic interventions for low back pain. (73,74) However, placebo effects in low back pain were demonstrated to be substantial. (75)

It would be worthwhile to investigate how these effects can be enhanced in order to optimise recovery, prevent chronicity (duration >3 months or ≥3 episodes a year) and limit the use of healthcare resources. (75-77) As mentioned above, healthcare professionals may be able to optimise these effects during the doctor-patient encounter.

We therefore developed a communicative strategy and examined its effectiveness (including cost-effectiveness) in patients presenting in general practice with non-chronic low back pain.

Aim and research questions

The aim of this thesis is to assess to what extent patients' recovery from a symptom-based illness such as non-chronic low back pain can benefit from training GPs in optimum use of shared decision-making followed by a positive reinforcement (SDM&PR).

This leads to the following research questions:

- What is known from literature about the relationship between enhanced patient participation and recovery from symptom-based illnesses in primary care, either from the perspective of the patient or in terms of the disease outcomes?
- What effects are there on the communicative behaviour (SDM, positive reinforcement) of GPs after training them in SDM&PR compared to untrained GPs?
- What are the effects of SDM and positive reinforcement by the GP (SDM&PR) compared to usual care on recovery rate in patients suffering from a new episode of non-chronic low back pain?
- What are the effects of patient participation and SDM on recovery from non-chronic low back pain and do these effects differ when perceived from the perspectives of the patients, the professionals or external observers?
- How cost-effective is SDM&PR in patients suffering from non-chronic low back pain?

Chapter 2 gives a description of a systematic study of controlled trials that aim to increase patient participation in face-to-face interactions between healthcare professionals and patients in primary care. As the primary outcome, we focused on recovery either from the perspective of the patient (i.e. pain) or in terms of the disease (i.e. blood pressure).

In Chapter 3, we evaluate a training course consisting of two training sessions that were each two and a half hours in duration. The training focused on the SDM process followed by a positive reinforcement and evidence-based treatment of non-chronic low back pain according to professional guidelines. We video-recorded 175 consultations of patients suffering from non-chronic low back pain and compared recorded GP behaviour from 23 trained GPs against 19 untrained GPs on SDM activities, providing positive reinforcement and on the level of patient participation in decision-making.

In Chapter 4, we describe the main trial of this thesis. In this study, 34 GPs trained in SDM followed by a positive reinforcement of the chosen therapy, included 110 patients consulting them for non-chronic low back pain. During the first 14 days and 2, 6, 12 and 26 weeks after the consultation for non-chronic low back pain, participants completed questionnaires on pain, restrictions and other outcome information. These outcomes were compared to 116 patients who visited an untrained GP who provided usual care.

Chapter 5 is a proof-of-concept study where SDM activities are associated with recovery, measured as a dichotomous outcome including both pain and restrictions. In this study, SDM performances are rated perceived from three different perspectives: the patient, the professional and an external observer.

Chapter 6 is a cost-effectiveness study of the intervention on costs 26 weeks after the consultation. Costs were estimated from both the healthcare and social perspectives because the resources are different and the bulk of the costs due to non-chronic low back pain are caused by sick leave.

In Chapter 7, the results of all the studies are summarised and reflected upon. Suggestions for future research and implementation of SDM are given related to modern developments on patient-professional communication.

REFERENCES

1) Lucassen P, Olesen F. Context as a drug: some consequences of placebo research for primary care. Scand J Prim Health Care 2016 December 01;34(4):428-433.
2) Evers AWM, Colloca L, Blease C, Annoni M, Atlas LY, Benedetti F, et al. Implications of Placebo and Nocebo Effects for Clinical Practice: Expert Consensus. Psychother Psychosom 2018;87(4):204-210.
3) Evers AWM, Colloca L, Blease C, Gaab J, Jensen KB, Atlas LY, et al. What Should Clinicians Tell Patients about Placebo and Nocebo Effects? Practical Considerations Based on Expert Consensus. Psychother Psychosom 2021;90(1):49-56.
4) Evers AWM, Colloca L, Blease C, Gaab J, Jensen KB, Atlas LY, et al. "Consensus on Placebo and Nocebo Effects Connects Science with Practice:" Reply to "Questioning the Consensus on Placebo and Nocebo Effects". Psychother Psychosom 2021;90(3):213-214.
5) Vase L, Wartolowska K. Pain, placebo, and test of treatment efficacy: a narrative review. Br J Anaesth 2019 August 01;123(2):e254-e262.
6) van Vliet LM, van Dulmen S, Mistiaen P, Bensing JM. The placebo effects of good communication. Ned Tijdschr Geneeskd 2016;160:D251.
7) Bensing JM, Verheul W. The silent healer: the role of communication in placebo effects. Patient Educ Couns 2010 Sep;80(3):293-299.
8) Atlas LY. A social affective neuroscience lens on placebo analgesia. Trends Cogn Sci 2021 November 01;25(11):992-1005.
9) Colagiuri B, Schenk LA, Kessler MD, Dorsey SG, Colloca L. The placebo effect: From concepts to genes. Neuroscience 2015 October 29;307:171-190.
10) Turner JA, LeResche L, Von Korff M, Ehrlich K. Back pain in primary care. Patient characteristics, content of initial visit, and short-term outcomes. Spine 1998 Feb 15;23(4):463-469.
11) Hunter T, Siess F, Colloca L. Socially induced placebo analgesia: a comparison of a pre-recorded versus live face-to-face observation. Eur J Pain 2014 August 01;18(7):914-922.
12) Dulmen van AM, Bensing JM. Contextwerking in de geneeskunde; een programmeringsstudie. Nivel (Nederlands instituut voor onderzoek van de gezondheidszorg) Postbus 1568, 3500 BN Utrecht 2000 August.
13) Friesen P. Placebos as a Source of Agency: Evidence and Implications. Front Psychiatry 2019 October 25;10:721.
14) Goossens ME, Vlaeyen JW, Hidding A, Kole-Snijders A, Evers SM. Treatment expectancy affects the outcome of cognitive-behavioral interventions in chronic pain. Clin J Pain 2005;21(1):18-72.
15) Peerdeman KJ, van Laarhoven AIM, Keij SM, Vase L, Rovers MM, Peters ML, et al. Relieving patients' pain with expectation interventions: a meta-analysis. Pain 2016 June 01;157(6):1179-1191.
16) van den Brink-Muinen A, van Dulmen AM, Jung HP, Bensing JM. Do our talks with patients meet their expectations? J Fam Pract 2007 Jul;56(7):559-568.
17) Kam-Hansen S, Jakubowski M, Kelley J, Kirsch I, Hoaglin D, Kaptchuk T, et al. Altered Placebo and Drug Labeling Changes the Outcome of Episodic Migraine Attacks. Science Translational Medicine 2014;6(218ra5).
18) Benedetti F, Amanzio M. Mechanisms of the placebo response. Pulm Pharmacol Ther 2013 Oct;26(5):520-523.
19) McKay KM, Imel ZE, Wampold BE. Psychiatrist effects in the psychopharmacological treatment of depression. J Affect Disord 2006 June 01;92(2-3):287-290.
20) Atlas LY, Wager TD. A meta-analysis of brain mechanisms of placebo analgesia: consistent findings and unanswered questions. Handb Exp Pharmacol 2014;225:37-69.

21) Amanzio M, Benedetti F. Neuropharmacological dissection of placebo analgesia: expectation-activated opioid systems versus conditioning-activated specific subsystems. J Neurosci 1999 January 01;19(1):484-494.
22) Smits RM, Veldhuijzen DS, van Middendorp H, van der Heijden, M J E, van Dijk M, Evers AWM. Integrating Placebo Effects in General Practice: A Cross-Sectional Survey to Investigate Perspectives From Health Care Professionals in the Netherlands. Front Psychiatry 2022 January 12;12:768135.
23) Houwen J, Lucassen PLBJ, Stappers HW, van Spaendonck K, van Duijnhoven A, Olde Hartman TC, et al. How to learn skilled communication in primary care MUS consultations: a focus group study. Scand J Prim Health Care 2021 March 01;39(1):101-110.
24) Mechanic D, Meyer S. Concepts of trust among patients with serious illness. Soc Sci Med 2000 Sep;51(5):657-668.
25) Stewart-Williams S, Podd J. The placebo effect: dissolving the expectancy versus conditioning debate. Psychol Bull 2004 Mar;130(2):324-340.
26) Verheul W, Sanders A, Bensing J. The effects of physicians' affect-oriented communication style and raising expectations on analogue patients' anxiety, affect and expectancies. Patient Educ Couns 2010 Sep;80(3):300-306.
27) Thomas KB. General practice consultations: is there any point in being positive? Br Med J (Clin Res Ed) 1987 May 9;294(6581):1200-1202.
28) Mistiaen P, van Osch M, van Vliet L, Howick J, Bishop FL, Di Blasi Z, et al. The effect of patient-practitioner communication on pain: a systematic review. Eur J Pain 2016 May;20(5):675-688.
29) Myers SS, Phillips RS, Davis RB, Cherkin DC, Legedza A, Kaptchuk TJ, et al. Patient expectations as predictors of outcome in patients with acute low back pain. J Gen Intern Med 2008 Feb;23(2):148-153.
30) Hallegraeff JM, Krijnen WP, van der Schans CP, de Greef MH. Expectations about recovery from acute non-specific low back pain predict absence from usual work due to chronic low back pain: a systematic review. J Physiother 2012;58(3):165-172.
31) Kapoor S, Shaw WS, Pransky G, Patterson W. Initial patient and clinician expectations of return to work after acute onset of work-related low back pain. J Occup Environ Med 2006 Nov;48(11):1173-1180.
32) Hayden JA, Wilson MN, Riley RD, Iles R, Pincus T, Ogilvie R. Individual recovery expectations and prognosis of outcomes in non-specific low back pain: prognostic factor review. Cochrane Database Syst Rev 2019 November 25;2019(11):10.1002/14651858.CD011284.pub2.
33) Couet N, Desroches S, Robitaille H, Vaillancourt H, Leblanc A, Turcotte S, et al. Assessments of the extent to which health-care providers involve patients in decision making: a systematic review of studies using the OPTION instrument. Health Expect 2013 Mar 4.
34) Stiggelbout AM, Van der Weijden T, De Wit MP, Frosch D, Legare F, Montori VM, et al. Shared decision making: really putting patients at the centre of healthcare. BMJ 2012 Jan 27;344:e256.
35) Legare F, Ratte S, Gravel K, Graham ID. Barriers and facilitators to implementing shared decision-making in clinical practice: update of a systematic review of health professionals' perceptions. Patient Educ Couns 2008 Dec;73(3):526-535.
36) Meijers MC, Noordman J, Spreeuwenberg P, Olde Hartman TC, van Dulmen S. Shared decision-making in general practice: an observational study comparing 2007 with 2015. Fam Pract 2019 May 23;36(3):357-364.
37) Charles C, Gafni A, Whelan T. Shared decision-making in the medical encounter: what does it mean? (or it takes at least two to tango). Soc Sci Med 1997 Mar;44(5):681-692.
38) Thompson AG. The meaning of patient involvement and participation in health care consultations: a taxonomy. Soc Sci Med 2007 Mar;64(6):1297-1310.

39) Shay LA, Lafata JE. Understanding patient perceptions of shared decision making. Patient Educ Couns 2014 September 01;96(3):295-301.
40) Elwyn G, Tsulukidze M, Edwards A, Legare F, Newcombe R. Using a 'talk' model of shared decision making to propose an observation-based measure: Observer OPTION 5 Item. Patient Educ Couns 2013 Nov;93(2):265-271.
41) Legare F, Politi MC, Drolet R, Desroches S, Stacey D, Bekker H, et al. Training health professionals in shared decision-making: an international environmental scan. Patient Educ Couns 2012 Aug;88(2):159-169.
42) Stacey D, Bennett CL, Barry MJ, Col NF, Eden KB, Holmes-Rovner M, et al. Decision aids for people facing health treatment or screening decisions. Cochrane Database Syst Rev 2011 Oct 5;10:CD001431.
43) Hack TF, Degner LF, Watson P, Sinha L. Do patients benefit from participating in medical decision making? Longitudinal follow-up of women with breast cancer. Psychooncology 2006 Jan;15(1):9-19.
44) Carlsen B, Aakvik A. Patient involvement in clinical decision making: the effect of GP attitude on patient satisfaction. Health Expect 2006 Jun;9(2):148-157.
45) Elwyn G, Edwards A, Mowle S, Wensing M, Wilkinson C, Kinnersley P, et al. Measuring the involvement of patients in shared decision-making: a systematic review of instruments. Patient Educ Couns 2001 Apr;43(1):5-22.
46) Clayman ML, Makoul G, Harper MM, Koby DG, Williams AR. Development of a shared decision making coding system for analysis of patient-healthcare provider encounters. Patient Educ Couns 2012 Sep;88(3):367-372.
47) Shay LA, Lafata JE. Where is the evidence? A systematic review of shared decision making and patient outcomes. Med Decis Making 2015 January 01;35(1):114-131.
48) Legare F, Stacey D, Turcotte S, Cossi MJ, Kryworuchko J, Graham ID, et al. Interventions for improving the adoption of shared decision making by healthcare professionals. Cochrane Database Syst Rev 2014 Sep 15;9:CD006732.
49) van der Weijden T, van der Kraan J, Brand PLP, van Veenendaal H, Drenthen T, Schoon Y, et al. Shared decision-making in the Netherlands: Progress is made, but not for all. Time to become inclusive to patients. Z Evid Fortbild Qual Gesundhwes 2022 June 01;171:98-104.
50) Meijers MC, Noordman J, Spreeuwenberg P, Olde Hartman TC, van Dulmen S. Shared decision-making in general practice: an observational study comparing 2007 with 2015. Fam Pract 2019 May 23;36(3):357-364.
51) Legare F, Adekpedjou R, Stacey D, Turcotte S, Kryworuchko J, Graham ID, et al. Interventions for increasing the use of shared decision making by healthcare professionals. Cochrane Database Syst Rev 2018 Jul 19;7:CD006732.
52) Bons SCS, Borg MAJP, Van den Donk M, Koes BW, Kuijpers T, Ostelo RWJG, et al. NHG-Standaard Aspecifieke lagerugpijn (Tweede herziening). 2017; Available at: https://www.nhg.org/standaarden/samenvatting/aspecifieke-lagerugpijn. Accessed 05/19, 2017.
53) van der Windt DA, Kuijpers T, Jellema P, van der Heijden GJ, Bouter LM. Do psychological factors predict outcome in both low-back pain and shoulder pain? Ann Rheum Dis 2007 Mar;66(3):313-319.
54) Lewis M, Morley S, van der Windt DA, Hay E, Jellema P, Dziedzic K, et al. Measuring practitioner/therapist effects in randomised trials of low back pain and neck pain interventions in primary care settings. Eur J Pain 2010 Nov;14(10):1033-1039.
55) Darlow B, Fullen BM, Dean S, Hurley DA, Baxter GD, Dowell A. The association between health care professional attitudes and beliefs and the attitudes and beliefs, clinical management, and outcomes of patients with low back pain: a systematic review. Eur J Pain 2012 January 01;16(1):3-17.

56) Farin E, Gramm L, Schmidt E. The patient-physician relationship in patients with chronic low back pain as a predictor of outcomes after rehabilitation. J Behav Med 2012 Apr 4.
57) Rosendal M, Carlsen AH, Rask MT. Symptoms as the main problem: a crosssectional study of patient experience in primary care. BMC Fam Pract 2016 March 10;17:29-8.
58) Bentsen BG. International Classification of Primary Care. Scandinavian Journal of Primary Health Care 1986;4(1):43-50.
59) Heins M, et al. Zorg door de huisarts; Nivel Zorgregistraties Eerste Lijn: jaarcijfers 2021 en trendcijfers 2017-2021. Nivel 2022 August,:3-181.
60) Maher C, Underwood M, Buchbinder R. Non-specific low back pain. Lancet (London, England) 2017;389:736-747.
61) Koes B, Van Tulder M. Acute low back pain. Am Fam Physician 2006 Sep 1;74(5):803-805.
62) Chavannes AW, Mens JMA, Koes BW, Lubbers WJ, Ostelo R, Spinnewijn WEM, et al. NHG-Standaard Aspecifieke lagerugpijn M54. 2005; Available at: http://www.springerlink.com.proxy.library.uu.nl/content/j12n60k507873536/, 2012.
63) Henschke N, Maher CG, Refshauge KM, Herbert RD, Cumming RG, Bleasel J, et al. Characteristics of patients with acute low back pain presenting to primary care in Australia. Clin J Pain 2009 January 01;25(1):5-11.
64) van Tulder MW, Koes BW, Bouter LM. A cost-of-illness study of back pain in The Netherlands. Pain 1995 Aug;62(2):233-240.
65) Schofield DJ, Callander EJ, Shrestha RN, Passey ME, Kelly SJ, Percival R. Back problems, comorbidities, and their association with wealth. Spine J 2015 January 01;15(1):34-41.
66) Maher C, Ferreira G. Time to reconsider what Global Burden of Disease studies really tell us about low back pain. Ann Rheum Dis 2022 March 01;81(3):306-308.
67) Murray CJ, Vos T, Lozano R, Naghavi M, Flaxman AD, Michaud C, et al. Disability-adjusted life years (DALYs) for 291 diseases and injuries in 21 regions, 1990-2010: a systematic analysis for the Global Burden of Disease Study 2010. Lancet 2012 Dec 15;380(9859):2197-2223.
68) Furlan AD, Imamura M, Dryden T, Irvin E. Massage for low back pain: an updated systematic review within the framework of the Cochrane Back Review Group. Spine (Phila Pa 1976) 2009 Jul 15;34(16):1669-1684.
69) Rubinstein SM, Terwee CB, Assendelft WJ, de Boer MR, van Tulder MW. Spinal manipulative therapy for acute low-back pain. Cochrane Database Syst.Rev. 2012;sept12(9):CD008880.pub2.
70) Chou R, Qaseem A, Snow V, Casey D, Cross JT,Jr, Shekelle P, et al. Diagnosis and treatment of low back pain: a joint clinical practice guideline from the American College of Physicians and the American Pain Society. Ann Intern Med 2007 Oct 2;147(7):478-491.
71) Roelofs PDDM, Deyo RA, Koes BW, Scholten RJPM, van Tulder MW. Non-steroidal anti-inflammatory drugs for low back pain. Cochrane Database of Systematic Reviews 2008(1):Art. No.: CD000396.
72) Pengel LH, Herbert RD, Maher CG, Refshauge KM. Acute low back pain: systematic review of its prognosis. BMJ 2003 Aug 9;327(7410):323.
73) Hall AM, Scurrey SR, Pike AE, Albury C, Richmond HL, Matthews J, et al. Physician-reported barriers to using evidence-based recommendations for low back pain in clinical practice: a systematic review and synthesis of qualitative studies using the Theoretical Domains Framework. Implement Sci 2019 May 07;14(1):49-4.
74) Flinterman, Linda, Schers H, Schermer T. Behandeling van klachten vanuit de lage rug in de eerste lijn. NIHES/Radboud University Medical Center 2019 “Sept, “:1-56.
75) van Lennep, J H P A, Trossel F, Perez RSGM, Otten RHJ, van Middendorp H, Evers AWM, et al. Placebo effects in low back pain: A systematic review and meta-analysis of the literature. Eur J Pain 2021 May 29.

76) Bishop A, Foster NE, Thomas E, Hay EM. How does the self-reported clinical management of patients with low back pain relate to the attitudes and beliefs of health care practitioners? A survey of UK general practitioners and physiotherapists. Pain 2008 Mar;135(1-2):187-195.
77) Lin CW, Haas M, Maher CG, Machado LA, van Tulder MW. Cost-effectiveness of general practice care for low back pain: a systematic review. Eur Spine J 2011 Jul;20(7):1012-1023.

CHAPTER 2

Effects of improved patient participation in primary care on health-related outcomes: a systematic review

Ariëtte R.J. Sanders
Inge van Weeghel
Maartje Vogelaar
William Verheul
Ron H.M. Pieters
Niek J. de Wit
Jozien M. Bensing

Family Practice. 2013 March 17;365–378.

ABSTRACT

Background. In primary care, many consultations address symptom-based complaints. Recovery from these complaints seldom exceeds placebo effects. Patient participation, because of its supposed effects on trust and patient expectancies, is assumed to benefit patients' recovery. While the idea is theoretically promising, it is still unclear what the effects of increased patient participation are on patient outcomes.

Aim. To review the effects of controlled intervention studies aiming to improve patient participation in face-to-face primary care consultations on patient-oriented and/or disease-oriented outcomes.

Methods. This study is a systematic review. A systematic search was undertaken for randomised controlled trials designed to measure the effects of interventions that aimed to improve adult patients' participation in primary care visits. The CINAHL, Cochrane, EMBASE, PsycINFO and PubMed databases were searched.

Results. Seven different trials fulfilled the inclusion criteria. Three of the studies were related to symptom-based complaints. Five studies measured patient-oriented outcomes, the primary outcome of interest for this review. All studies suffered from substantial bias. Studies varied widely in their aims, types of complaints/diseases, strength of the interventions and their outcomes. The effects on patient-oriented outcomes and disease-oriented outcomes were ambiguous.

Conclusion. Little research has been performed on health outcomes of interventions aiming to increase patient participation in general practice visits among patients suffering from symptom-based complaints. The results still are non-conclusive. The quality of the trials has been weak, possibly due to the complexity of the concept. This weak quality may explain the lack of conclusive results. Proposals for future research designs are offered.

INTRODUCTION

In primary care, patients consult their GP for complaints rather than for diseases. (1) The effectiveness of the therapies for these illnesses seldom exceeds placebo effects, but placebo effects alone can be substantial. (3,4,5) Usually these illnesses are evaluated by assessing patient-reported outcomes, which are highly dependent on patients' perspectives. (6-10)

In 1999, Crow concluded that the evidence justifies techniques that facilitate patient participation in consultations. (11) To encourage patient participation, providers need to recognise the patient's 'expert role' in matters such as the patient's preferences, concerns and self-efficacy. (12-14) When earlier positive patient experiences are incorporated into the treatment plan, this can subconsciously lead patients to beneficial outcomes. (15) In addition to the subconscious reaction, patients interpret this listening to their perspective as trust (16) Providing a caring, respectful and empowering context, thereby influencing the patient's affective state, seems to be effective in producing positive health outcomes. (10) Finally, incorporating patients' preferences in medical decision-making will lead to higher treatment expectations and higher adherence to therapy. (11,15) These expectations showed a positive effect on recovery although the literature shows some conflicting results. (17,18) The most recent Cochrane review on the effect of patient-centred approaches on health-related outcomes concluded that there was limited proof of beneficial effects. (19) The search was performed through 1999 using a restricted search strategy. Since then, shared decision-making (SDM), a technique in which patients are empowered to make healthcare choices jointly with the practitioner, has come to be considered crucial for patient-centred care. (13,20,21) However, previous reviews on SDM do not consider health-related outcomes as primary outcomes. (22,23)

Given the limitations of earlier reviews and their inconclusive results, we performed a systematic review on the effects of controlled interventions that aimed to improve patient participation in complaint-driven face-to-face primary care consultations on patient-oriented and/or disease-oriented outcomes.

METHODS

Search strategy

In March 2011, a pilot search was performed by MV in PubMed, using the search strategy developed by Légaré supplemented with terms from Lewin's work. (24,19) The pilot search was conducted backwards and forwards. This technique resulted in our definitive search strategy (see Appendix 1). A broad search strategy including the domain (providers, patients and provider–patient interactions) and determinant (promoting patient participation) was chosen to ensure that no applicable studies were missed. We used two filters (one to identify quasi-randomised and randomised

controlled trials and one to restrict the search to primary care). The PubMed search strategy is shown in Appendix 1.

The following electronic databases were searched on 7 October 2012: CINAHL, Cochrane, EMBASE through Embase.com, PsycINFO and PubMed (incorporating MEDLINE and Old MEDLINE). The searches were not restricted by language or by date. The included studies were forward and backward searched.

Inclusion and exclusion criteria

Types of studies. Randomised controlled intervention studies were included. We excluded controlled before-and-after studies, interrupted time series studies and all non-experimental studies. We did not exclude studies on the basis of allocation concealment or blinding.

Participants. The patients were over 18 years. We excluded studies of people with serious psychiatric symptoms, defined as patients requiring help from secondary care. The healthcare professionals were those responsible for patient care, including professionals in training.

Types of interventions. We included all patient-centred interventions aimed at affecting patients' ability to influence treatment decisions during primary care encounters. The interventions could occur before or during the clinical encounter.

Eligible interventions included educational meetings, audit and feedback (i.e. any summary of clinical healthcare performance over a specified period of time), reminders (i.e. information provided verbally, on paper or on a computer screen that prompts a professional to recall information), patient-mediated interventions (i.e. any intervention aimed at changing healthcare professionals' behaviour through interactions with or information provided by or to patients, which could include providing patients with information about the effectiveness and/or appropriateness of particular health technologies) and the distribution of printed educational material (i.e. published or printed recommendations about clinical care and evidence to improve practices, including clinical practice guidelines, journals and monographs). (24) Patient decision aids were considered patient-mediated interventions because one of their purposes is to foster patients' participation in decisions made during the clinical encounter. (25) Interventions conducted after the clinical encounter or studies that trained healthcare providers to achieve specific treatment or preventive goals (e.g. providers' adherence to guidelines or behavioural changes by the patients) that were initiated by the healthcare provider were excluded. Routine consultations for controlling chronic diseases were included.

Types of outcome measures

Primary outcomes. All patient-oriented outcomes, such as morbidity, mortality, symptoms, quality of life or personal costs were included. (26)

Secondary outcomes. All disease-oriented outcomes, such as histopathologic, physiologic or surrogate indicators (e.g. clinical assessments, body mass index (BMI), blood pressure or blood glucose) that may reflect changes in the disease course or health risks were included. (26)

We excluded studies that did not include any of the outcomes listed above or that measured only the patient's lifestyle behaviours; studies that measured only the provider's knowledge, attitudes or intentions; and studies that used simulated patients.

Data collection and analysis

Selection of trials. After excluding duplicates, two authors (IW and AS) screened the titles and abstracts of the articles obtained from the search and excluded studies according to the predetermined exclusion criteria (see Fig. 1). For studies that were questionable, one author (IW) scanned the full text before flagging the study for full text reading. The full text copies of all potentially relevant studies, except for one, were retrieved. The author of the study, which was not retrievable from several international libraries, was emailed but did not respond.

Eligible trials were screened for quality, the types of interventions they included and outcomes. The relevant data were extracted from the eligible studies by one author (AS), and the data from a sample of the studies (one-third) were independently confirmed by the second author (IW).

Quality assessment. All eligible trials were screened for the following characteristics: randomisation procedure, allocation concealment, blinded assessment of the outcome(s), intention-to-treat (ITT) analysis, differences at baseline between the groups of professionals and patients, missing data for providers and patients and protection against contamination. We also determined whether there was potential for error related to the unit of analysis and, if so, whether it was acknowledged and adjustments were made.

Disagreements were resolved through discussion between AS and IW; in cases of persistent disagreement, consensus was reached with the entire research team. To reduce quality assessment to one criterion, all of the articles were assessed by two independent researchers (AS and JvdZ) using the SORT criteria. (26) The SORT criteria for high-quality randomised controlled clinical trials (RCT) include concealed allocation, blinding, ITT analysis, adequate statistical power and adequate percentage of participants completing follow-up (greater than 80 per cent).

Data analysis. We planned to combine the studies with common outcomes when possible and to conduct a subgroup analysis for trials about complaint-based consultations versus routine check-ups for chronic diseases. We assumed that the effects would be lower in the latter group because preventive visits do have different contextual effects. (27) We estimated that there would be only moderate publication bias for the patient-oriented results because secondary patient-reported outcomes from studies that measured process outcomes as their primary outcomes were also included in the review.

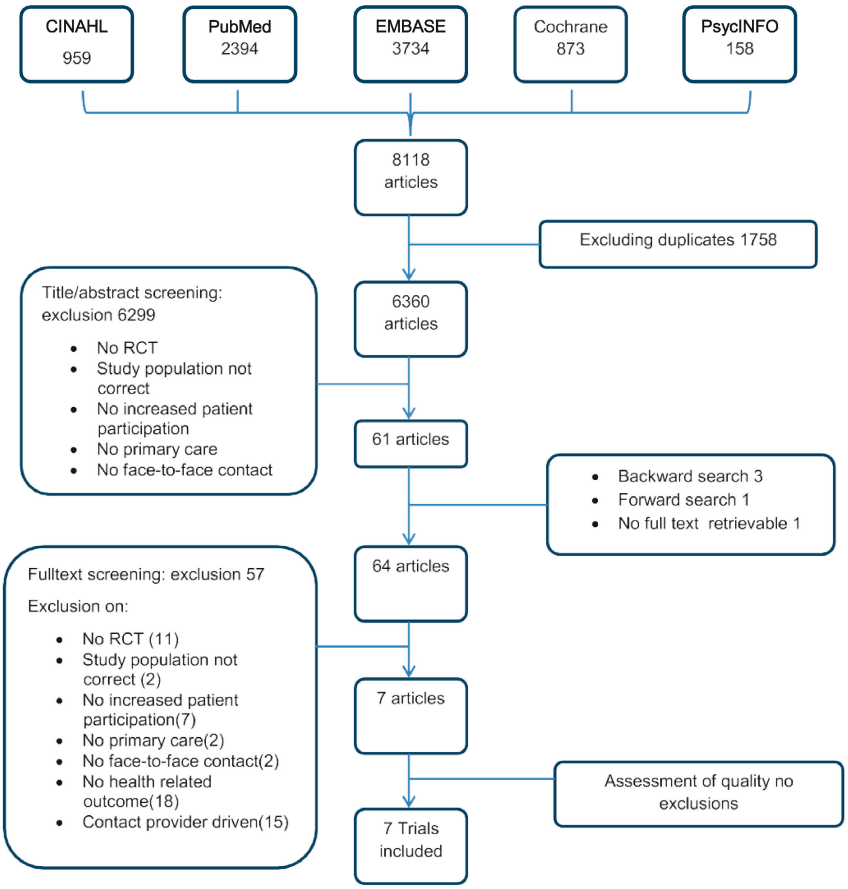


Figure 1. Flow chart

RESULTS

Study selection

The search identified a total of 8118 potentially relevant articles. After excluding duplicates, this number was reduced to a total of 6360 articles. The titles and abstracts were screened based on all of the exclusion criteria except for the outcomes and the purpose of contact. The correlation between both assessors was 0.645 (standard error = 0.032; 95% confidence interval = 0.582–0.708).

A total of 61 studies were flagged for full text screening. Three articles were added based on the backward search, and one was added from the forward search. Fifty-seven publications were excluded because they did not meet the inclusion criteria. Thirty-three of these were excluded based on the two criteria that were not applied during the screening of the titles and abstracts (i.e. they had no health-related outcomes or they included provider-driven contact; see Fig. 1). Ultimately, seven trials were included in this review.

Characteristics of the included studies

Two studies were conducted in England (28,29), two in Germany (30,31) one in the USA (32), one in Canada (33) and one in France (34). The data collection occurred in various time periods between 1993 and 2012. All of the studies were cluster randomised trials; the unit of randomisation was either the GP or the practice. The providers were predominantly GPs or primary care physicians (PCP). In four studies, PCP practice teams received the intervention training as a unit (28,29,32,33) The number of providers varied from 30 (31) to 162 (33) The number of included patients ranged from 165 (29) to 926 (30). The patients were seen for acute respiratory infections (33), osteoarthritis (OA) (34), depression (31) diabetes (28,29), hypertension (32) and cholesterol measurement (30). In three studies, patient complaints prompted contact with a provider (31,33,34); the other four studies involved routine visits to control chronic diseases. (28-30,32)

Only in one study was the primary aim of the intervention to use patient-centred methods to relieve patients' complaints. (34) All others, except for one (31), used patient participation as an instrument to address a disease-oriented measure, such as antibiotic use (33), A1C, lipid levels, etc. (28-30) Four of the studies combined this focus with process outcomes. (29,30,32,33) Loh focused on the beneficial effects on process outcomes. (31) Disease-oriented outcomes such as adherence to therapy were considered secondary outcomes. In all except one study, the control subjects received care as usual. (28) Table 1 summarises the characteristics of the included studies.

Table 1. Characteristics of the included studies

Author	Publication year	Country	Period of data collection	Unit of randomisation	Provider participants
Chassany (34)	2006	France	May 2001–April 2002	GP	GPs
Cooper (32)	2011	USA	January 2002–August 2005	Practice	General internists and family physicians
Kinmonth (28)	1998	England	April 1994–June 1995	Practice	PC practice teams[e]
Krones (30)	2008	Germany	May 2005–March 2006	CME-group	PC practice teams[e]
Légaré (33)	2012	Canada	July 2010–April 2011	Practice	Family physicians
Loh (31)	2007	Germany	October 2002–December 2004	PCP	PCPs[h]
Pill (29)	1998	England	April 1993–April 1996	Practice	PCPs[i]

IV, intervention; NIDDM, non-insulin-dependent diabetes mellitus; CME, continuous medical education.
[a] c: face-to-face contact is complaint-driven; r: routine visit for controlling chronic diseases.
[b] c: primary aim to relieve patient-oriented outcomes (complaints); d: disease-oriented outcomes (disease); p: process outcomes.
[c] 1: patient participation as an end in itself; 0: as an instrument to reach another goal.
[d] 1: controls care as usual; 2: other.
[e] PC practice teams consisted of doctors and nurses.

Number of providers (IV/controls)	Patient participants	Number patients (IV/ controls)	Contact complaint driven[a]	Primary aim[b]	Patient participation[c]	Controls[d]
180 (84/96)	Chronic complaints of osteoarthritis	842 (414/428)	c	c	1	1
41 (22/19)	Hypertensive patients in underserved PC	279 (83/57/84/55)	r	d + p	1	1
41 (21/20)[f]	Newly diagnosed NIDDM	250 (142/108)	r	d	0	2[g]
87 (44/47)	Cholesterol measurement	926 (460/466)	r	d + p	0	1
149 (77/72)	Acute respiratory infections	359 (181/178)	c	d + p	1	1
30 (20/10)	Newly diagnosed depressive disorder	405 (263/142)	c	p	0	1
29 general practices[j]	NIDDM	165 (77/88)	r	d + p	0	1

[f] 21 practices consisted of 23 doctors and 32 nurses; 20 practices consisted of 20 doctors and 32 nurses.
[g] The nurses in the comparison group were offered similar support sessions focusing on the use of guidelines and materials.
[h] PCP: primary care physicians, all teaching practices.
[i] PCPs and their practice nurses committed for at least 2 years to an annual peer reviewed clinical audit of diabetic care.
[j] Provider number, discrimination between IV versus controls not made.

Table 2 Quality assessment based on SORT-criteria (26)

Article	Randomisation[a]	Blinding providers[b]	Blinding patients[c]	ITT[d]	Adequate size	Missing data[e]	Level of evidence[f]
Chassany (34)	–	–	–	+	+	+	2
Cooper (32)	+	–	+	+	–	+	2
Kinmonth (28)	+	–	+	+	+	–	2
Krones (30)	+	–	+	+	+	–	2
Légaré (33)	+	–	+	+	+	–	2
Loh (31)	–	–	+	–	–	–	2
Pill (28)	+	–	–	–	+	–	2

[a] +: low risk of bias; –: high risk of bias or insufficient information to judge the risk.
[b] +: providers blinded; –: providers not blinded or insufficient information to judge provider blinding.
[c] +: patients blinded; –: patients not blinded or insufficient information to judge patient blinding.
[d] +: ITT analysis performed; –: ITT analysis not performed or insufficient information available to judge ITT analysis.
[e] +: missing data <20%; –: missing data> 20%.
[f] Quality level based on the SORT criteria. (26)

Quality assessment

All of the trials are *cluster* randomised controlled trials and therefore provided a moderate level of evidence according to the SORT criteria. (26) However, a considerable risk of bias hampered all of the studies. For two studies, the randomisation procedure could not be definitively determined from the article. (31,34) Information about patient blinding and outcome blinding could not be retrieved from three of the articles. Only one trial had a low risk of bias based on missing outcomes. (31) In the study by Loh, the follow-up time was short (2 weeks) and the GPs who did not include patients in the study were excluded from the analysis. (31) Thus, the internal validity of the overall results of this review is low. For all of the trials except the study by Légaré, recruitment bias was considered high. Consequently, the external validity must also be considered low. Table 2 summarises the quality assessment. Table 3 summarises the interventions.

Outcome measures

The outcome measures varied across different studies (see Table 4). All health-related outcomes were included. The patient-oriented outcomes varied from pain intensity (34) to perceived health (overall or disease-related) measured with questionnaires. (28,29,31,33,34) Most of the instruments were validated. The disease-oriented outcomes varied from biochemical results (28,29) to physical measurements (28,29,32) to risk scores. (30,34) The outcomes were measured between 2 weeks (33,34) and 1 year (32) after the intervention.

Effects

A meta-analysis on effect estimates could not be performed because of the wide variety of effects. There were both positive and negative effects for intervention groups, but only a few were significant (see Table 4).

Chassany found significant positive results for all outcome measures (pain, disability and global perceptions of osteoarthritis). Cooper found no effect on blood pressure control. Kinmonth found positive effects on all three primary outcomes (quality of life, depression and well-being) in newly diagnosed diabetic patients. The impact on well-being, however, was the only significant result. The effects on secondary outcomes were positive in some cases (A1C and total cholesterol) and negative in other cases (triglycerides, BMI and blood pressure). The impact on two indicators, BMI and triglycerides, was significantly negative; overall, there was a non-significant negative trend. Krones only measured secondary outcomes.

The mean change in cardiovascular risk for patients who had their cholesterol measured tended to be negative, albeit non-significant. Légaré found no effect on perceived health status. Loh found only a non-significant negative effect on depression as a primary review outcome. Pill found a significant negative effect on one out of eight primary outcomes (self-reported health status for diabetic patients); there was no effect on

the other seven measures. The effect on secondary outcomes (glyco-haemoglobin, BMI, blood pressure and cardiovascular complications) was positive but not significant.

In summary, three studies showed positive results for both groups. (28,30,31) Four studies suffered from recruitment bias, which led to a very good level of performance by all of the providers (28-31), or differences in the baseline characteristics of providers (31) or patients. (28,30) Of all five studies that measured primary outcomes, two found positive effects: one all significant (34) and one partly significant (28). One study found a negative result for one of eight included outcomes (29), and one study showed a non-significant negative result (31). The remaining study showed no effect. (33) The measurements in the significant positive study were taken over a short time span. (34)

Of the three studies with consultations based on patients' complaints (31,33,34), two of them (31,34) measured patient-related outcomes as primary outcomes. One showed a significant positive effect (34), and one found a non-significant negative effect. (31)

Subgroup analysis. Four studies were based on routine visits for controlling chronic diseases. (28-30,32) Two of them measured patient-oriented outcomes as primary outcomes. One study found a significant positive effect out of three positive effects (28), and one study found one significant negative effect out of eight outcomes. (29) For the secondary review outcomes, non-significant positive effects were observed in one study. (29) In addition to these effects, there were significant negative effects in Kinmonth's study and a non-significant negative effect in Krones' study. There were no effects observed in the study by Cooper (see Table 4).

Table 3. Summary of interventions

Author	Description of intervention and supportive material(s)	Training strategies	Concep tual basis[a]	Multi-faced intervention[b]
Chassany (34)	Pragmatic, interactive, centred on the patient–physician relationship and based on the specific biopsychosocial model of chronic pain; the training focused on 3 themes: the patient–physician relationship, analysis and evaluation of pain and prescribing and the negotiation of the therapeutic contract with the patient	Workshops, group discussion, reminders	1	1
Cooper (32)	Physician training focused on increasing patient engagement and pre-visit patient coaching to improve patient communication with clinicians and outcomes	Physicians: feedback on simulated patient contact and workbook or CD-ROM exercise; patients: coaching and telephone	1	1
Kinmonth (28)	Doctor training: didactic only; nurses training: didactic instruction and skills, including active listening and negotiation of behaviour change; for patients: a booklet that encouraged patients to ask questions and an optional leaflet for patients encouraging discussion of complications and concerns and a booklet for practitioners describing approaches to behaviour change	Doctors and nurses receive theory and nurses practiced skills and were supported by a facilitator	1	1

Trainers	Behaviour[c]	Duration of training[d]	Supportive material[e]	Strength of intervention[f]	How manipulation was measured[g]	Manipulation check result[h]
Facilitators and experts	1	240	0	2	–	–
Physicians: unknown; patients: by community health workers	1	>1480	1	2	Videotape and PQ, PQ on patient ratings of physicians' participatory decision-making style, PICS50	-/+
Experienced facilitator	1	270	1	2	PQ recall of supportive material and DQ on use of skills	+

Table 3. (Continued)

Author	Description of intervention and supportive material(s)	Training strategies	Concep tual basis[a]	Multi-faced intervention[b]
Krones (30)	Training for family doctors on SDM, script-like decision aid, booklet for doctors, individual summary sheet for the patient	CME groups in which family doctors were trained to moderate the training for the participants	1	1
Légaré (33)	Online self-tutorial and an interactive workshop addressing key components of the clinical decision-making process about antibiotic treatment for acute respiratory infections in PC	Videos, exercises, decision aid	1	1
Loh (31)	Multi-faceted intervention, physician training, decision aid and patient information leaflet	Theory, role play, discussions, modelling	1	1
Pill (29)	Training at surgery was tailored to the needs of the practice but at least 2 sessions of 3 hours with simple visual aids designed to assist the clinician in encouraging active patient participation, newsletters every 3–4 months	Discussion, demonstration of technology and often role play, continuing support by a research nurse, who mostly visited the practice nurses, 2 group meetings during the course of the study	1	1

PICS=patient perceived involvement in care scale.
[a] Conceptual basis—1: yes; 0: no;?: unclear
[b] Multi-faced intervention—1: yes; 0: no;?: unclear.
[c] Intended target of behaviour change—1: provider; 0: patient.
[d] Duration of training in minutes.
[e] Supportive material for patient—1: yes; 0: no.

Trainers	Behaviour[c]	Duration of training[d]	Supportive material[e]	Strength of intervention[f]	How manipulation was measured[g]	Manipulation check result[h]
Moderators of CME group and members of the research team	1	240	1	2	PQ (50-53)	+
Facilitators trained during the pilot trial	1	240	1	2	PQ: the modified Control Preference Scale(54-56) and a single question with a Likert scale to assess the quality of the decision made	+
Unclear	1	Unknown	1	2	PQ (PICS)(50) results	+
Intervention team (GP, research nurse and clinical psychologist)	1	>360	0	3	Audiotape and telephone interview	+/−

[f] Strength of the intervention—1: weak (1 session, 1 day, teaching didactics); 2: intermediate (all other interventions with training sessions between sessions); 3: strong (3≥ sessions, >1 day, opportunity to practice skills between sessions and at least one of next 3 items: follow-up support, additional materials or a supportive tool); 0: no trained intervention.
[g] How manipulation was measured; 0: no manipulation check; PQ: patient questionnaire after the encounter; DQ: provider questionnaire after the encounter.
[h] Manipulation check results—0= no manipulation check;? = unclear because the manipulation check failed; + = increased patient participation in the intervention group; − = less patient participation in the intervention group compared with the control group.

Table 4. Summary of patient-oriented outcomes (primary review outcome) and disease-oriented outcomes (secondary review outcome)

Author	Outcome	Instrument	Study duration in weeks	Outcome[a]	Significance[b]	Adverse effects[c]	Summary of results
Chassany (34)	Pain intensity with motion (change from baseline over 2 weeks)	VAS-scale expressed as the sum of pain intensity differences	2	+	<0.00001	None	Overall significant positive effect on the primary review outcome measurement point after 2 weeks only
	Pain intensity with motion by VAS (change from baseline to study end)	VAS-scale	2	+	0.01		
	Functional disability	WOMAC index(58)	2	+	<0.0001		
	Global perception	A 7-point Likert scale	2	+	0.002		
Cooper (32)	*Blood pressure change, systolic and diastolic*	*Automatic oscillometric monitor (Omron HEM 907)*	*52*	*0*	*ns*	*None*	No primary review outcome; *no effect on secondary review outcomes*; positive effect on process outcomes

Kinmont (28)	Quality of life	ADDQoL(57)	52	+	ns	Knowledge score below 0.03	Of 3 primary review outcomes one showed a significant positive effect, 2 were positive but non-significant. *Secondary review outcomes: all patients reached good clinical A1C levels. 2 out of 6 showed significant negative effects. Combined, all other clinical measures (i.e., cardiovascular risk factors) showed a negative trend.* Both groups showed improved performances on process and health outcomes
	Depressed well-being	Depressed well-being questionnaire(57)	52	+	ns		
	Well-being	Bradley well-being questionnaire(57) d	52	+	0.03		
	A1C	*Percentage of glycated haemoglobin*(30)	*52*	*+*	*ns*		
	Total cholesterol	*Cholesterol oxidase concentration*	*52*	*+*	*ns*		
	Triglycerides	*Plasma triglyceride concentration*	*52*	*–*	*0.02*		
	Body mass index (kg/m²)	*Harpenden pocket stadiometer and Seca 835 electronic scales*	52	–	0.03		
	Blood pressure, systolic and diastolic	*Omron electronic*	*52*	–	*ns*		
	Urinary albumin/ creatinine ratio	*Jaffé reaction*	*52*	*+*	*ns*		

Table 4. Continued

Author	Outcome	Instrument	Study duration in weeks	Outcome[a]	Significance[b]	Adverse effects[c]	Summary of results
Krones (30)	*Mean change in* cardiovascular disease *risk (%)*	*Framingham scoring system calibrated for European populations*	*26*	–	*ns*	None	No primary review outcome. *Secondary outcome: no significant difference between groups*. Cardiovasculair risk status decreased in both groups
Légaré (33)	Self-reported healthstatus	Short form SF-12 questionnaire (59)	2	0	*ns*	None	No effect on primary review outcome. There was a positive effect on process outcomes
Loh (31)	Decline in depression severity (PHQ baseline-PHQ treatment/PHQ baseline x 100%)	Measured by brief PHQ-D	6-8	-	-0.078	None	Primary review outcome: negative non-significant decrease in depression, but controls had a worse starting point. Depression scores decreased in both groups

Pill (29)	Self-reported health status	Short form SF-36 questionnaire (60)	39	-/0	1 of 8 was significant -0.02	None	Primary review outcome: 1 out of 8 showed a significant negative effect. Others showed no effect. *Positive effect on secondary review outcomes (not significant).* Low self-reported ability to maintain behaviour over time
	Glyco-Hb	*Lab test*	*39*	*+*	*n.s.*		
	Body mass index (kg/m²)	*Routinely collected audit data*	*39*	*+*	*n.s.*		
	Blood pressure,systolic and diastolic	*Routinely collected audit data*	*39*	*+*	*n.s.*		
	Other clinical complications	*Routinely collected audit data*	*39*	*0*	*n.s.*		

PHQ= the short form of the patient health questionnaire.
Note: gray shade = consultations based on patients' complaints. Secondary review outcomes are in italics.
VAS-scale, visual analogue scale; WOMAC, Western Ontario and McMaster Universities osteoarthritis index.
[a] + = in favour of the intervention group; − = in favour of the control group; 0 = no difference.
[b] Significance expressed in *P*-values, if possible.
[c] Adverse effect = effects not mentioned as primary or secondary review outcomes but (potentially) harmful other effects (i.e. patient withdrawal from care), expressed in *P*-values.
[d] Diabetes specific measures of well-being and satisfaction with treatment.
[e] Clinical complications, 1 point for each complication: amputation, foot ulcer, cardiovascular accident, ischemic heart disease, retinopathy and neuropathy.

DISCUSSION

Main findings

Despite an elaborate search strategy, we found only seven controlled intervention studies that related to our study aim despite the growing emphasis on patient participation in the literature. (10,15,36) The seven studies in this review show ambiguous results. Despite the underlying theory, we see no significant effect (a suggestion of a positive impact at most) of patient participation on patient-related outcomes. This is similar to or even weaker than the results of other reviews. (10,19,24,37) For disease-related outcomes, no overall effect of patient participation can be demonstrated; some studies even revealed deterioration in disease-oriented outcomes. Moreover, in several studies, the control patients also improved sometimes even more than the patients in the experimental condition. These results deserve some reflection.

First, why did we find so few studies in this area even though few people will deny the importance of the subject? In the literature, many different terms are used for patient participation. We may have missed certain terms and consequently missed trials. Backward and forward searches and a review of the reference lists of the selected publications resulted in only three new publications. Thus, this will be not the main reason.

A more convincing reason might be that there simply are not many controlled intervention studies on the effect of patient participation on health outcomes. The concepts used can be considered 'fuzzy concepts' (19,38): everyone understands what is meant by the concept generally, but there is no precise definition, which hampers the operationalisation of the term. (19,38) Therefore, researchers may shy away from choosing this topic for an RCT on health outcomes rather than process outcomes. Many other topics are easier to research.

The yield of this review suggest that researchers show more interest in disease-specific goals such as adherence, (31) antibiotic use or lowering risk factors (28,30), using patient participation as an instrument. Training physicians hampers blinding them. Yet, well-performed, state-of-the-art empirical research is needed to discriminate between sense and nonsense of specific approaches to improve patient participation.

We must also conclude that the results are contradictory. The lack of actual behavioural differences between the experimental groups and the control groups may have contributed to the non-significant results. Informing participating GPs about the aim of the study might affect communication in both groups. In addition, conflicting motives (i.e. increasing patient participation versus changing the patients' unhealthy behaviours) might have counteracted the providers' performance of the skills they had been taught. Changing provider behaviour, especially related to patient participation, is not easily achieved. (19,24,39,40-42) All 4 trials (28,30-32) that measured a positive intervention effect used questionnaires, but the results of video-recorded or audiotaped encounters were ambiguous. (29,32)

Patient-related elements, such as patients' former failed attempts to change harmful behaviours, can lead to negative effects on health-related outcomes such as blood pressure. (11,43)

Learning from the pitfalls mentioned above, we offer several recommendations for future research in this field.

We recommend choosing an illness rather than a disease and measuring patient-oriented outcomes rather than disease-oriented outcomes. To design a properly blinded RCT, one could consider training participants in task-oriented behaviour (i.e. guideline adherence) instead of providing care as usual. The patients could be recruited in proximity to each training session, but there should be sufficient time between trainings to prevent problematic carry-over effects. To measure training effects, one could use baseline measurements before both patient recruitment phases. We assume carry-over effects are lower than generally accepted because the persistence of trained behaviours is repetitively lower than expected. (6,29,39,44) To prevent the influence of affective behaviour, providers should not be informed about the aims of health-oriented trials, only about process goals. This can best be accomplished during the recruitment of professionals because a particular interest in the topic affects associated outcomes. It is generally recommended to assess the feasibility of the trial with a pilot test, but it seems wise to pay extra attention to the behavioural effects of the trainings and observe subsequent behaviour for changes. (45,47)

LIMITATIONS OF THE STUDY

We only found one trial (32) in which an intervention (decision aid) was performed prior to consulting a physician. This might be due to logistic problems implementing this type of design in primary care. (48) However, if an intervention preceding a clinical encounter was not described in the abstract, we may have missed it.

For the sake of time, we did not contact researchers in the field. Instead, we based our search on well-established search strategies with proven merit. We piloted the search strategy in PubMed. When leading articles were not identified, we broadened the strategy using terms that cover these non-identified articles. Many studies in this field focus on process outcomes and do not mention health-related outcomes in the abstract. Therefore, we did not exclude outcomes during the abstract screening.

Publication bias may have influenced the results, but we believe that this risk was low because studies that had secondary outcomes relevant to the review were included. There were no restrictions on time or language.

Authors were not contacted, which may have led to an underestimation of trial quality. (49) Overall, more than 570 providers and 3244 patients were included, and the primary outcome was measured in 1824 patients.

CONCLUSION

The trials were heterogeneous in their populations, interventions and measures. The theoretical concept that patient participation has a beneficial effect on patient-reported health outcomes has still not been proven. The trials that concentrated on relieving patient complaints were scarce and suffered from quality issues. The results of studies on disease-oriented health outcomes were ambivalent. However, including patients in trials aimed at improving patient participation tends to benefit all patients. Research on the effects of improved patient participation should devote extra attention to developing unbiased designs and invest more in changing the affective behaviour of providers.

ACKNOWLEDGEMENTS

The authors wish to thank Jouke van der Zee for critically appraising all included articles.

Ethical approval: none. Conflict of interest: none.

REFERENCES

1) Nivel, Netherlands institute for health services and research. Incidentie- en prevalentiecijfers in de huisartsenpraktijk; incidentie alle symptomen en klachten ICPC (R01-R29). *Afkomstig uit LINH*. 2012. https://www.nivel.nl/nl/nivel-zorgregistraties-eerste-lijn (accessed on 10 march 2023).
2) Biondi DM. Physical treatments for headache: a structured review.*Headache* 2005; **45**: 738–46.
3) Montgomery P, Dennis J. Cognitive behavioural interventions for sleep problems in adults aged 60+. *Cochrane Database Syst Rev* 2003; **1**: CD003161.
4) Di Blasi Z, Harkness E, Ernst E, Georgiou A, Kleijnen J. Influence of context effects on health outcomes: a systematic review. *Lancet* 2001; **357**: 757–62.
5) Enck P, Horing B, Weimer K, Klosterhalfen S. Placebo responses and placebo effects in functional bowel disorders. *Eur J Gastroenterol Hepatol* 2012; **24**: 1–8.
6) Jellema P, van der Windt DA, van der Horst HE, Twisk JW, Stalman WA, Bouter LM. Should treatment of (sub)acute low back pain be aimed at psychosocial prognostic factors? Cluster randomised clinical trial in general practice. *BMJ* 2005; **331**: 84.
7) van der Windt DA, Kuijpers T, Jellema P, van der Heijden GJ, Bouter LM. Do psychological factors predict outcome in both low-back pain and shoulder pain?*Ann Rheum Dis* 2007; **66**: 313–9.
8) Khan AA, Khan A, Harezlak J, Tu W, Kroenke K. Somatic symptoms in primary care: etiology and outcome. *Psychosomatics* 2003; **44**: 471–8.
9) Ursin H, Eriksen H. Cognitive activation theory of stress, sensitization, and common health complaints. *Ann N Y Acad Sci* 2007; **1113**: 304-10.
10) Stewart MA. Effective physician-patient communication and health outcomes: a review. *CMAJ* 1995; **152**: 1423–33.
11) Crow R, Gage H, Hampson S, Hart J, Kimber A, Thomas H. The role of expectancies in the placebo effect and their use in the delivery of health care: a systematic review. *Health Technol Assess* 1999; **3**: 1–96.
12) Ong LM, de Haes JC, Hoos AM, Lammes FB. Doctor-patient communication: a review of the literature. *Soc Sci Med* 1995; **40**: 903–18.
13) Towle A, Godolphin W. Framework for teaching and learning informed shared decision making. *BMJ* 1999; **319**: 766–71.
14) Elwyn G, Hutchings H, Edwards A *et al.* The OPTION scale: measuring the extent that clinicians involve patients in decision-making tasks. *Health Expect* 2005; **8**: 34–42.
15) Bensing JM, Verheul W. The silent healer: the role of communication in placebo effects. *Patient Educ Couns* 2010; **80**: 293–9.
16) Mechanic D, Meyer S. Concepts of trust among patients with serious illness. *Soc Sci Med* 2000; **51**: 657–68.
17) Thomas KB. General practice consultations: is there any point in being positive? *Br Med J* 1987; **294**: 1200–2.
18) Knipschild P, Arntz A. Pain patients in a randomized trial did not show a significant effect of a positive consultation. *J Clin Epidemiol* 2005; **58**: 708–13.
19) Lewin SA, Skea ZC, Entwistle V, Zwarenstein M, Dick J. Interventions for providers to promote a patient-centred approach in clinical consultations. *Cochrane Database Syst Rev* 2001; **4**: CD003267.
20) Rockenbauch K, Schildmann J. Shared decision making (SDM): a systematic survey of terminology use and concepts. *Gesundheitswesen* 2011; **73**: 39408.

21) Weston WW. Informed and shared decision-making: the crux of patient-centered care. *CMAJ* 2001; **165**: 438–9.
22) Joosten EA, DeFuentes-Merillas L, de Weert GH, Sensky T, van der Staak CP, de Jong CA. Systematic review of the effects of shared decision-making on patient satisfaction, treatment adherence and health status. *Psychother Psychosom* 2008; **77**: 219–26.
23) Légaré F, Turcotte S, Stacey D, Ratté S, Kryworuchko J, Graham ID. Patients' perceptions of sharing in decisions: a systematic review of interventions to enhance shared decision making in routine clinical practice. *Patient* 2012; **5**: 1–19.
24) Légaré F, Bekker H, Desroches S *et al.* Effective continuing professional development for translating shared decision making in primary care: A study protocol. *Implement Sci* 2010; **5**: 83.
25) O'Connor AM, Bennett CL, Stacey D *et al.* Decision aids for people facing health treatment or screening decisions. *Cochrane Database Syst Rev* 2009; **3**: CD001431.
26) Ebell MH, Siwek J, Weiss BD *et al.* Simplifying the language of evidence to improve patient care: Strength of recommendation taxonomy (SORT): a patient-centered approach to grading evidence in medical literature. *J Fam Pract* 2004; **53**: 111–20.
27) Brownson RC, Haire-Joshu D, Luke DA. Shaping the context of health: a review of environmental and policy approaches in the prevention of chronic diseases. *Annu Rev Public Health* 2006; **27**: 341–70.
28) Kinmonth AL, Woodcock A, Griffin S, Spiegal N, Campbell MJ. Randomised controlled trial of patient centred care of diabetes in general practice: impact on current wellbeing and future disease risk. The Diabetes Care From Diagnosis Research Team. *BMJ* 1998; **317**: 1202–8.
29) Pill R, Stott NC, Rollnick SR, Rees M. A randomized controlled trial of an intervention designed to improve the care given in general practice to Type II diabetic patients: patient outcomes and professional ability to change behaviour. *Fam Pract* 1998; **15**: 229–35.
30) Krones T, Keller H, Sönnichsen A *et al.* Absolute cardiovascular disease risk and shared decision making in primary care: a randomized controlled trial. *Ann Fam Med* 2008; **6**: 218–27.
31) Loh A, Simon D, Wills CE, Kriston L, Niebling W, Härter M. The effects of a shared decision-making intervention in primary care of depression: a cluster-randomized controlled trial. *Patient Educ Couns* 2007; **67**: 324–32.
32) Cooper LA, Roter DL, Carson KA *et al.* A randomized trial to improve patient-centered care and hypertension control in underserved primary care patients. *J Gen Intern Med* 2011; **26**: 1297–304.
33) Légaré F, Labrecque M, Cauchon M, Castel J, Turcotte S, Grimshaw J. Training family physicians in shared decision-making to reduce the overuse of antibiotics in acute respiratory infections: a cluster randomized trial. *CMAJ* 2012; **184**: E726–34.
34) Chassany O, Boureau F, Liard F *et al.* Effects of training on general practitioners' management of pain in osteoarthritis: a randomized multicenter study. *J Rheumatol* 2006; **33**: 1827–34.
35) Kinmonth AL, Spiegal N, Woodcock A. Developing a training programme in patient-centred consulting for evaluation in a randomised controlled trial; diabetes care from diagnosis in British primary care. *Patient Educ Couns* 1996; **29**: 75–86.
36) Salzburg Global Seminar. Salzburg statement on shared decision making. *BMJ* 2011; **342**: d1745.
37) Griffin SJ, Kinmonth AL, Veltman MW, Gillard S, Grant J, Stewart M. Effect on health-related outcomes of interventions to alter the interaction between patients and practitioners: a systematic review of trials. *Ann Fam Med* 2004; **2**: 595–608.
38) Entwistle VA, Watt IS. Patient involvement in treatment decision-making: the case for a broader conceptual framework. *Patient Educ Couns* 2006; **63**: 268–78.

39) Edwards A, Elwyn G. Involving patients in decision making and communicating risk: a longitudinal evaluation of doctors' attitudes and confidence during a randomized trial. *J Eval Clin Pract* 2004; **10**: 431–7.
40) Oxman AD, Thomson MA, Davis DA, Haynes RB. No magic bullets: a systematic review of 102 trials of interventions to improve professional practice. *CMAJ* 1995; **153**: 1423–31.
41) Lee RG, Garvin T. Moving from information transfer to information exchange in health and health care. *Soc Sci Med* 2003; **56**: 449–64.
42) Gysels M, Richardson A, Higginson IJ. Communication training for health professionals who care for patients with cancer: a systematic review of training methods. *Support Care Cancer* 2005; **13**: 356–66.
43) Stewart-Williams S, Podd J. The placebo effect: dissolving the expectancy versus conditioning debate. *Psychol Bull* 2004; **130**: 324–40.
44) Elwyn G, Edwards A, Hood K *et al.* Achieving involvement: process outcomes from a cluster randomized trial of shared decision making skill development and use of risk communication aids in general practice. *Fam Pract* 2004; **21**: 337–46.
45) Campbell M, Fitzpatrick R, Haines A *et al.* Framework for design and evaluation of complex interventions to improve health. *BMJ* 2000; **321**: 694–6.
46) Caldwell PH, Hamilton S, Tan A, Craig JC. Strategies for increasing recruitment to randomised controlled trials: systematic review. *PLoS Med* 2010; **7**: e1000368.
47) Treweek S, Pitkethly M, Cook J *et al.* Strategies to improve recruitment to randomised controlled trials. *Cochrane Database Syst Rev* 2010; **4**: MR000013.
48) van der Wouden JC, Blankenstein AH, Huibers MJ, van der Windt DA, Stalman WA, Verhagen AP. Survey among 78 studies showed that Lasagna's law holds in Dutch primary care research. *J Clin Epidemiol* 2007; **60**: 819–24.
49) Soares HP, Daniels S, Kumar A *et al.* Bad reporting does not mean bad methods for randomised trials: observational study of randomised controlled trials performed by the Radiation Therapy Oncology Group. *BMJ* 2004; **328**: 22–4.
50) Lerman CE, Brody DS, Caputo GC, Smith DG, Lazaro CG, Wolfson HG. Patients' Perceived Involvement in Care Scale: relationship to attitudes about illness and medical care. *J Gen Intern Med* 1990; **5**: 29–33.
51) Man-Son-Hing M, Laupacis A, O'Connor AM *et al.* A patient decision aid regarding antithrombotic therapy for stroke prevention in atrial fibrillation: a randomized controlled trial. *JAMA* 1999; **282**: 737–43.
52) Krones T, Keller H, Sönnichsen AC, Sadowski EM, Baum E, Donner-Banzhoff N. Partizipative Entscheidungsfindung in der Cardiovasculären Risicoprävention: Ergebnisse der Pilotstudie von ARRIBA-Hertz, einer consultationbezogenen Entscheidungshilfe für die allgemeinmedizinische Praxis (Shared decision-making in cardiovasculair prevention: results from the pilotting of ARRIBA-Herz, a transactional decision aid for general practice). *Z Med Psychol* 2006; **15**: 61–70.
53) Loh A, Leonhart R, Wills CE, Simon D, Härter M. The impact of patient participation on adherence and clinical outcome in primary care of depression. *Patient Educ Couns* 2007; **65**: 69–78.
54) Strull WM, Lo B, Charles G. Do patients want to participate in medical decision making? *JAMA* 1984; **252**: 2990–4.
55) Melbourne E, Roberts S, Durand MA, Newcombe R, Légaré F, Elwyn G. Dyadic OPTION: Measuring perceptions of shared decision-making in practice. *Patient Educ Couns* 2011; **83**: 55–7.
56) Melbourne E, Sinclair K, Durand MA, Légaré F, Elwyn G. Developing a dyadic OPTION scale to measure perceptions of shared decision making. *Patient Educ Couns* 2010; **78**: 177–83.

57) Bradley C (ed.). *Handbook of Psychology and Diabetes: A Guide to Psychological Measurement in Diabetes Research and Management*. Chur, Zwitserland: Harwood Academic, 1994.
58) López Alonso SR, Martínez Sánchez CM, Romero Cañadillas AB, Navarro Casado F, González Rojo J. Metric properties of WOMAC questionnaires-original and reduced versions-to measure symptoms and Physical Functional Disability. *Aten Primaria* 2009; **41**: 613–20.
59) Ware J Jr, Kosinski M, Keller SD. A 12-Item Short-Form Health Survey: construction of scales and preliminary tests of reliability and validity. *Med Care* 1996; **34**: 220–33.
60) Ware JE Jr, Sherbourne CD. The MOS 36-item short-form health survey (SF-36). I. Conceptual framework and item selection. *Med Care* 1992; **30**: 473–83.

APPENDIX 1 SEARCH STRATEGY

No. Synonyms

#1 Domain: providers, patients and provider–patient interactions, patient OR patient[MeSH] OR subject OR subjects OR participant* OR client* OR inpatient* OR outpatient* OR hospitalized* OR
institutionalized* OR institutionalised* OR survivor* OR men OR woman OR women OR man OR consumer* OR people

#2 health personnel[MeSH] OR doctor* OR physician* OR provider* OR practitioner* OR gp OR gps OR health-professional* OR nurse* OR carer* OR caregiver* OR clinician* OR health-care-professional OR health-care-professionals OR healthcare professional* OR health-careworker* OR healthcare-worker* OR hospitalist* OR resident* NOT (veterinarian*)

#3 communication[MeSH] OR interact* OR communicat* OR relation* OR instruct* OR verbal* OR nonverbal OR smiling OR "facial expression" OR advis* OR talk* OR contact* OR conversation* OR consult OR consultation

#4 Professional-Patient Relations[MeSH] OR Physician's Role[MeSH] OR nurse's role[MeSH] OR "Professional Patient" OR "patient professional" OR "therapeutic alliance" OR doctor-patient OR patient-doctor OR clinician-patient OR patient-clinician OR physician-patient OR patient-physician OR nurse-patient OR patient-nurse OR patient-practitioner OR practitioner-patient OR biopsychosocial*

#5 (#1 AND #2 AND #3) OR #4
Determinant: promoting patient participation

#6 decision making[MeSH] OR decision support techniques[MeSH:noexp] OR decision support systems, clinical[MeSH] OR choice behaviour[MeSH] OR choice behavior[MeSH] OR decision making[tiab] OR decision counselling[tiab] OR decision support[tiab] OR choice-behaviour*[tiab] OR choice-behavior*[tiab] OR ((decision*[ti] OR choice*[ti]) AND (making*[ti] OR support*[ti] OR behaviour*[tiab] OR behavior*[tiab])) OR shared-decision[tiab] OR sharing-decision*[tiab] OR informed-decision*[tiab] OR informed-choice*[tiab] OR treatment-choice*[tiab] OR decision-autonomy[tiab] OR decisional-autonomy[tiab]

#7 consumer participation[MeSH] OR patient participation[MeSH] OR patient-participation*[tiab] OR consumer-participation*[tiab] OR patient involvement*[tiab] OR consumer-involvement[tiab] OR ((patient*[ti] OR consumer*[ti]) AND (involvement*[ti] OR involving*[ti] OR participation*[ti] OR participating*[ti])) OR patient-centered[tiab]
OR patient-centred[tiab] OR patient-oriented[tiab] OR patient-focused[tiab] OR client-focused[tiab] OR client-oriented[tiab] OR patient preference[MeSH] OR patient-centered care[MeSH] OR patient preference[tiab] OR patient-centered care[tiab]

#8 decision-aid*[tiab] OR consultation-leaflet[tiab] OR patient education handout[tiab] OR patient education as topic[MeSH Terms] OR "patient education"[All Fields] OR decision trees[MeSH] OR decision-support[tiab] OR audiotape*[tiab] OR brochure[tiab] OR booklet[tiab] OR flyer[tiab] OR folder[tiab] OR handout[tiab] OR leaflet[tiab] OR pamphlet[tiab] OR guide[tiab]

#9 #6 OR #7 OR #8
RCT-filter

#10 (randomized-controlled-trial OR controlled-clinical-trial OR randomized[tiab] OR randomised[tiab] OR placebo[tiab] OR drug-therapy[sh] OR randomly[tiab] OR trial[tiab] OR groups[tiab]) NOT (animals[mh] NOT (animals[mh] AND humans[mh]))
General practice

#11 primary health care[MeSH] OR physicians, primary care[MeSH] OR primary care nursing[MeSH] OR general practice[MeSH] OR general practitioners[MeSH] OR primary-care[tiab] OR general-
practice[tiab] OR family-practice[tiab] OR primary-health-care[tiab] OR primary-care-physician*[tiab]
Domain AND Determinant AND RCT AND General Practice

#12 #5 AND #9 AND #10 AND #11

CHAPTER 3

Does training general practitioners result in more shared decision-making during consultations?

Ariëtte R.J. Sanders
Jozien M. Bensing
Margan A.L.U. Essed,
Tessa Magnée,
Niek J. de Wit,
Peter F.M Verhaak

Patient Educ Couns. 2017 Mar; 563-574.

ABSTRACT

Objective. We conducted a clustered randomised controlled trial to study the effects of shared decision-making (SDM) on patient recovery. This study aims to determine whether GPs trained in SDM and reinforcing patients' treatment expectations showed more trained behaviour during their consultations than untrained GPs.

Method. We compared 86 consultations conducted by 23 trained GPs with 89 consultations completed by 19 untrained GPs. The primary outcomes were SDM, as measured by the OPTION scale, and positive reinforcement, as measured by global observation. Secondary outcomes were the level of autonomy in decision making and the duration of the consultation.

Results. Intervention consultations scored significantly higher on most elements of the OPTION scale, and on the autonomy scale; however, they were three minutes longer in duration, and the mean OPTION score of the intervention group remained below average.

Conclusion. Training GPs resulted in more SDM behaviour and more autonomy for the patient; however, this increase is not attributable to the adoption of a patient perspective. Furthermore, while we aimed to demonstrate that SDM facilitates the reinforcement of patients' positive expectations, the measurement of this behaviour was not reliable.

Practice implications. In supporting SDM, professionals should give greater attention to *patients*' treatment expectations.

INTRODUCTION

In medical decisions, little attention is devoted to the patient perspective, and patients' expectations often remain unnoticed. (1–3) This may have negative implications for recovery. (4). The concept of shared decision-making (SDM), i.e., both the patient and the professional participate in the decision-making process and come to joint conclusions, is considered crucial for empowering patients to manage their healthcare problems and for overcoming this deficiency. (5)

SDM may reinforce patients' pre-existing ideas about recovery in treatment choices, and recovery may be facilitated if they have positive expectations. (4,6) Thus, health professionals can contribute to better health outcomes by positively reinforcing patients' recovery expectations through discussions of the benign spontaneous course. (7) Furthermore, health professionals can use a therapeutic approach to positively reinforce patients' pre-existing positive ideas about recovery.

The aim of SDM is to increase patients' autonomy in decisions about their personal health by shifting the doctor-patient relationship from a paternalistic to a more equal relationship. (5) Glyn Elwyn operationalised this concept into a 12-step process. (8,9) In this broadly accepted model, patients are informed about the decision process and the pros and cons of treatment options. Then, patients' concerns and expectations are explicitly explored and incorporated into the treatment choice before the treatment plan is mutually determined. (8,9)

Despite impressive scientific efforts, effective methods of implementing this approach remain unclear. (9–11) Further, current knowledge on effective methods of directing professional behaviour towards more patient-centred care and SDM is scarce and inconsistent. (9,11) Effective methods of teaching physicians communication skills generally combine role-playing and feedback with small group discussions, and they should take at least one day. (11). Multifaceted interventions that include educating health professionals and decision aids, defined as instruments that prepare people to participate in decisions, are promoted to increase SDM behaviour. (9). Although these training sessions increase professionals' performance in SDM process elements, such as listing options, patient care is not adequately adjusted to include patient preferences. (3)

Time investment seems to be a necessary condition for implementing SDM because it is the most frequently mentioned barrier to introducing SDM into daily practice and because professionals' level of performance is associated with the consultation duration. (3,10)

To promote general practitioners' (GPs') positive reinforcement of *patients' expectations*, we developed a training program to teach GPs to implement SDM techniques and to positively reinforce the chosen therapy. This training program was part of an intervention study that compared the recovery of patients with non-chronic low back pain who consulted a GP trained in SDM and in positively reinforcing the chosen therapy with the recovery of similar patients who consulted untrained GPs.

We assessed whether GPs who were trained in SDM and in positively reinforcing treatment expectations demonstrated better SDM and reinforcement skills during consultations with patients with non-chronic low back pain than untrained GPs.

METHODS

Design

This study was embedded in a clustered randomised trial that evaluated the effectiveness of SDM among patients with non-chronic low back pain. For the trial, 68 GPs were recruited and randomly assigned to the intervention (n = 34) or control (n = 34) group. All participating GPs were asked to recruit 10 patients with non-chronic low back pain and to video-record their consultations with those patients. Of the consultations completed with 226 recruited patients, 175 consultations were video-recorded and used for this secondary analysis. (Fig. 1)

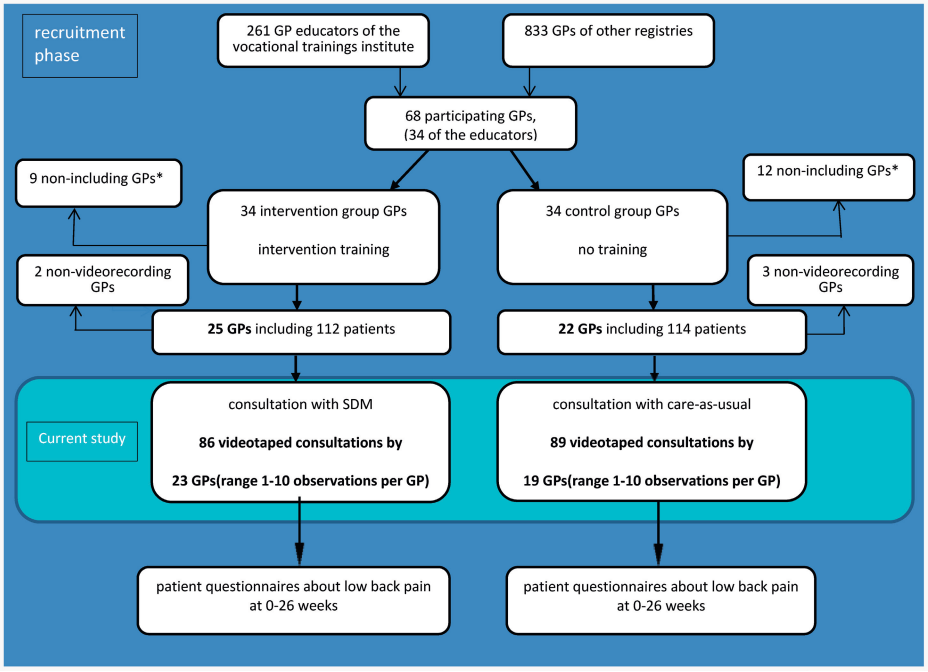


Figure 1. Flowchart

GP = general practitioner, SDM = shared decision-making, PR = positive reinforcement of the chosen therapy, * these GPs did not include any patient.

Participants

GPs were recruited from the vocational training institute in Utrecht and affiliated GP registries.

The training program

GPs in the intervention group received two training sessions that were each two and a half hours in duration and were held in small groups of approximately three to five participants. The training focused on the SDM process and evidence-based treatment of non-chronic low back pain according to professional guidelines. The GPs were encouraged to discuss the favourable prognosis of non-chronic low back pain with the patient and to positively reinforce the treatment that was jointly selected. The training was based on the learning principles described by Kolb and the SDM behavioural process elements developed by Elwyn. (Table 3) (12,13) In the training sessions, group discussion, theory, role-playing and reflections on personal behaviour were alternated. (13) A decision aid for non-chronic low back pain was developed based on the internationally accepted IPADS guidelines (Appendix A) (14) To stimulate their use of SDM skills during the actual consultations, we provided the GPs with a desktop tool containing group-formulated open-ended questions applicable to the consecutive SDM process elements and standard sentences that could be used to positively reinforce patients' treatment expectations. This tool was generated in the first session when the GPs reflected on their training experiences. The participants were advised but not obligated to use the plasticised A3 decision aid or the desktop tool in their encounters with patients. The training program was developed and implemented by a peer GP with expertise in training skills. (first author, AS) (Box 1)

In addition to receiving training sessions, the GPs in the intervention group received personalised feedback on each video-recorded consultation for a maximum of two consultations between the training sessions and for all consultations with recruited patients. The feedback was sent via email within 24 h after the trainer received the video-record, and it focused on performance on all SDM process elements, the extent to which the 'wait and see' approach was discussed, the extent to which the chosen treatment was positively reinforced and the level of autonomy in decision making.

Controls

GPs in the control group were not trained and provided usual care. They were informed that the intervention group was trained in 'communicative skills to maximise placebo effects' but were unaware of the content of the intervention.

Data collection

Practice staff, patients and observers were not informed of the training or group allocation. After GPs were recruited for the trial, they completed a questionnaire concerning their age, gender, educator status and years of GP experience. Practice staff invited patients to participate in the trial, and after the patients provided informed

consent, the patients completed one questionnaire before the consultation (baseline measurement). The GPs video-recorded the consultation about non-chronic low back pain. (Fig. 1)

Outcomes

The primary outcomes were as follows:

- level of *SDM* and
- level of *positive reinforcement of the chosen therapy*.

The secondary outcomes were as follows:

- level of *autonomy in decision making*,
- extent to which the *'wait and see'* approach was discussed,
- performance on the *consecutive process elements of the OPTION scale*, and
- *duration of the consultation*, which was divided into the intake, the physical examination and the evaluation and plan.

We compared the outcomes between the control and intervention group.

Measurement instruments

The OPTION scale and its process elements

We used the OPTION scale, a validated instrument used to measure GPs' performance on the SDM process elements. (Table 2) (12) Scores (ranging from 0 to 4) are given on 12 SDM process elements, and these scores are summed to obtain an overall score. (Table 2) A score of zero corresponds to no behaviour observed, and a score of four indicates that the behaviour is exhibited to a high degree. The scores for the 12 process elements are summed and transformed into a scale using the following formula: summed score/48x100. The scale exhibits adequate properties (Cronbach's alpha = 0.728).

Box 1

Training steps, aim and format both training sessions.

	Step	Aim	Format
	FIRST AND SECOND SESSION		
1	Introduction	Create a save environment	Introduction round
2	Inventory of attitude	1. Focus on placebo knowledge on recovery 2. Focus on learned skills of SDM&PR	Plenary short discussion
3	Reflection on daily practice	1. Create a shift from 'unaware unskilled' to 'aware unskilled' in SDM&PR 2. Create a situation of constructive friction in experienced skills or a lack of skills in SDM&PR	1. Group discussions of 2-3 trainees 2. Group discussions of 2-3 trainees coached by the trainer
4	Theory on placebo effects and SDM	1. Knowledge transfer on positive expectations on recovery and SDM process steps as instrument to implement patient's treatment expectations 2. Knowledge transfer on decision aid and double positive reinforcement [a]	Frontal presentation altered with plenary discussion based on questions on practical issues
5	Practical implications	To reach proportional understanding	
6	Practical work-out	1. Skills development in SDM process steps and positive reinforcement 2. Skills development in usage decision aid and double positive reinforcement [a]	1. Role play on simple LBP consultation 2. Role play with decision aid (obligatory) and decision tool [b] (voluntary)
7	Reflection on experiences	Solve problems, attack reluctance and increase self-efficacy	1. Plenary discussion and creation of desktop tool [b] 2. Plenary discussion ending with expressing feelings of 'aware skilled'

Training steps, aim and format both training sessions. Continued.

	Step	Aim	Format
	FIRST AND SECOND SESSION		
8	Ending	To thank and give a positive reinforcement on trained behaviour	A personal remark on best performances

1 = first session; 2 = s session SDM = shared decision-making, PR = positive reinforcement of the chosen therapy, LBP = non-chronic low back pain.

[a] Participants were taught to positive reinforce the benign course as recommended in the guideline but also positive reinforce the chosen therapy.

[b] Desktop tool contained group self-formulated open questions applicable to consecutive SDM process elements and sentences that could be pronounced to discuss the benign course or positively reinforce the treatment expectations.

Positive reinforcement of the chosen therapy

The value assigned to the level of positive reinforcement of the chosen therapy was unrelated to the individual who made the choice. The chosen therapy was defined as any selected treatment plan, including 'staying active', and the level of positive reinforcement was measured using the following rating scale (with some examples):

0. no remark on the effect of treatment
1. insufficient attempt (e.g. "I think it is good to do something about it")
2. basic skills (e.g., "It seems good to follow . . . therapy")
3. good level (e.g., "I fully expect that your complaints will fade within . . . weeks when you adhere to this therapy")
4. high level (i.e., the GP discussed the expected positive effect with the patient; e.g., "This therapy will take away your complaints. What do you think about it?" (The patient also responded to this question))

Level of autonomy in decision making

The Control Preference Scale was transformed into a scale to assess the observed level of autonomy in decision making with the following scores: 4 (autonomous), 3 (dominated by the patient), 2 (SDM), 1 (informed choice) and 0 (paternalistic). Observers determined scores based on their personal perception of the level of autonomy during the actual decision-making process while taking into account the preparation for the choice (which varied from GPs engaging in a more unidirectional informing approach to more active listening and dialogue in the decision phase) and the individual who made the actual choice. (15)

Some examples of each level of this scale are as follows:

4. Autonomous: "OK, I will start with physiotherapy since that suites me best."
3. Dominated by the patient: "Thank you for sharing this information about the different options. I think I will start taking painkillers because I have good experiences with taking pills."

2. SDM: “If I understand you correctly, the best choice for me/you could be just to wait and see because I/you had side effects from painkillers the last time I/you used them. What do you think?”
1. Informed choice: “It seems good for you to start taking pills because physiotherapy will cost too much of your precious time.”
0. Paternalistic: “This therapy is best for you. Let’s start with it now.”

Extent to which the ‘wait and see’ approach was discussed

The Dutch guidelines on low back pain recommend that medical professionals mention the benign spontaneous course of non-chronic low back pain (‘wait and see’) to increase patients’ expectations regarding recovery. (16) GPs in the intervention group were trained to explicitly express positive expectations about the spontaneous course of non-chronic low back pain and the ‘wait and see’ approach. The extent to which the ‘wait and see’ approach was discussed was measured in the same manner as positive reinforcement.

0. No remark on the ‘wait and see’ approach
1. Insufficient attempt: “I hope you will recover in the next few weeks”
2. Basic skills: “I expect you will recover in a couple of weeks.”
3. Good level: “Normally, your complaints fade away in days or weeks”
4. High level: “Your complaints will disappear in the coming days or in weeks at the most. Is that what you expected, too?” (The patient also responded to this question.)

Duration of the consultation

The total duration of the consultation was measured from the video-recording. The recording started when the patient entered the room and ended when the GP stopped the video-recording because the consultation about non-chronic low back pain had concluded. Time spent explaining trial-related issues or addressing other external interruptions, such as telephone calls, was subtracted from the overall time of the observation. The consultation was divided into three parts – namely, intake, physical examination and evaluation and plan – and the duration of each separate part was measured in seconds. The physical examination started when the patient stood to be examined and ended when the patient was re-seated. In cases in which no examination occurred, the intake ended when the GP began to discuss the diagnosis, therapeutic options, or plan. Time was measured in seconds.

Coding reliability

Two blind observers (AL and IvdE) scored the video-recordings using Observer (Noldus, 7th edition), a program designed to aid in the observation of video-recordings. (17) The observers scored a random sample of 17 consultations a second time to control for intra-observer reliability, and both observers scored another random sample of 32 consultations to control for inter-observer reliability. We calculated Pearson correlation coefficients and Cohen’s kappas for both inter-observer and intra-observer reliability.

The results are presented in Appendix B. Inter-observer reliability (Cohen's Kappa and Pearson correlation coefficient) for two of the twelve Option scale items was below 0.60; both items had a low frequency. However, to maintain the integrity of the OPTION scale, we kept these items for the analysis.

The level of positive reinforcement was not assessed reliably, and the inter-observer reliability for the level of autonomy in decision making was substantial (0.67).

Statistical analysis

Differences in the baseline characteristics of patients and GPs in the observed sample were tested against differences in the trial sample and between the group allocations. The effects on the OPTION sum score, level of positive reinforcement, level of involvement, extent to which the 'wait and see' approach was discussed and performance on the consecutive SDM process elements were assessed using *t*-tests based on the means. To test the differences in consultation duration, averages were subjected to Student's *t*-test. Because of the nested design (clustering of patients per GP), we employed a multilevel model to correct for clustering at the GP level. To control for learning effects due to the receipt of personalised feedback on each video-recording, we ranked each observation according to the order in which the GP conducted the session. Further, to correct for the patient's age, gender and educational level; the GP's age and gender; the gender of the patient-GP dyad; and rank of the video-recording and duration of the consultation, we built a multilevel linear regression model based on the means of the process elements and on the level of involvement separately by using the total score of the observer per allocation and tested whether the variables in this model had a significant influence on the outcomes by using a X^2 test.

Ethics

The study protocol for the trial (Netherlands National Trial Register (NTR) number: NTR1960) was assessed by the Ethical Committee of the University Medical Centre of Utrecht and exempted from full assessment. GPs asked patients to provide permission to be video-recorded.

RESULTS

GPs and consultations

Twenty-three trained GPs video-recorded 86 consultations, and 19 untrained GPs video-recorded 89 consultations. (Table 1) Three GPs in the intervention group did not complete both training sessions because of personal circumstances and did not recruit any patients. Further, two patients did not agree to the observation, 15 consultations were not video-recorded because the GP was not able to do so, 33 consultation observations failed and 1 observation was not scored for logistical reasons. (Fig. 1) Non-recruiting GPs did not differ from recruiting GPs in terms of age, gender, status as an educator of GP trainees or years of GP experience. In addition, video-recorded patients

did not differ from all 226 included patients in terms of age, gender, native country, educational level or pain characteristics. In the observed sample, there were also no significant differences in the baseline characteristics between the GPs and patients in the intervention group and the GPs and patients in the control group.

Table 1. Univariate analysis on baseline characteristics for untrained and trained GPs and their patients. For continuous variables, means and standard deviations tested by T-test and for dichotomous variables, percentages tested by X^2-distributions are given. Absolute numbers are in brackets.

	trained GP group		untrained GP group		
	mean	SD	mean	SD	p-value
GP CHARACTERISTICS					
number of GPs	23		19		
GP age	52.7	6.4	49.0	7.0	0.081
male	48% (11)		63% (12)		0.320
educator	65% (15)		84% (16)		0.163
number of years of experience as GP	18.8	7.0	19.1	8.9	0.597
number of patient inclusions	4.83	2.5	6.11	3.4	0.168
5 inclusions or more	48% (11)		63% (12)		0.320
PATIENT CHARACTERISTICS					
number of patients	86		89		
age	45.5	14.0	44.5	14.1	0.655
male	47% (40)		52% (45) †		0.546
Dutch origin	96%(75) ¶		99%(79) ‡		0.926
educational level	¶		†		0.969
low	18% (15)		17% (15)		
middle	49% (41)		49% (43)		
high	31% (26)		33% (29)		
VAS pain (0–100)	48.2††	15.6	46.9††	16.7	0.620

* = p<0.05; †=2 missing cases; ‡= 3 missing cases; ¶=4 missing cases; ††=14 missing cases; VAS=visual analogue scale.

Intervention effects

SDM and positive reinforcement of the chosen therapy

In the multilevel multivariate analyses, the trained GPs scored significantly higher on the OPTION scale and the level of positive reinforcement than the control GPs. Both scores were less than half the value recommended for best practice based on the maximum scores per scale. Nevertheless, the level of positive reinforcement was not assessed reliably and should be interpreted with caution.

Table 2. Multilevel multivariate analysis of SDM, as measured by OPTION sum scores (scale 0–100) and positive reinforcement (range 0–4), mean scores (confidence intervals) and differences. We corrected for patients' age, gender and educational level; GPs' age and gender; the gender of the patient-GP dyad, the rank of the video-recording and the consultation duration. The GP was the level of clustering.

	trained GP group (n = 86)		untrained GP group (n = 89)		
	mean	CI	mean	CI	difference
OPTION sum score on scale (0–100)	38.53	35.31–41.74	23.66	20.25–27.08	14.86*
positive reinforcement of the chosen therapy (0–4)	1.16	0.82–1.50	0.50	0.14–0.87	0.77*

* = $p < 0.05$ difference in scores between the two groups in the multilevel multivariate outcomes corrected for GP level, CI = confidence interval.

Level of autonomy in decision making and extent to which the 'wait and see' approach was discussed

In the multilevel multivariate analyses, the trained GPs exhibited significantly less paternalistic decision making (corresponding to a zero score on the Control Preference scale) than the control GPs but did not engage in SDM (corresponding to a score of two on the Control Preference scale). Control GPs discussed the 'wait and see' approach more explicitly, but the difference between the groups was not significant in the fully corrected analysis.

Table 3. Multilevel multivariate analysis of level of autonomy, discussion of the 'wait and see' approach and SDM consecutive process elements, mean scores (standard error). All scores range from 0 to 4.

		trained GP group (n = 86)		untrained GP group (n = 89)		
SECONDARY OUTCOMES		**mean**	**SE**	**mean**	**SE**	**difference**
level of autonomy (2 = SDM)		1.74	0.11	0.86	0.11	0.91*
discussion of the 'wait and see' approach		1.55	0.25	1.50	0.27	-0.05
item	*OPTION process elements*					
1	GP draws attention to a decision-making stage	1.18	0.15	0.52	0.16	0.66*
2	equipoise	1.21	0.23	0.17	0.25	1.03*
3	information format	0.84	0.12	0.25	0.13	0.59*
4	lists options	3.59	0.19	1.97	0.21	1.62*
5	explanation of pros and cons of options	1.65	0.14	0.98	0.15	0.67*
6	exploration of the patient's expectations	1.82	0.13	1.41	0.13	0.42*
7	exploration of the patient's concerns	0.13	0.05	0.02	0.05	0.11
8	check of patient's understanding	1.09	0.10	0.96	0.11	0.12
9	offering opportunities to ask questions	1.20	0.14	0.90	0.15	0.30
10	elicitation of patient's preferred level of involvement	2.18	0.18	0.96	0.19	1.22*
11	indication of a decision-making stage	0.88	0.09	0.38	0.10	0.50*
12	indication of the need to review the decision	2.67	0.33	2.82	0.35	-0.072

* = $p < 0.05$ difference in scores between the two groups in the multilevel multivariate outcomes.

Consecutive process elements of the OPTION scale

In the fully corrected model, the trained GPs performed significantly better on eight of the twelve elements of the OPTION scale, and the control GPs did not perform significantly better on any element.

Consultation duration effects

We found a significantly longer consultation duration for the trained group (15 min and 54 s) than for the untrained group (13 min and 6 s). In particular, the trained physicians spent significantly more time on the intake phase and the evaluation and plan phase but significantly less time on the physical examination.

The duration of the consultation had an important and significant positive effect ($p = 0.000$) on the summed score of the process elements.

Table 4. Duration (in seconds) of the observed consultations, mean scores (standard deviation).

	trained GP group (n = 86)		untrained GP group (n = 89)		
CONSULTATION PHASES	mean	SD	mean	SD	p-value
intake*	339	170	282	114	0.001
physical examination*	177	152	222	129	0.037
evaluation and plan*	437	246	283	164	0.000
total consultation duration*	947	359	786	271	0.001
total duration in minutes	15.8	6.0	13.1	4.5	0.001

* = $p < 0.05$ between the two groups.

DISCUSSION AND CONCLUSION

Discussion

This study evaluated the effects of training GPs in *SDM and positively reinforcing the chosen therapy*. Trained GPs engaged in SDM and positively reinforced the therapeutic choice significantly more than untrained GPs. However, the results regarding positive reinforcement are insufficiently reliable to draw firm conclusions. The training resulted in a 13% increase in the OPTION scale score and a 19% increase in the level of positive reinforcement, but the observed levels remained low. Further, consultations with trained GPs were 3 min longer than those with untrained GPs.

Our study confirms the findings of Couët's review of studies assessing GPs use of the OPTION scale. In the four comparison studies conducted in general practice settings in the review, untrained GPs exhibited comparably lower scores (a mean score on the OPTION magnitude scale of 26), and GPs who were trained in SDM achieved higher sum scores than controls (a mean score of 36 on the OPTION magnitude scale). (3) Similar to Couët, we found that the duration of the consultation was strongly correlated with performance in terms of SDM. However, patient-centred behaviour, such as questioning patients about the preferred information format, their concerns or their understanding, was rarely observed. (3)

Although results regarding the separate elements of the OPTION scale should be interpreted with caution, a comparison of the current findings with other relevant findings might unveil some similarities or differences between these elements. Similar to the EXACTE2 study, we found higher scores on 'announcing the decision-making stage', 'listing the options', 'discussing pros and cons' and 'exploring the patient's expectations'. (18,19) The last item was primarily observed in the intake phase. The Dutch guideline for low back pain recommends that GPs explores patients' treatment expectations in the intake phase. (16) Similar to the studies of Bensing et al. and Butalid et al., which evaluated changes in GP behaviour over time, we observed that GPs often did not ask patients questions in the decision phase. (20,21) Overall, the training directed GP behaviour towards more SDM and more positive reinforcement of treatment expectations; however, the major changes were found in the informative items of the OPTION scale, such as information transfer of therapeutic options or introduction of process steps. The training did not significantly increase patient-centred behaviour, as measured by more receptive items on the OPTION scale, such as listening to patients' concerns.

Training sessions are typically more effective at changing behaviour than attitudes. (22) This notion is confirmed by Couët's claim that behavioural change might be partially due to communication tools rather than clinicians' attempts to inquire about patients' preferences. (3) His interpretation is in line with our observations from the video-recordings and training. Specifically, the trained GPs became more aware of the need to better inform patients about their treatment options and to incorporate patients' expectations; however, they often made inquiries during the intake phase of the consultation and provided information in the form of a paternalistic, guideline-oriented choice. As Braddock suggested, GPs are more involved in the preparation for the decision than in the actual choice itself. (23) Furthermore, Butalid et al. and Bensing et al. concluded that in modern medicine, GPs' communication is more oriented towards biomedical tasks and characterised by lower levels of patient involvement. (20,21) Based on this interpretation and the limited inquiries about patients' expectations, concerns, need for questioning and preferred levels of involvement in the decision-making stage, we conclude that GPs in the intervention group engaged in more informative process elements of SDM and better inquired about patients' desired level of involvement in the decision making process than the controls; however, in terms of patients' involvement

in the actual treatment choice, there was considerable room for improvement. Thus, we expect that patients did not have more positive expectations about the treatment choices and that the GPs' treatment choice was reinforced in a typical manner.

However, why did the GPs adhere to a more paternalistic attitude? In the training, GPs considered it 'unprofessional' to accept a therapeutic option as the treatment choice when this option was not in line with their interpretation of the guidelines. Interestingly, GPs displayed individual differences in their preferred illness-specific treatment option even though they followed the same guidelines. During the training, the GPs became aware of the variety of preferred treatment options across GPs; however, they still expressed difficulties in accepting equipoise. The trained GPs were aware that placebo effects from expectations are substantial and that the guidelines reject these therapies because they do not exceed the placebo effect. They also experienced difficulties in accepting patients' personal preference of having an equal role in selecting the treatment. A sense of professionalism may have led them to feel the need to convince patients of their medical viewpoint rather than to consider the patient perspective. This attitude implies that SDM may be acceptable only from a medical perspective, i.e., when the patient prefers a treatment in line with illness-centred guidelines. Furthermore, professionals do not consider patients' treatment preferences based on placebo effects to be acceptable, even when the beneficial effects are based on evidence and the underlying mechanism of the effects is openly discussed. (24) These observations raise questions about the concept of SDM and professionals' and patients' perceptions of SDM. How do patients' preferences for treatment options relate to GPs' considerations as defined in clinical practice guidelines? Do considerations in illness-centred guideline reflect the patient's or the GP's perspective? To what extent does SDM imply an equal relationship in decision making? The answers to these questions can help to determine the best strategy for training providers to devote greater attention to the patient perspective in decision making.

The consultations with trained GPs lasted, on average, three minutes longer than those with untrained GPs, who spent 13 min delivering 'usual care'. Consultations in general practice in the Netherlands take, on average, 10.2 min. (25). Typically, a longer consultation duration is associated with a higher SDM-process score, and after duration is controlled for, the association between the intervention and the level of SDM becomes nonsignificant. (26) This was not the case in our study. Thus, in addition to the effects of a longer consultation duration, the training itself directed GPs towards more SDM. However, a 20% increase in consultation time is substantial in daily practice. One might question whether SDM could be implemented more efficiently if professionals invest less time engaging in task-oriented behaviour and more time exhibiting receptive behaviour, such as listening to patients' needs, to reach a treatment decision.

Strengths of the study

We developed a training and supportive tool that conforms to recent internationally accepted guidelines. (9,11) The duration of the training was sufficient considering

recently gained knowledge. (11,27) We used the OPTION scale to evaluate the training because it is a well-validated and broadly accepted instrument for measuring the conceptual construct of SDM. (3,12,28) Upon calculating the intra- and inter-observer rating reliability, we found reasonable scores for most of the OPTION scale items and the level of autonomy. The difference in the summed process elements between groups was congruent with the difference in the process outcome between the groups. We focused on not only reinforced positive expectations about treatment but also the benign natural course.
Time investment, a recognised barrier to SDM, was also take into account. (10)

Limitations of the study

The inter-observer reliability of the measurement of the level of positive reinforcement of the chosen therapy was insufficient – potentially because of the relatively low scores for this element. Therefore, conclusions regarding these results cannot be drawn with acceptable certainty.

We measured the effect of the training by comparing trained and untrained GPs but did not assess the behavioural change of each individual GP.

Not all trained GPs video-recorded consultations. However, we do not believe that this limitation induced selection bias because we did not find differences in the baseline variables between recruited and non-recruited patients. Although one-third of the GPs in the intervention group completed less than eight hours of training and, thus, did not receive feedback on more than four video-recordings, the observed behaviour differed significant compared to the GPs of the control group.

GPs were recruited and trained by a peer GP (the first author of the article), which might have triggered all GPs to perform better in terms of communication skills, contributed to the longer duration of the consultations and decreased the differences between the groups. Control GPs were not blind to the allocation or scope of the intervention, but they were not familiar with the content of the intervention. In two cases, GPs in the intervention group and in the control group worked on the same premises, but they worked on different days during the week. The majority of the GPs were educators, and they were recruited based on their interest in placebo effects on low back pain. Further, training intensity differed across trained GPs because not all GPs video-recorded consultations or reached the goal of including ten patients in the sample. These factors might have decreased the differences between the groups, especially with respect to the use of more general communication skills, such as ‘reviewing the decision’; performance on elements emphasised in the Dutch guidelines on low back pain, such as ‘exploring patient’s expectations’; and the extent to which the ‘wait and see’ approach was discussed. However, overall, the control GPs did not exceed the average scores found in Couët’s review, which highlights the difficulty of implementing SDM in daily practice. (16)

CONCLUSION

The results of the study indicate that training in *SDM and in positively reinforcing the chosen therapy* significantly increases SDM, especially with respect to behaviour related to informative items, such as the transfer of information from the GP to the patient. SDM did not lead GPs to incorporate the patient perspective in the actual decision-making process, for instance, by making inquiries about patients' concerns or need for questioning, to a greater extent. Unfortunately, the measurement of the level of positive reinforcement proved to be unreliable. Further, the consultations conducted by GPs in the intervention group were, on average, three minutes longer, and the GPs expressed reluctance in engaging in SDM when the patient's preferences were not in line with their interpretations of the clinical guideline.

PRACTICE IMPLICATIONS

In selecting a therapy, professionals should devote greater attention to the patient's perspective, including the patient's concerns, understanding and questions. To increase the level of patient involvement in actual decision making, scholars should conduct more research to understand patients' and professionals' perceptions regarding SDM.

CONTRIBUTORS

All authors were involved in the critical review of the manuscript and have read and approved the final version. The authors' specific contributions are as follows: study conception and design: JB, PV, William Verheul (Nivel), Margan Essed, NW and AS; sample acquisition and sequence data processing: Emily Fick, Margan Essed, Inge van Weeghel, Mijke van Gijn, Marieke van Noord, Jan Willem van Uffen, Lisanne Louisse, Annemarie Schatsnabel (all Julius Institute) and AS; analysis of epidemiological and sequence data: PS, PV, TM, and AS; drafting the manuscript: NW, JB, PV and AS. All authors had full access to all the study data and take responsibility for the integrity of the data and the accuracy of the data analysis.

CONfliCT OF INTEREST

None.

INFORMED CONSENT

The authors confirm that all patient/personal identifiers have been removed or disguised such that the patients/persons described are not identifiable.

ACKNOWLEDGEMENTS

We thank all patients and GPs who voluntarily participated in the trial and Hans van Lennep and Joeky Senders for piloting the observations of the video-recordings. We thank Amy Los (AL) and Ilse van de Ende (IvdE) for observing the video-recordings, Jan Maessen for transporting the observation data to statistical programs and Peter Spreeuwenberg for performing the statistical analyses.

APPENDIX A. DECISION AID (SUB)ACUTE ASPECIfiC LOW BACK PAIN

Front-side of plasticised A3 format decision aid card

What is wrong with me?

You suffer from an innocent low back pain. There is no mention of a serious abnormal situation. The cause of your pain is most probable situated in your muscles, ligaments and bones in the back, which are temporarily 'out of order'. This pain does not mean that there is a disease or any kind of damage.

What will happen to the symptoms of my pain?

The pain and complaints will vanish in short time. The more sever pain will go away in a few days. Three into four people will regain pain within a year.

Which therapeutic possibilities exist? Bed rest

Passive therapy for example: corset, acupuncture, etc

Manual therapy

Active therapy for example: physiotherapy, back schools, etc

Medication for example: NSAID's (ibuprofen, diclofenac), muscle relaxants (like Diazepam, Oxazepam etc.) or antidepressants.

Stay active this means: keep up your daily routine as much as possible. The best way to do this is to have your daily activities extended a little, in spite of the pain.

Which therapy is best?

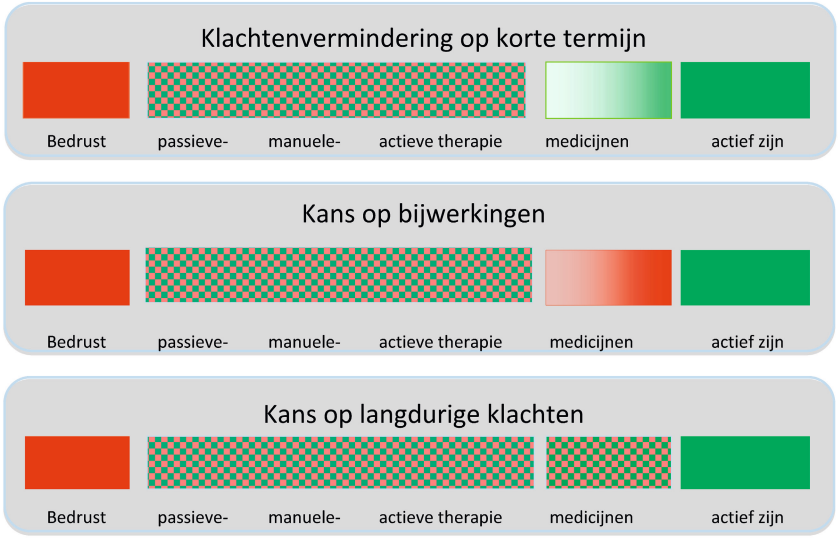
All therapies have their advantages and disadvantages. As a physician we cannot tell you which one will be best for you.

You can choose to wait and see whether pain will go away by itself. What we do know is that your recovery will be sooner if you stay active as much as possible and that staying in bed has negative results.

Back-side of plasticised A3 format decision aid card

In the tables mentioned below you can see how effective different therapies are for the majority of people on short term. What are chances on adverse effects and on chronic complaints. The tables can help you to decide together with your physician how you want to cope with your back pain. You can also choose to combine different therapies.

This color means that we don't know if this therapy will help. Maybe you have your own preferences



Additional background information on A4 paper for GPs

Conclusion based on Cochrane reviews
Presented are summaries of results. Differences between *acute/sub-acute or chronic are highlighted by italics*.
For authors and content we refer to Cochrane website. http://www.thecochranelibrary.com

Advice to rest in bed versus advice to stay active for acute low back pain and sciatica
June 2010
We included ten RCTs with varying risk of bias. For patients with *acute low back pain*, results from two trials (N = 401) suggest small improvements in pain relief (SMD 0.22 (95% CI: 0.02–0.41)) and functional status (SMD 0.29 (95% CI: 0.09–0.49)) in favor of advice to stay active. For patients with *sciatica*, there is moderate quality evidence of little or no difference in pain relief (SMD 0.03 (95% CI: 0.24 to 0.18)) or functional status (SMD 0.19 (95% CI: 0.02 to 0.41)), between advice to rest in bed or stay active.
Low quality evidence (3 RCTs, N = 931) suggests little or no difference between exercises, advice to rest in bed or stay active for patients with *acute low back pain*. Low quality evidence (1 RCT, N = 250) suggests little or no difference between physiotherapy, advice to rest in bed or stay active for patients with *sciatica*. No trials that compared different ways of delivering advice.

Lumbar supports for prevention and treatment of low back pain
February 2011
Seven preventive studies (14,437 people) and eight treatment studies (1361 people) were included in this updated review. Overall, the methodological quality of the studies was rather low. Only five of the fifteen studies met 50% or more of the internal validity items. There was moderate evidence that lumbar supports are not more effective than no intervention or training in preventing low back pain, and conflicting evidence whether lumbar supports are effective supplements to other preventive interventions. It is still unclear if lumbar supports are more effective than no or other interventions for the treatment of low back pain.

Superficial heat or cold for low back pain
February 2011
Nine trials involving 1117 participants were included. In two trials of 258 participants with a mix of acute and sub-acute low back pain, heat wrap therapy significantly reduced pain after five days (weighted mean difference (WMD) 1.06, 95% confidence interval (CI) 0.68–1.45, scale range 0–5) compared to oral placebo. One trial of 90 participants with *acute low back pain* found that a heated blanket significantly decreased acute low-back pain immediately after application (WMD 32.20, 95% CI 38.69 to 25.71, scale range 0–100). One trial of 100 participants with a mix of *acute and subacute low back*

pain examined the additional effects of adding exercise to heat wrap, and found that it reduced pain after seven days. There is insufficient evidence to evaluate the effects of cold for low-back pain, and conflicting evidence for any differences between heat and cold for low back pain.

Acupuncture and dry needling for low back pain
February 2011
Thirty-five RCTs were included; 20 were published in English, seven in Japanese, five in Chinese and one each in Norwegian, Polish and German. There were only three trials of acupuncture for *acute low back pain*. They did not justify firm conclusions, because of small sample sizes and low methodological quality of the studies. For *chronic low back pain* there is evidence of pain relief and functional improvement for acupuncture, compared to no treatment or sham therapy. These effects were only observed immediately after the end of the sessions and at short-term follow-up. There is evidence that acupuncture, added to other conventional therapies, relieves pain and improves function better than the conventional therapies alone. However, effects are only small. Dry needling appears to be a useful adjunct to other therapies for chronic low back pain. No clear recommendations could be made about the most effective acupuncture technique.

Massage for low back pain
June 2010
Thirteen randomised trials were included. Eight had a high risk and five had a low risk of bias. One study was published in German and the rest in English. Massage was compared to an inert therapy (sham treatment) in two studies that showed that massage was superior for pain and function on both short and long-term follow-ups. In eight studies, massage was compared to other active treatments. They showed that massage was similar to exercises, and massage was superior to joint mobilisation, relaxation therapy, physical therapy, acupuncture and self-care education. One study showed that reflexology on the feet had no effect on pain and functioning. The beneficial effects of massage in patients with chronic low back pain lasted at least one year after the end of the treatment. Two studies compared two different techniques of massage. One concluded that acupuncture massage produces better results than classic (Swedish) massage and another concluded that Thai massage produces similar results to classic (Swedish) massage.
Author conclusions: Massage might be beneficial for patients with *subacute and chronic aspecific low back pain*, especially when combined with exercises and education. The evidence suggests that acupuncture massage is more effective than classic massage, but this need confirmation. More studies are needed to confirm these conclusions, to assess the impact of massage on return-to-work, and to determine cost-effectiveness of massage as an intervention for low back pain.

Combined chiropractic interventions for low back pain
February 2011
We included 12 studies involving 2887 participants with low back pain. Three studies had low risk of bias. Included studies evaluated a range of chiropractic procedures in a variety of sub-populations of people with low back pain. No trials were located of combined chiropractic interventions compared to no treatment. For *acute and subacute low back pain*, chiropractic interventions improved short- and medium-term pain (SMD 0.25 (95% CI 0.46 to 0.04) and MD 0.89 (95% CI 1.60 to 0.18)) compared to other treatments, but there was no significant difference in long-term pain (MD 0.46 (95% CI 1.18 to 0.26)). Short-term improvement in disability was greater in the chiropractic group compared to other therapies SMD 0.36 (95% CI 0.70 to 0.02)). However, the effect was small and all studies contributing to these results had high risk of bias. There was no difference in medium- and long-term disability. No difference was demonstrated for combined chiropractic interventions for *chronic low back pain* and for studies that had a mixed population of low back pain.

Exercise therapy for treatment of aspecific low back pain
February 2011
Sixty-one randomised controlled trials (6390 participants) met inclusion criteria: *acute (11), subacute (6) and chronic (43) low back pain (1 unclear)*. Evidence was found of effectiveness in *chronic* populations relative to comparisons at all follow-up periods; pooled mean improvement was 7.3 points (95% CI, 3.7–10.9) for pain (out of 100), 2.5 points (1.0–3.9) for function (out of 100) at earliest follow-up. In studies investigating patients (i.e. presenting to healthcare providers) mean improvement was 13.3 points (5.5– 21.1) for pain, 6.9 (2.2–11.7) for function, representing significantly greater improvement over studies where participants included those recruited from a general population (e.g. with advertisements). There is some evidence of effectiveness of graded-activity exercise program *in subacute low back pain* in occupational settings, though the evidence for other types of exercise therapy in other populations is inconsistent. There was evidence of equal effectiveness relative to comparisons in *acute populations* (pain: 0.03 points (95% CI, 1.3 to 1.4).
Limitations: This review largely reflects limitations of the literature, including low quality studies with heterogeneous outcome measures, inconsistent and poor reporting, and possibility of publication bias.

Back schools for non-specific low back pain
March 2010
Nineteen RCTs (3584 patients) were included in this updated review. Overall, the methodological quality was low, with only six trials considered to be high quality. It was not possible to perform relevant subgroup analyses for low back pain with radiation versus low back pain without radiation. The results indicate that there is moderate evidence suggesting that back schools have better short and intermediate-term

effects on pain and functional status than other treatments for patients with *recurrent and chronic low back pain*. There is moderate evidence suggesting that back schools for *chronic low back pain* in an occupational setting, are more effective than other treatments and placebo or waiting list controls on pain, functional status and return to work during short and intermediate-term follow-up. In general, the clinical relevance of the studies was rated as insufficient.

Non-steroidal anti-inflammatory drugs for low back pain
March 2010
In total, 65 trials (total number of patients = 11,237) were included in this review. Twenty-eight trials (42%) were considered high quality. Statistically significant effects were found in favor of NSAIDs compared to placebo, but at the cost of statistically significant more side effects. There is moderate evidence that NSAIDs are not more effective than paracetamol for *acute low back pain*, but paracetamol had fewer side effects. There is moderate evidence that NSAIDs are not more effective than other drugs for *acute low-back pain*. There is strong evidence that various types of NSAIDs, including COX-2 NSAIDs, are equally effective for *acute low back pain*. COX-2 NSAIDs had statistically significantly fewer side-effects than traditional NSAIDs.

Muscle relaxants for aspecific low back pain
October 2008
Thirty trials met the inclusion criteria. Twenty-three trials (77%) were of high quality, 24 trials (80%) were on *acute low back pain*. Four trials studied benzodiazepines, 11 non-benzodiazepines and two anti-spasticity muscle relaxants in comparison with placebo. Results showed that there is strong evidence that any of these muscle relaxants are more effective than placebo for patients with *acute low back pain* on short-term pain relief. The pooled RR for non-benzodiazepines versus placebo after two to four days was 0.80 (95% CI; 0.71 to 0.89) for pain relief and 0.49 (95% CI; 0.25 to 0.95) for global efficacy. Adverse events, however, with a relative risk of 1.50 (95% CI; 1.14 to 1.98) were significantly more prevalent in patients receiving muscle relaxants and especially the central nervous system adverse effects (RR 2.04; 95% CI; 1.23 to 3.37). The various muscle relaxants were found to be similar in performance.

Antidepressants for aspecific low back pain
October 2010
Ten trials that compared antidepressants with placebo were included in this review. The pooled analyses showed no difference in pain relief (six trials (one trial with two treatment arms and a second trial with 3 treatment arms); standardised mean difference (SMD) 0.04 (95% confidence interval (CI) 0.25 to 0.17)) or depression (two trials; SMD 0.06 (95% CI 0.29 to 0.40)) between antidepressant and placebo treatments. The qualitative analyses found conflicting evidence on the effect of antidepressants on pain intensity in *chronic low back pain*, and no clear evidence that antidepressants reduce

depression in chronic low back pain patients. Two pooled analyses showed no difference in pain relief between different types of antidepressants and placebo. Our findings were not altered by the sensitivity analyses, which varied the risk of bias allowed for inclusion in the meta-analyses to allow data from additional trials to be examined.

Individual patient education for low back pain
February 2010
Of the 24 studies included in this review, 14 (58%) were of high quality. Individual patient education was compared with no intervention in 12 studies; with non-educational interventions in 11 studies; and with other individual educational interventions in eight studies. Results showed that for patients with *subacute low back pain*, there is strong evidence that an individual 2.5 h oral educational session is more effective on short-term and long-term return-to-work than no intervention. Educational interventions that were less intensive were not more effective than no intervention. Furthermore, there is strong evidence that individual education for patients with *(sub)acute low back pain* is as effective as non-educational interventions on long-term pain and global improvement and that for *chronic patients*, individual education is less effective for back pain-specific function when compared to more intensive interventions. Comparison of different types of individual education did not show significant differences.

APPENDIX B.

Inter-observer and intra-observer reliability of SDM process elements, positive reinforcement of the chosen therapy, level of autonomy and discussion of the benign course measured by Spearman's correlation efficient and Cohen's kappa of dichotomised scores.

All inter-observer scores are on 32 observations. Scores were dichotomised based on a frequency closest to 50%.

	INTER-OBSERVER		INTRA-OBSERVER			
COMMUNICATIVE ELEMENT	Spearman's correlation	Cohen's kappa	Spearman's correlation¶	Cohen's kappa¶	Mean per observer¶	SD per observer¶
GP draws attention to a decision-making stage	0.63	0.65‡	0.89	0.89‡	0.88	0.75
			0.48	0.16‡	0.69	0.59
equipoise	0.44	0.35‡	0.21	0.29‡	0.88	1.24
			0.36	0.14‡	0.84	1.14
information format	0.86	0.86‡	0.90	0.89‡	0.34	0.48
			0.99	1.00‡	0.34	0.48
lists options	0.88	0.75†	0.99	1.000†	2.53	1.27
			0.86	0.88†	1.16	0.85
explanation of pros and cons of options	0.70	0.55	0.74	0.78‡	1.16	0.88
			0.87	0.82‡	1.72	0.58
exploration of the patient's expectations	0.70	0.78†	0.60	0.59†	1.72	0.58
			0.27	0.41†	1.59	0.67
exploration of the patient's concerns	n.a.	n.a.	n.a.	n.a.	0.00	0.00
					0.00	0.00
check of patient's understanding	0.94	0.87‡	0.79	n.a. ‡	0.97	0.54
			0.84	n.a. ‡	1.00	0.51
offering opportunities to ask questions	0.79	0.71‡	0.92	0.88‡	1.16	0.72
			0.68	0.65‡	1.13	0.76
elicitation of patient's preferred level of involvement*	0.83	0.75†	0.73	0.49†	1.31	1.06
			0.91	1.00†	1.53*	1.05*
indication of a decision-making stage	0.13	0.13‡	0.24	0.22‡	0.41	0.50
			0.98	1.00‡	0.31	0.47
indication of the need to review the decision	0.79	0.93‡	0.72	0.70‡	2.88	1.83
			1.00	1.00‡	2.28	1.84
total score*	0.92	0.61††	0.92	0.67††	14.42	4.98
			0.93	0.75††	13.42	4.86
positive reinforcement of the chosen therapy	0.39	0.18‡	0.80	0.80‡	0.63	0.79
			0.67	0.64‡	0.66	0.83
level of autonomy (SDM = 2)	0.67	0.59†	0.84	0.80†	1.19	0.82
			0.81	0.77†	1.16	0.77

Continued

	INTER-OBSERVER		INTRA-OBSERVER			
COMMUNICATIVE ELEMENT	Spearman's correlation	Cohen's kappa	Spearman's correlation¶	Cohen's kappa¶	Mean per observer¶	SD per observer¶
discussion of the benign	0.89	1.00‖	0.853	0.80‖	2.09	1.20
course			0.96	1.00‖	1.97	1.31

¶ = number above is IvEe and number under is AI,
n.a. = not applicable because of a constant (zero for both scores),
* = one observation is missing,
‡ = dichotomisation scores 0 = 1 and scores 1 thru 4 = 2,
† = dichotomisation scores 0 thru 1 = 1 and scores 2 thru 4 = 2,
‖ = dichotomisation scores 0 thru 2 = 1 and scores 3 thru 4 = 2,
†† = dichotomisation scores 1–13 = 1 and scores 14–27 = 2.

REFERENCES

1) J. Bensing, Bridging the gap. The separate worlds of evidence-based medicine and patient-centered medicine, Patient Educ. Couns. 39 (2000) 17–25.
2) P.F.M. Stalmeijer, J.J. van Tol-Geerdink, E.N.J.T. van Lin, E. Schimmel, H. Huizenga, W.A.J. van Daal, J.W.H. Leer, De patient kiest: Werkbaar en effectief, Ned. Tijdschr. Geneeskd. 153 (2009) 600–606.
3) N. Couët, S. Desroches, H. Robitaille, H. Vaillancourt, A. Leblanc, S. Turcotte, G. Elwyn, F. Légaré, Assessments of the extent to which health-care providers involve patients in decision making: a systematic review of studies using the OPTION instrument, Health Expect. 18 (2015) 542–561.
4) J.S. Khan, P.J. Devereaux, Y. LeManach, J.W. Busse, Patient coping and expectations about recovery predict the development of chronic post-surgical pain after traumatic tibial fracture repair, Br. J. Anaesth. 117 (2016) 365–370.
5) A.M. Stiggelbout, T. Van der Weijden, M.P. De Wit, D. Frosch, F. Legare, V.M. Montori, et al., Shared decision making: really putting patients at the centre of healthcare, BMJ 344 (January (27)) (2012) e256.
6) F. Benedetti, M. Amanzio, Mechanisms of the placebo response, Pulm. Pharmacol. Ther. 26 (2013) 520–523.
7) K.B. Thomas, General practice consultations: is there any point in being positive? Br. Med. J. (Clin. Res. Ed.) 294 (1987) 1200–1202.
8) G. Elwyn, A. Edwards, P. Kinnersley, R. Grol, Shared decision making and the concept of equipoise: the competences of involving patients in healthcare choices, Br. J. Gen. Pract. 50 (2000) 892–899.
9) F. Legare, D. Stacey, S. Turcotte, M.J. Cossi, J. Kryworuchko, I.D. Graham, A. Lyddiatt, M.C. Politi, R. Thomson, G. Elwyn, N. Donner-Banzhoff, Interventions for improving the adoption of shared decision making by healthcare professionals, Cochrane Database Syst. Rev. 9 (2014) CD006732.
10) F. Légaré, S. Ratté, K. Gravel, I.D. Graham, Barriers and facilitators to implementing shared decision-making in clinical practice: update of a systematic review of health professionals' perceptions, Patient Educ. Couns. 73 (2008) 526–535.
11) M. Berkhof, H.J. van Rijssen, A.J. Schellart, J.R. Anema, A.J. van der Beek, Effective training strategies for teaching communication skills to physicians: an overview of systematic reviews, Patient Educ. Couns. 84 (2011) 152–162.
12) G. Elwyn, H. Hutchings, A. Edwards, F. Rapport, M. Wensing, W.Y. Cheung, R. Grol, The OPTION scale: measuring the extent that clinicians involve patients in decision-making tasks, Health Expect. 8 (2005) 34–42.
13) A.Y. Kolb, D.A. Kolb, Learning styles and learning spaces: enhancing experiential learning in higher education, Acad. Manage. Learn. Educ. 4 (2005) 193–212.
14) G. Elwyn, A.M. O'Connor, C. Bennett, R.G. Newcombe, M. Politi, M.A. Durand, E. Drake, N. Joseph-Williams, S. Khangura, A. Saarimaki, S. Sivell, M. Stiel, S.J. Bernstein, N. Col, A. Coulter, K. Eden, M. Härter, M.H. Rovner, N. Moumjid, D. Stacey, Assessing the quality of decision support technologies using the International Patient Decision Aid standards instrument (IPDASi), PLoS One 4 (2009) e4705.
15) L.F. Degner, J.A. Sloan, P. Venkatesh, The control preferences scale, Can. J. Nurs. Res. 29 (1997) 21–43.
16) Nederlands Huisartsen Genootschap, NHG-Standaard Aspecifieke lagerugpijn M54 (Internet).; 2005. http://www.springerlink.com.proxy.library.uu.nl/ content/j12n60k507873536/.

17) L.P. Noldus, R.J. Trienes, A.H. Hendriksen, H. Jansen, R.G. Jansen, The Observer Video-Pro: new software for the collection, management, and presentation of time-structured data from videotapes and digital media files, Behav. Res. Methods Instrum. Comput. 32 (2000) 197–206.
18) G. Elwyn, P.J. Barr, S.W. Grande, R. Thompson, T. Walsh, E.M. Ozanne, Developing CollaboRATE: a fast and frugal patient-reported measure of shared decision making in clinical encounters, Patient Educ. Couns. 93 (2013) 102– 107.
19) G. Elwyn, M. Tsulukidze, A. Edwards, F. Légaré, R. Newcombe, Using a 'talk' model of shared decision making to propose an observation-based measure: observer OPTION 5 Item, Patient Educ. Couns. 93 (2013) 265–271.
20) J. Bensing, F. Tromp, S. van Dulmen, A. van den Brink-Muinen, W. Verheul, F. Schellevis, De zakelijke huisarts en de niet-mondige patiënt: veranderingen in communicatie Een observatieonderzoek naar consulten met hypertensiepatiënten, Huisarts Wetenschap 51 (2008) 6–12.
21) L. Butalid, J.M. Bensing, P.F. Verhaak, Talking about psychosocial problems: an observational study on changes in doctor-patient communication in general practice between 1977 and 2008, Patient Educ. Couns. 94 (2014) 314–321.
22) K. Bombeke, L. Symons, E. Vermeire, L. Debaene, S. Schol, B. De Winter, P. Van Royen, Patient-centredness from education to practice: the 'lived' impact of communication skills training, Med. Teach. 34 (2012) e338–e348.
23) C.H. Braddock 3rd, The emerging importance and relevance of shared decision making to clinical practice, Med. Decis. Making 30 (5 Suppl) (2010) 5S–7S.
24) S. Kam-Hansen, M. Jakubowski, J.M. Kelley, I. Kirsch, D.C. Hoaglin, T.J. Kaptchuk, R. Burstein, Altered placebo and drug labeling changes the outcome of episodic migraine attacks, Sci. Transl. Med. 6 (2014) 218ra5.
25) M. Deveugele, A. Derese, A. van den Brink-Muinen, J. Bensing, J. De Maeseneer, Consultation length in general practice: cross sectional study in six European countries, BMJ 325 (2002) 472.
26) M.A. Pellerin, G. Elwyn, M. Rousseau, D. Stacey, H. Robitaille, F. Légaré, Toward shared decision making: using the OPTION scale to analyze resident-patient consultations in family medicine, Acad. Med. 86 (2011) 1010–1018.
27) F. Dwamena, M. Holmes-Rovner, C.M. Gaulden, S. Jorgenson, G. Sadigh, A. Sikorskii, S. Lewin, R.C. Smith, J. Coffey, A. Olomu, Interventions for providers to promote a patient-centred approach in clinical consultations, Cochrane Database Syst. Rev. 12 (2012) CD003267.
28) G. Elwyn, A. Edwards, M. Wensing, K. Hood, C. Atwell, R. Grol, Shared decision making: developing the OPTION scale for measuring patient involvement, Qual. Saf. Health Care 12 (2003) 93–99.
29) J. Bensing, Bridging the gap. The separate worlds of evidence-based medicine and patient-centered medicine, Patient Educ. Couns. 39 (2000) 17–25.
30) P.F.M. Stalmeijer, J.J. van Tol-Geerdink, E.N.J.T. van Lin, E. Schimmel, H. Huizenga, W.A.J. van Daal, J.W.H. Leer, De patient kiest: Werkbaar en effectief, Ned. Tijdschr. Geneeskd. 153 (2009) 600–606.
31) N. Couët, S. Desroches, H. Robitaille, H. Vaillancourt, A. Leblanc, S. Turcotte, G. Elwyn, F. Légaré, Assessments of the extent to which health-care providers involve patients in decision making: a systematic review of studies using the OPTION instrument, Health Expect. 18 (2015) 542–561.
32) J.S. Khan, P.J. Devereaux, Y. LeManach, J.W. Busse, Patient coping and expectations about recovery predict the development of chronic post-surgical pain after traumatic tibial fracture repair, Br. J. Anaesth. 117 (2016) 365–370.

33) A.M. Stiggelbout, T. Van der Weijden, M.P. De Wit, D. Frosch, F. Legare, V.M. Montori, et al., Shared decision making: really putting patients at the centre of healthcare, BMJ 344 (January (27)) (2012) e256.
34) F. Benedetti, M. Amanzio, Mechanisms of the placebo response, Pulm. Pharmacol. Ther. 26 (2013) 520–523.
35) K.B. Thomas, General practice consultations: is there any point in being positive? Br. Med. J. (Clin. Res. Ed.) 294 (1987) 1200–1202.
36) G. Elwyn, A. Edwards, P. Kinnersley, R. Grol, Shared decision making and the concept of equipoise: the competences of involving patients in healthcare choices, Br. J. Gen. Pract. 50 (2000) 892–899.
37) F. Legare, D. Stacey, S. Turcotte, M.J. Cossi, J. Kryworuchko, I.D. Graham, A. Lyddiatt, M.C. Politi, R. Thomson, G. Elwyn, N. Donner-Banzhoff, Interventions for improving the adoption of shared decision making by healthcare professionals, Cochrane Database Syst. Rev. 9 (2014) CD006732.
38) F. Légaré, S. Ratté, K. Gravel, I.D. Graham, Barriers and facilitators to implementing shared decision-making in clinical practice: update of a systematic review of health professionals' perceptions, Patient Educ. Couns. 73 (2008) 526–535.
39) M. Berkhof, H.J. van Rijssen, A.J. Schellart, J.R. Anema, A.J. van der Beek, Effective training strategies for teaching communication skills to physicians: an overview of systematic reviews, Patient Educ. Couns. 84 (2011) 152–162.
40) G. Elwyn, H. Hutchings, A. Edwards, F. Rapport, M. Wensing, W.Y. Cheung, R. Grol, The OPTION scale: measuring the extent that clinicians involve patients in decision-making tasks, Health Expect. 8 (2005) 34–42.
41) A.Y. Kolb, D.A. Kolb, Learning styles and learning spaces: enhancing experiential learning in higher education, Acad. Manage. Learn. Educ. 4 (2005) 193–212.
42) G. Elwyn, A.M. O'Connor, C. Bennett, R.G. Newcombe, M. Politi, M.A. Durand, E. Drake, N. Joseph-Williams, S. Khangura, A. Saarimaki, S. Sivell, M. Stiel, S.J. Bernstein, N. Col, A. Coulter, K. Eden, M. Härter, M.H. Rovner, N. Moumjid, D. Stacey, Assessing the quality of decision support technologies using the International Patient Decision Aid standards instrument (IPDASi), PLoS One 4 (2009) e4705.
43) L.F. Degner, J.A. Sloan, P. Venkatesh, The control preferences scale, Can. J. Nurs. Res. 29 (1997) 21–43.
44) Nederlands Huisartsen Genootschap, NHG-Standaard Aspecifieke lagerugpijn M54 (Internet).; 2005. http://www.springerlink.com.proxy.library.uu.nl/ content/j12n60k507873536/.
45) L.P. Noldus, R.J. Trienes, A.H. Hendriksen, H. Jansen, R.G. Jansen, The Observer Video-Pro: new software for the collection, management, and presentation of time-structured data from videotapes and digital media files, Behav. Res. Methods Instrum. Comput. 32 (2000) 197–206.
46) G. Elwyn, P.J. Barr, S.W. Grande, R. Thompson, T. Walsh, E.M. Ozanne, Developing CollaboRATE: a fast and frugal patient-reported measure of shared decision making in clinical encounters, Patient Educ. Couns. 93 (2013) 102– 107.
47) G. Elwyn, M. Tsulukidze, A. Edwards, F. Légaré, R. Newcombe, Using a 'talk' model of shared decision making to propose an observation-based measure: observer OPTION 5 Item, Patient Educ. Couns. 93 (2013) 265–271.
48) J. Bensing, F. Tromp, S. van Dulmen, A. van den Brink-Muinen, W. Verheul, F. Schellevis, De zakelijke huisarts en de niet-mondige patiënt: veranderingen in communicatie Een observatieonderzoek naar consulten met hypertensiepatiënten, Huisarts Wetenschap 51 (2008) 6–12.

49) L. Butalid, J.M. Bensing, P.F. Verhaak, Talking about psychosocial problems: an observational study on changes in doctor-patient communication in general practice between 1977 and 2008, Patient Educ. Couns. 94 (2014) 314–321.
50) K. Bombeke, L. Symons, E. Vermeire, L. Debaene, S. Schol, B. De Winter, P. Van Royen, Patient-centredness from education to practice: the 'lived' impact of communication skills training, Med. Teach. 34 (2012) e338–e348.
51) C.H. Braddock 3rd, The emerging importance and relevance of shared decision making to clinical practice, Med. Decis. Making 30 (5 Suppl) (2010) 5S–7S.
52) S. Kam-Hansen, M. Jakubowski, J.M. Kelley, I. Kirsch, D.C. Hoaglin, T.J. Kaptchuk, R. Burstein, Altered placebo and drug labeling changes the outcome of episodic migraine attacks, Sci. Transl. Med. 6 (2014) 218ra5.
53) M. Deveugele, A. Derese, A. van den Brink-Muinen, J. Bensing, J. De Maeseneer, Consultation length in general practice: cross sectional study in six European countries, BMJ 325 (2002) 472.
54) M.A. Pellerin, G. Elwyn, M. Rousseau, D. Stacey, H. Robitaille, F. Légaré, Toward shared decision making: using the OPTION scale to analyze resident-patient consultations in family medicine, Acad. Med. 86 (2011) 1010–1018.
55) F. Dwamena, M. Holmes-Rovner, C.M. Gaulden, S. Jorgenson, G. Sadigh, A. Sikorskii, S. Lewin, R.C. Smith, J. Coffey, A. Olomu, Interventions for providers to promote a patient-centred approach in clinical consultations, Cochrane Database Syst. Rev. 12 (2012) CD003267.
56) G. Elwyn, A. Edwards, M. Wensing, K. Hood, C. Atwell, R. Grol, Shared decision making: developing the OPTION scale for measuring patient involvement, Qual. Saf. Health Care 12 (2003) 93–99.

CHAPTER 4

The effectiveness of shared decision-making followed by positive reinforcement on physical disability in the long-term follow-up of patients with non-chronic low back pain in primary care: a clustered randomised controlled trial

Ariëtte R.J. Sanders
Jozien M. Bensing
Tessa Magnée
Peter F.M. Verhaak
Niek J. de Wit

BMC Family Practice. 2018 June 28; 102-8

ABSTRACT

Background. Although the recovery of patients suffering from non-chronic low back pain is highly context dependent, patient preferences about treatment options are seldom incorporated into the therapeutic plan. Shared decision-making (SDM) offers a tool to overcome this deficiency. The reinforcement by the general practitioner (GP) of a 'shared' chosen therapy might increase patients' expectations of favourable outcomes and thus contribute to recovery.

Methods. In the Netherlands, a clustered randomised controlled trial was performed to assess the effectiveness of shared decision-making followed by positive reinforcement of the chosen therapy (SDM&PR) on patient-related clinical outcomes. Overall, 68 GPs included 226 patients visiting their GP for a new episode of non-chronic low back pain. GPs in the intervention group were trained in implementing SDM&PR using a structured training programme with a focus on patient preferences in reaching treatment decisions. GPs in the control group provided care as usual. The primary outcome was the change in physical disability measured with the Roland-Morris disability questionnaire (RMD) during the six-month follow-up after the first consultation. Physical disability (RMD), pain, adequate relief, absenteeism and healthcare consumption at 2, 6, 12 and 26 weeks were secondary outcomes. A multivariate analysis with a mixed model was used to estimate the differences in outcomes.

Results. Of the patients in the intervention and the control groups, 66% and 62%, respectively, completed the follow-up. Most patients (77%) recovered to no functional restrictions due to back pain within 26 weeks. No significant differences in the mean scores for any outcome were observed between intervention patients and controls during the follow-up, and in multivariate analysis, there was no significant difference in the main outcome during the six-month follow-up. Patients in the intervention group reported more involvement in decision-making.

Conclusion. This study did not detect any improvement in clinical outcome or in healthcare consumption of patients with non-chronic low back pain after the training of GPs in SDM&PR. The implementation of SDM merely introduces task-oriented communication. The training of the GPs may have been more effective if it had focused more on patient-oriented communication techniques and on stressing the expectation of favourable outcomes.

Trial registration. The Netherlands National Trial Register (NTR) number: NTR1960. The trial was registered in the NTR on August 20, 2009.

BACKGROUND

Low back pain is defined as back pain localised below the costal margin and the inferior gluteal folds, with or without referred leg pain, and without a specific somatic origin. (1, 2)

Low back pain can be divided into acute, with a duration of complaints < 6 weeks, subacute, with a duration of complaints between 6 and 12 weeks, and chronic, with complaints lasting longer than 3 months. (1–3)

Low back pain has a lifetime prevalence of 60–85%. (4) Most episodes of low back pain resolve after two weeks, but the recurrence rate is high; three-quarters of patients have a second episode within one year. (4) Because of related health costs, absenteeism and disability, low back pain is a substantial economic burden to society. (2, 5)

The therapeutic guidelines on low back pain focus on the continuation of physical activity, as the effectiveness of most therapeutic interventions does not exceed the placebo effect. (1) In addition, the guidelines recommend considering patient preferences in the choice of the therapeutic regimen because contextual factors determine the speed of recovery. (1, 6) Contextual factors include the patient, the physician, and their relationship. (6)

The illness perceptions of a patient, such as avoidance beliefs and fear of the duration of the illness, predict the patient's recovery and their return to work. (7–9) In medical decision-making, little attention is paid to the patient perspective, and even if considered, it is often misinterpreted. (9–11) However, the patient perspective is generally considered essential for medical decision-making, as stated in the Salzburger Statement on shared decision-making (SDM). (12)

SDM is defined as follows: A situation in which the professional and patient share their perspective and jointly decide on a treatment plan. SDM provides the possibility of incorporating patient preferences into clinical decision-making. (13) The philosophy of this concept is that patients will have more autonomy in decisions about their personal health if the doctor-patient relationship shifts from paternalistic to a more equal relationship. (14) Glyn Elwyn operationalised this concept into a three-talk model of shared decision-making. *Team talk* places an emphasis on the need to provide support to patients when they are made aware of choices, and *Option talk* refers to the task of comparing alternatives by using risk communication principles. *Decision talk* refers to the task of arriving at decisions that reflect the informed preferences of patients, guided by the experience and expertise of health professionals. In this broadly accepted model, patients are informed about the decision process and the pros and cons of treatment options.(15)

Since the introduction of SDM in clinical care, research has focused on the process of SDM implementation and its effect on clinical outcomes. (16) At present, the findings related to clinical outcomes are scarce and unconvincing. (16)

For patients with low back pain, SDM could improve the prognosis if patients were more adherent to treatment, as the expectation of a favourable outcome is incorporated into the treatment decisions. (17)

It has been empirically proven that the positive outcome expectations of the patient benefit the health status of the patient, and the reinforcement of these treatment expectations could endorse these effects. (18, 19)

Although widely advocated in guidelines, the effectiveness of SDM in the management of low back pain has not been evaluated in general practice. (20)

Therefore, we conducted a large randomised controlled trial among primary care patients with non-chronic low back pain in the Netherlands and report the effectiveness of SDM followed by positive reinforcement of the therapeutic choice (SDM&PR) on recovery and healthcare consumption.

METHODS

Aim

The aim is to assess the effectiveness of shared decision-making followed by a positive reinforcement of the chosen therapy (SDM&PR) on patient-related clinical outcomes in patients with non-chronic low back pain in general practice.

Design and setting

A cluster-randomised controlled trial was performed in the practices of 68 general practitioners (GPs) in the academic primary care network around Utrecht in the Netherlands.

Participants

GPs were recruited between August 2009 and May 2011. Each participating GP was requested to include ten patients with non-chronic aspecific low back pain.

The inclusion criteria were as follows:

1. between 18 and 65 years of age, and
2. in consultation for a new episode of non-chronic aspecific low back pain (as defined by the guidelines of the Dutch College of General Practitioners and the Cochrane Collaboration). (2, 21)

The exclusion criteria were as follows:

1. duration of low back pain longer than three months,
2. any previous episode of low back pain within the three months prior to the onset of the present episode,
3. pregnancy, and
4. insufficient mastery of the Dutch language.

Because the causes and pathophysiology of low back pain might be different in patients younger than 18 or older than 65 years, those who are pregnant or in those with a longer disease duration, we excluded these patients. (1, 21)

Randomisation, data collection and blinding

GPs were randomly assigned to the usual care (UC) group or the intervention (IV) group immediately after consenting to participate in the trial. Randomisation was done by research staff members who were not otherwise in the research project. Allocation was blinded using allocation cards in sealed envelopes in an initial block of 40 followed by blocks of ten envelopes. GPs in the control group were kept unaware of the communicative techniques that were trained. Auxiliary staff members recruited the patients. Patients and auxiliary practice staff members were not informed about the allocation of the GP or about the communicative techniques in the training programme. Auxiliary practice staff members collected questionnaires from the patients after inclusion. A follow-up questionnaire with a pre-paid envelope was given to each patient with instructions on when to complete it and send it to the research team. Patients were reminded to send the questionnaires two, six, twelve and twenty-six weeks after the consultation by email or phone just before the correct time and, if necessary, again two weeks later.

Intervention

GPs in the intervention group were trained to perform SDM&PR during their consultations with the included patients. SDM followed the following process steps: inform the patient about therapeutic options, discuss the patient's preferences, concerns and expectations, confirm the patient's understanding, assess the patient's preferred level of involvement in decision-making and finally make a joint decision about the optimal therapeutic regimen. GPs were trained to positively reinforce treatment outcomes after SDM.

Training

GPs in the intervention group received two training sessions of two and a half hours. Training sessions were held in small groups of approximately three to five participants and were given by a peer GP with expertise in training SDM skills (AS).

The training was based on the learning principles of Kolb and the behavioural process elements of Elwyn. (22) To support SDM performance during consultations, the participating GPs received a desktop card summarising all consecutive process elements for SDM and a decision aid specifically developed for this trial according to the International Patient Decision Aids Standards (IPDAS)-guidelines. (Additional file 1 Appendix 1) (23, 24) Finally, they received individual feedback on their SDM performance based on observation by the trainer (AS) of video-recordings of the consultation of each included patient. Details of the training are reported elsewhere. (23) The fidelity of the intervention was checked by measuring behavioural changes and consultation duration differences between the intervention and the control group using the OPTION instrument on video-recorded consultations. (23)

Control group

In the control group, the GPs provided the usual standard of care. Although routine management was not predefined in the instructions for the study, GPs in the Netherlands are reported to follow the professional guidelines on low back pain in 70% of patients. (25) Discussion of the favourable prognosis of non-chronic low back pain is part of the suggested management in the guideline, but SDM is not. (2)

Outcome

The primary outcome was the difference between the intervention and the control group in the course of functional disability during the six-month follow-up. Functional disability was measured daily during the first two weeks and at two, six, twelve and twenty-six weeks after the first consultation.

As secondary outcomes, we assessed the difference in functional disability at the time of each of the separate measurements (2, 6, 12 and 26 weeks), the difference in severity of back pain and the percentage of patients with adequate relief on separate measurement dates and at the end of the study.

As indicators of economic effect, we evaluated the differences in absenteeism and healthcare consumption between groups over the complete study period and on the separate measurement dates.

To be able to test for potential confounding, we measured illness perceptions at the baseline. To check the fidelity of the intervention, we questioned patients after the consultation about the level of involvement in decision-making.

Measurements and instruments

Functional disability was assessed by the Dutch version of the Roland Morris disability questionnaire (RMD). This validated questionnaire contains 24 closed questions about restrictions in daily activities during the previous day. The score is the total number of positive answers. (26) Pain severity was quantified by the validated continuous visual analogue scale (VAS), in which the patients indicate the level of pain during the past week on a continuous line that ranges from zero (no pain) to ten cm (the most terrible pain I can imagine), with outcomes measured in mm. (27) Adequate relief of pain was assessed with one closed question referring to the recovery experienced since the previous questionnaire and was expressed as the percentage of patients with adequate relief in each group.

Absenteeism was measured by the response to a question referring to the time before the baseline or since the previous questionnaire: ‘Because of my back pain, I refrained from work (absenteeism),’ yes/no/not applicable (expressed as a percentage of patients), followed by an inquiry about the number of days of sick leave.

Healthcare consumption data were derived from patient questionnaires by counting follow-up contacts, via telephone or at the practice, and expressed as the mean number of contacts per patient since the previous questionnaire. Illness perceptions were assessed by the Dutch version of the abbreviated illness perception questionnaire

(IPQ). (28) This instrument measures eight separate dimensions of perceptions about low back pain using a scale rating from zero to ten. This instrument has been proven to be valid. (29)

The actual level of shared decision-making, as experienced by the patient, was evaluated by their response to one simple question immediately after the consultation: 'Were you involved in decision-making?' Mean scores were calculated from a range of one to four points corresponding to the answers 'no,' 'mostly no,' mostly yes' or 'yes'.

The observed effects of the training were reported in a separate article on the evaluation of the training. (23)

The primary outcome (RMD) and the VAS were assessed daily during the first 14 days by a diary and at two, six, twelve and twenty-six weeks after consultation by questionnaires. All other secondary outcomes were assessed at two, six, twelve and twenty-six weeks. (Fig. 1) Baseline measurements, potential confounders and the manipulation check were assessed through questionnaires completed by all patients before and immediately after the consultation.

Sample size

To reach a minimum standardised difference of 0.3 in the primary outcome between the intervention and control groups, which is more than 1 point on RMD scores with a standard deviation of 5, using a beta of 0.80 and an alpha of 0.05, 352 patients would be required. (26) As we randomised at the level of the GP but measured patient outcomes, we controlled for clustering effects. Based on clustering effects reported in earlier trials, we applied an intra-class correlation of 0.03. (30) Presuming a 10% dropout rate, we calculated that 426 patients should be included by 60 GPs.

Statistical analysis

Differences in baseline characteristics between dropouts and patients who completed the follow-up were tested for significance using a t-test for continuous variables and a X2-test for dichotomous and categorical variables.

The effect of the intervention on the separate measurement dates at two, six, twelve and twenty-six weeks was tested univariately. In multivariate analysis, differences in primary and secondary outcomes were estimated with a mixed model corrected for potential confounders: age, sex, educational level and the corresponding baseline value of the outcome. A random intercept was included for clustering at the level of the GP, and a random intercept and a random effect for time at the patient level were included to incorporate the effect over time. All analysis was performed on an intention-to-treat basis.

The potential confounding effect of each of the illness perception dimensions was assessed with mixed models, with restrictions at 12 weeks as the outcome variable and correction for all confounders.

Mixed models are robust for individual patients with missing follow-up measurements. In the analysis, we originally included the baseline measurement of

the corresponding outcome as a covariate. Consequently, any measurement of any patient with a missing baseline measurement would be excluded from the analysis, thus reducing power and potentially introducing bias. We therefore decided to use multiple imputation to impute baseline variables and missing confounders measured at the baseline. (31, 32)

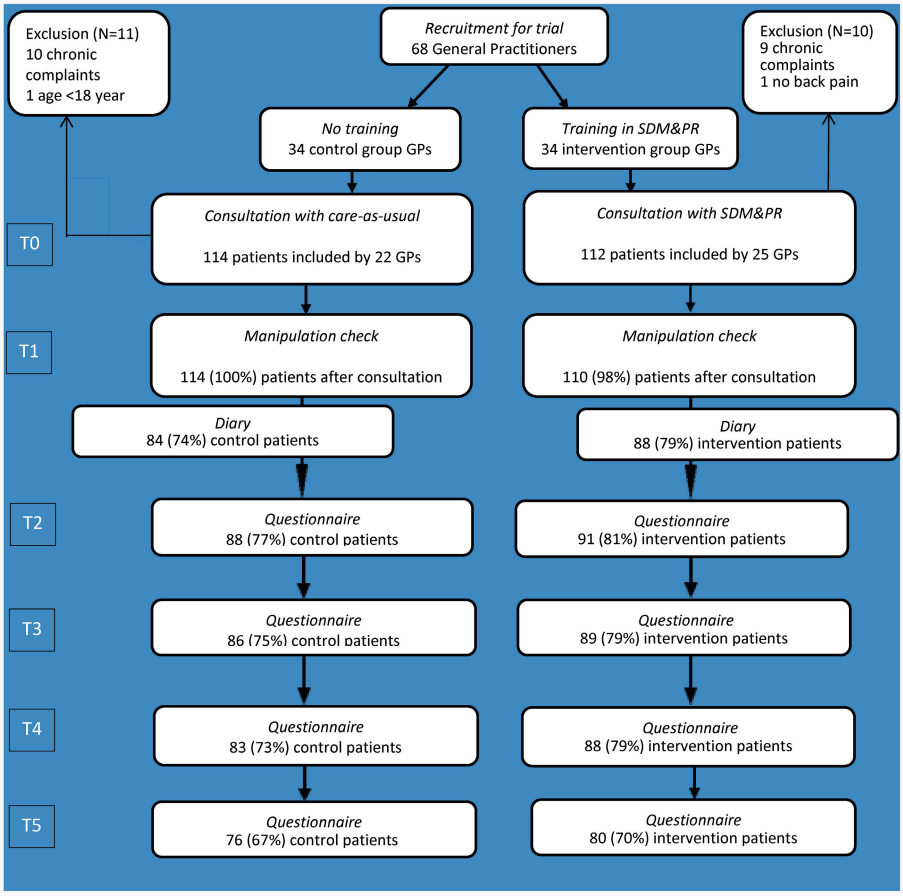


Figure 1 Flow chart of the participants in different phases through the trial.

T0 = directly before the consultation; T1 = directly after the consultation; T2 = 2 weeks after consultation; T3 = 6 weeks after consultation; T4 = 12 weeks after consultation; T5 = 26 weeks after consultation

Age, sex, educational level, absenteeism, all illness perceptions, the treatment allocation and all baseline measurements of primary and secondary outcomes were included in the multiple imputation. The numbers of imputed missing variables per baseline variable are described in Table 1. Five imputed datasets were created. Univariate and multivariate analyses were performed on each imputation; the results reported here were combined with Rubin's rule. (33)

RESULTS

Participants

Sixty-eight GPs agreed to participate and were randomised to the intervention (n = 34) or the control group (n = 34). GPs in the intervention group did not differ from control GPs with regard to sex, age, professional age, number of included patients or percentage of GP trainers per group.

Between January 2010 and January 2012, forty-seven GPs included 247 patients (range 1–10). Twenty-one of these patients did not meet the inclusion criteria because they did not align with the definition of non-chronic low back pain (n = 19), did not have back complaints (n = 1) or were younger than 18 years old (n = 1). Ultimately, 114 patients in the control group and 112 in the intervention group were included in the analysis. (Fig. 1)

Patients in the intervention and control groups were comparable across most baseline measurements. (Table 1 and Additional file 2 Appendix 2 for the imputed dataset) Patients in the intervention group reported more absenteeism from work due to their low back pain, and they more frequently expected their pain to last longer than did the controls.

During the follow-up, 76 (67%) of the control patients and 80 (70%) of the intervention patients completed all questionnaires. Overall, 71 (62%) patients in the control group and 75 (66%) patients in the intervention group completed the diary and all questionnaires. (Fig. 1) Patients who did not complete the follow-up were more frequently of non-Dutch origin (15% non-Dutch natives in dropouts versus 5% non-Dutch natives in the analysed group; p = 0.017) and were younger (a mean age of 39.2 years for dropouts versus 47.4 years for the analysed group; p = 0.000). They did not differ significantly in other baseline measurements.

Table 1. Baseline demographic and clinical characteristics of patients *in the complete* dataset. Continuous variable values are represented as means (standard deviation). Dichotomous variable values are represented as numbers (percentages)

	intervention group (*n* = 112)	control group (*n* = 114)
PATIENT CHARACTERISTICS		
mean age (years)	45.4 (13.2)	44.3 (14.4)
male†	52 (47%)	55 (49%)
Dutch origin‡	97 (91%)	103 (93%)
educational level‡		
primary only	15 (14%)	19 (17%)
secondary	56 (52%)	53 (48%)
college, university	36 (34%)	39 (35%)
employed§	73 (70%)	71 (70%)
BASELINE CLINICAL CHARACTERISTICS		
functional disability score (RMD range 0–24) (primary measure)¶	10.7 (5.0)	10.3 (5.2)
pain severity at baseline (VAS scale 0–100 mm)||	48.6 (16.0)	46.7 (16.7)
absenteeism (yes/no)||	39 (35%)	19 (20%)
illness perception dimensions (IPQ range 0–10)§		
consequences	6.3 (2.3)	6.1 (2.5)
timeline	4.2 (2.8)	3.5 (2.4)
personal control	5.0 (2.2)	5.4 (2.1)
treatment control	6.6 (1.9)	6.9 (1.9)
identity	6.9 (1.6)	7.2 (1.6)
concerns	4.5 (2.5)	4.8 (2.6)
illness comprehensibility	6.0 (2.3)	6.1 (2.3)
emotional response	5.0 (2.5)	5.2 (2.6)

RMD = Roland Morris disability questionnaire (a higher score indicates a more favourable outcome). VAS = visual analogue scale combined score of low back pain, leg pain and both (a lower score indicates a more favourable outcome). IPQ = Illness Perception Questionnaire. ¶ *n* = 3 missing. || *n* = 34 missing. † *n* = 4 missing. ‡ *n* = 8 missing. § *n* = 19 missing.

Intervention effect

The mean disability score among the patients in the intervention and the control groups declined to 4.1 (SDM&PR group) and 4.3 (control group) after 2 weeks (difference 0.2; p-value 0.789), 2.1 (SDM&PR) and 2.3 (controls) after 12 weeks (difference 0.2; p-value 0.720) and 2.0 for both groups after 26 weeks (difference 0.0; p-value 0.949). (Table 2 and Fig. 2)

The mean pain score in the two groups was 18.9 (SDM&PR) and 20.3 (controls) after 2 weeks (difference 1.4; p-value 0.675), 14.2 (SDM&PR) and 12.4 (controls) after 12 weeks (difference 1.8; p-value 0.577) and 13.6 (SDM&PR) and 16.3 (controls) after 26 weeks (difference 2.7; p-value 0.385). The percentage of patients with adequate relief was 70% (SDM&PR) and 62% (controls) after 2 weeks (p-value 0.888), 49% in both groups after 6 weeks, 69% (SDM&PR) and 62% (controls) after 12 weeks, and 66% (SDM&PR) and 64% (controls) after 26 weeks. The mean number of days of absenteeism and mean healthcare consumption at 2, 6, 12 and 26 weeks did not differ between the two groups. (Table 2)

In the multilevel, multivariate analysis, correcting for baseline differences, patient characteristics and the clustering effect, the mean difference in disability scores between the intervention and control groups during the six-month follow-up was − 0.259 (p-value 0.582) (Table 3). The mean difference in pain score between the two groups in the six-month follow-up was − 2.269 (p-value 0.306). During the follow-up, the two groups did not differ in the percentage of patients with adequate relief, the number of days of absenteeism or in healthcare consumption. (Table 3)

Of the 8 dimensions of illness perception, only consequences (β = 1.24 confidence interval (CI) 0.14–2.35), timeline (β = 1.38 CI 0.27–2.48) and concern (β = 1.45 CI 0.34–2.56) were significantly associated with disability at 12 weeks. However, when the interaction term of each of these three items with the intervention was added to the multivariate model, no significant effect of illness perception on disability at 12 weeks was found.

In both groups, patients reported a substantial degree of involvement in decision-making. However, the patients in the intervention group reported a significantly higher level of patient involvement (2.92 standard deviation (SD) 1.21; range 1–4) than the controls (2.44 SD 1.23) (difference 0.48; *p*-value 0.005).

When studying the fidelity of the intervention, we measured significant differences in the SDM behaviour in favour of the intervention group and a mean duration of the consultation of 16 min for the intervention group versus 13 min for the control group. (23)

Table 2. Univariate mean score per group in primary and secondary outcomes in the imputed dataset without correction for clustering. Continuous variable values are represented in means (standard deviation). Dichotomous variable values are represented as numbers (percentage).

	mean score at 2 weeks			mean score at 12 weeks			mean score at 26 weeks		
	intervention group	control group	*p*-value	intervention group	control group	*p*-value	intervention group	control group	*p*-value
clinical parameters									
disability (RMD range 0–24)	4.1 (5.3)	4.3 (4.8)	0.789	2.1 (4.0)	2.3 (3.7)	0.720	2.0 (3.7)	2.0 (3.6)	0.949
pain (VAS scale 0–100 mm)	18.9 (21.7)	20.3 (20.9)	0.675	14.2 (22.6)	12.4 (17.5)	0.577	13.6 (17.3)	16.3 (21.2)	0.385
adequate relief (yes/no)†	70 (81%)	62 (81%)	0.888	45 (69%)	38 (62%)	0.416	35 (66%)	32 (64%)	0.830
societal impact									
absenteeism (days)	1.47 (3.35)	2.05 (4.03)	0.359	0.93 (5.65)	0.800 (3.63)	0.888			*
absenteeism (yes/no)‡	18 (28%)	20 (27%)	0.924	3 (5%)	4 (8%)	0.552	7 (11%)	6 (14%)	0.650
healthcare consumption									
telephone consultations (per patient)	0.35 (0.71)	0.29 (0.60)	0.556	1.18 (0.50)	1.18 (0.51)	0.672			*
practice consultations (per patient)	0.21 (0.51)	0.11 (0.39)	0.134	1.12 (0.43)	1.15 (0.52)	0.965	1.11 (0.45)	1.10 (0.38)	0.914

RMD = Roland Morris disability questionnaire (a lower score indicates a more favourable outcome). VAS = visual analogue scale (a lower score indicates a more favourable outcome). Mean score of low back pain, leg pain and both. IPQ = illness perception questionnaire. * = Cannot be computed because the group is zero.

Table 3. Difference in mean scores between the control and intervention groups *in the imputed dataset* during the six-month follow-up

	univariate analysis*			multivariate analysis	
	mean difference/ rate ratio	confidence interval	*p*-value	mean difference/ rate ratio	*p*-value
ENDPOINT					
disability (RMD range 0–24) ¶	−0.233	−1.258 to 0.791	0.655	−0.259	0.582
secondary outcomes					
pain (VAS scale 0–100 mm) ¶	−1.120	−6.133 to 3.893	0.662	−2.269	0.306
adequate relief (yes/no) †	1.118	0.510–1.567	0.696	1.119	0.730
absenteeism (in days) ‡	1.032	0.927–1.338	0.249	0.332	0.506
healthcare consumption					
telephone consultations (number per patient) §	1.0142	1.001–1.018	0.845	1.020	0.789
practice consultations (number per patient) §	1.0143	0.880–1.169	0.845	1.067	0.245

RMD = Roland Morris disability questionnaire (a lower score indicates a more favourable outcome). VAS = visual analogue scale (a lower score indicates a more favourable outcome) mean score of low back pain, leg pain and both. * = corrected for clustering effect. ¶ = mean difference between control and intervention groups over 26 weeks. † = odds ratio without baseline correction. ‡ = rate ratio in the multilevel model corrected for dichotomous baseline value. § = rate ratio without baseline correction.

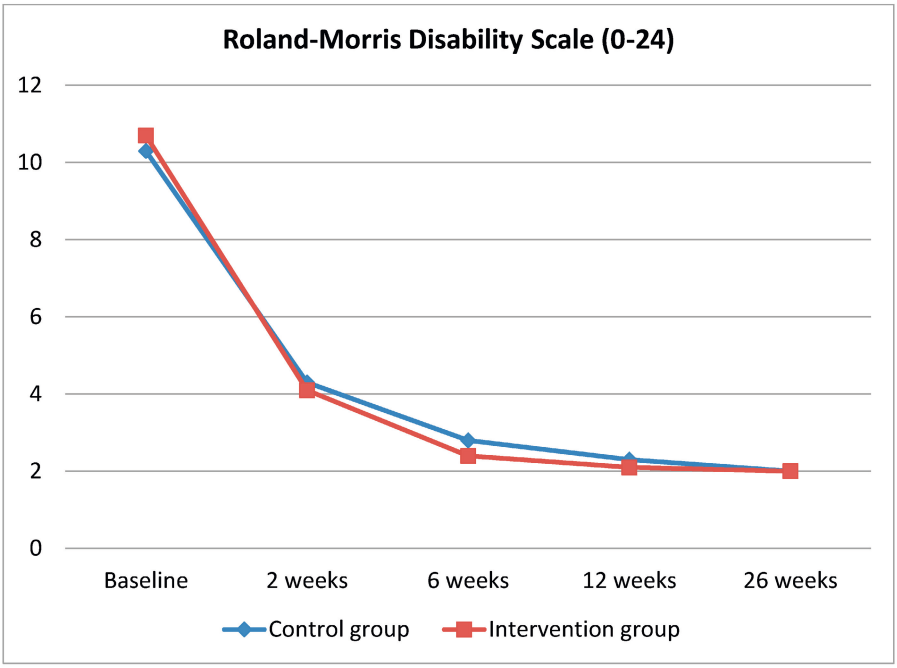


Figure 2. Patients experiencing disabilities after 2, 6, 12 and 26 weeks. Observed mean scores in experienced disabilities (measured by RMD scale) per group without correction for clustering between baseline and 26 weeks in time. At baseline, 226 patients, at 2 weeks, 179 patients, at 6 weeks, 175 patients, at 12 weeks, 171 patients and at 26 weeks, 156 patients. Minimal clinically important change = 3.5 (34)

DISCUSSION

This training of general practitioners in SDM&PR did not improve the symptom recovery of patients with non-chronic low back pain in primary care, even though the GPs effectively involved patients in the choice of treatment after the training. At no point in the follow-up did the mean disability or pain score of the patients whose GP was trained in SDM&PR differ from those patients whose GP provided the usual standard of care (Figure 2). Patients in both groups reported that pain and physical limitations gradually declined and returned to a normal population level at 26 weeks. (34)

The comparable clinical discourse in the two groups was also reflected in pain-related absenteeism from work and in healthcare consumption during the follow-up.

Patients who attributed much importance to the consequences of their back pain, those who expected the pain to last long and those who had many concerns about the pain, had a poorer prognosis for symptom recovery. However, the prognosis was independent of the performance of SDM&PR by the GPs.

Strengths of the study

Most research on SDM thus far has focused on process outcomes and not on patient-related clinical outcomes (20). We performed a randomised controlled trial among patients with non-chronic low back pain recruited in daily primary care and evaluated the effectiveness of SDM on relevant clinical outcomes. Based on current knowledge, we constructed a multifaceted intervention and training programme that was grounded in a theoretical concept of SDM, involving both participants in the decision process. (20)

Participating GPs were well trained, and the positive SDM performance during consultations after the training was acknowledged by the patients. (23) We used a mixed model analysis because these are robust for individual patients with missing follow-up measurements under the assumption of missing values completely at random or missing values at random dependent on another variable included in the mixed model. As in most studies, the correctness of this assumption of 'random missing' cannot be proven in our study. In our view, however, it is very unlikely that a treatment effect was not observed due to the multiple imputation procedure because the complete case analyses confirmed the lack of a treatment effect. (Additional file 3 Appendix 3)

Limitations of the study

The patient recruitment met only half (53%) of the pre-set sample size. Participating GPs experienced problems that hampered recruitment, possibly due to unforeseen changes in the healthcare system, such as the introduction of direct access physiotherapy. However, the fact that the results of the patients in the intervention and control groups did not differ in any of the outcome measures at any moment in time demonstrates in our view the consistency of the results, and even if differences were demonstrated at the pre-set sample size, these differences would have been small and of questionable clinical relevance.

Dropout rates of GPs and patients are similar in the intervention and control groups and are in line with other studies on patients suffering from non-chronic low back pain in primary care in the Netherlands. (35) We cannot think of any reason why the intervention should have influenced the dropout rate of patients, but we estimate that dropout is not related to the intervention but is rather related to the complex disease course of non-chronic low back pain and the mismatch between patient expectations and the professional's management.

Although we have observed significant differences in the perception of SDM between the experimental groups, we believe that the difference should be evaluated in the context of treatment fidelity. In a recent publication evaluating the intervention from an observational perspective, we found significant differences in the use of physical examination and in the consultation duration between groups. (23) Moreover, we question whether the patient perspective was sufficiently considered to incorporate the patient's positive expectations into the actual decision despite significant differences in the SDM behaviour of GPs between the groups.

Unfortunately, as with many other studies on the effects of SDM on health outcomes, the effects are too small to allow conclusions about the impact of the separate communicative techniques, SDM and positive reinforcement of treatment expectations. Theoretically, we expected a positive interaction between the two techniques. The consistent pattern of very small potential positive treatment effects of the combined intervention above the usual standard of care could be explained by a stronger positive effect of one technique counteracted by the effect of the other. (Fig. 2) For instance, in a study by a physiotherapist on the effects of SDM on the prognosis of low back pain, even negative expectations of patients were suggested to be responsible for poorer health outcomes after SDM than the usual standard of care. (36) Conversely, in a study evaluating the placebo effect on chronic low back pain, positive treatment expectations and a supportive environment were considered responsible for short-term relief from complaints. (16, 23)

Studies on the effects of learned communicative strategies frequently face problems with blinding. We downsized the risk of non-blinding by sorting patients per GP, by recruiting patients via auxiliary staff members unaware of the allocation, and by not providing details of the trained communicative strategy to control GPs, auxiliary staff members and patients.

We did not perform a full health economic assessment but restricted the economic impact analysis to measuring the absenteeism of workers. However, in a detailed cost-effectiveness analysis, the reported 20% difference in the duration of the consultation time should be considered. (23)

Possible explanations

As in other studies on interventions for non-chronic low back pain, we did not find substantial or significant effects. (37) This finding could be attributed to different factors. Such as an excessively diverse study population in the duration of the complaints or in patient characteristics. (36–38)

Although we think that the risk of contamination was limited, GPs in the control group may have incorporated SDM in their consultations as well. This might be reflected in the fact that the mean score for the question of whether patients felt involved in the decision-making was between ‘mostly yes’ and ‘mostly no’. The difference in the results of the intervention group, where the mean score was ‘mostly yes’, was significant but limited. However, the observational study demonstrates low levels of SDM in both groups despite the significant effect of training in SDM behaviour. (11, 23, 39, 40)

Because most patients quickly recover from their back pain, the intervention simply might not have had sufficient discriminative content above the spontaneous course. In their review of psychosocial interventions for non-chronic low back pain in primary care, Ramond et al. (38) advise the integration of several psychosocial factors with multicomponent interventions to overcome this problem.

Contextual factors play an important role in the symptom perception, prognosis and recovery of low back pain. (1, 4) We identified three subgroups of patients with a

poorer prognosis for symptom recovery in the analysis of the effect of each of the illness perception dimensions on the restrictions at twelve weeks. Patients with more negative illness perceptions or with a longer duration of complaints before they contact their GP might be better helped by more positive treatment expectations, but unfortunately, our dataset did not allow subgroup analysis on the effect of the intervention for particular patient characteristics. (38) Differences between the contextual factors of patients in the intervention and control groups may have influenced the results. Although the patients were not randomised, we have no indication that the recruitment to the intervention and control groups resulted in selection bias. (18, 28, 41, 42)

Although the GPs were extensively trained before participating in the intervention group and the patients recognised SDM during the actual consultations, we question, based on the results of the evaluation of the training, whether the training did result in adequate SDM performance. (40, 43) Trained GPs became more aware of the need to better inform patients about treatment options and to incorporate patients' expectations during the intake phase of the consultation, but they persisted in providing a paternalistic, guideline-oriented choice.

In a review on the effects of the implementation of SDM in clinical encounters measured by an external observer, Couët et al. reported similar training effects and noticed that only incidentally clinical management is adjusted to patient preferences. (40) In the evaluation of the training, we confirm this observation and conclude that patient preferences were insufficiently considered in the actual decision-making to incorporate the patient's positive treatment expectations into the treatment choice. (23) When patient preferences are not reflected in treatment choices, the impact of positive reinforcement of the therapeutic plan on patient recovery will diminish.

So far, task-oriented issues, such as performing process steps and information exchange, are emphasised in the implementation of SDM. (20) However, the effects of knowledge transfer on proportional understanding are questionable, and the effects on recovery are unclear. (20, 43)

Future research on the involvement of patients in treatment decisions should therefore focus more on professional attitude and equality in the patient professional relationship as a condition for successful SDM.

CONCLUSION

Training of GPs in the application of SDM&PR during consultations with patients with non-chronic low back pain did not significantly improve clinical recovery. Although it may have improved the 'knowledge and rationalise expectations' of the patients, this did not lead to less functional impairment, shorter pain duration or less absenteeism from work than routine practice. Most patients recovered from their low back pain within 12 weeks, and this positive effect was persistent at the 26-week follow-up, which confirms the benign natural course of low back pain as reported in the literature. A potential small positive effect of either SDM or positive reinforcement of treatment

expectations cannot be excluded. As the prognosis of low back pain is predominantly determined by psychosocial factors, we suggest that further research on the positive health effects of communicative techniques should focus on a more patient-oriented approach, combined with the reinforcement of positive recovery expectations, than on task-oriented techniques such as SDM.

ABBREVIATIONS

AR: Adequate relief;
AS: Ariëtte Sanders;
CI: Confidence interval;
GP: General practitioner;
IPDAS: International Patient Decision Aids Standards;
IPQ: Illness perception questionnaire;
IV: Intervention;
NTR: The Netherlands National Trial Register;
RMD: Roland Morris disability questionnaire;
SD: Standard deviation;
SDM: Shared decision-making;
SDM&PR: Shared decision-making followed by a positive reinforcement of the chosen therapy;
UC: Usual care;
VAS: Visual analogue scale

ACKNOWLEDGEMENTS

The authors wish to thank all patients and GPs who voluntarily participated in the trial, William Verheul for developing the concept, Peter Zuithoff and Peter Spreeuwenberg for statistical support and Emily Fick and all students who supported the sample acquisition and sequence data processing.

AVAILABILITY OF DATA AND MATERIALS

Data sharing is possible; researchers should send a request and motivation for (parts of) the database or remaining sample material to the first author. A committee will then decide whether the data or samples can be used for the specific research purpose.

AUTHORS' CONTRIBUTIONS

All authors were involved in critical review of the manuscript and have seen and approved the final version. Specific contributions are as follows: study conception and design: JB and AS; analysis of epidemiological and sequence data: PV, TM, AS; drafting

the manuscript: NW, JB, PV and AS. All authors had full access to all the study data and take responsibility for the integrity of the data and the accuracy of the data analysis. All authors read and approved the final version of the manuscript.

ETHICS APPROVAL AND CONSENT TO PARTICIPATE

The study protocol for the trial was assessed by the Ethics Committee of the University Medical Centre of Utrecht and exempted from full assessment because participants are not subject to medical proceedings or behavioural changes as referred to in the definition of medical scientific research in the Medical Research Involving Human Subjects Act (article 1b). Patients with back pain were individually informed by the medical staff about the trial and the consequences of participation, including the video-recording. They signed a written informed consent form in the waiting room. Before the recording started, the GP checked the permission together with the patient.
The manuscript has been drafted according to the CONSORT guidelines.

COMPETING INTERESTS

The authors declare that they have no competing interests to disclose.

REFERENCES

1) van Tulder M, Becker A, Bekkering T, Breen A, del Real MT, Hutchinson A, et al. Chapter 3. European guidelines for the management of acute nonspecific low back pain in primary care. Eur Spine J. 2006;15(Suppl 2): S169–91.
2) Bons SCS, Borg MAJP, Van den Donk M, et al. NHG-Standaard Aspecifieke lagerugpijn (Tweede herziening). In: Richtlijnen en praktijk. Het Nederlands Huisartsen Genootschap. 2017. https://www.nhg.org/standaarden/volledig/ nhg-standaard-aspecifieke-lagerugpijn. Accessed 20 June 2018.
3) Dionne CE, Dunn KM, Croft PR, Nachemson AL, Buchbinder R, Walker BF, et al. A consensus approach toward the standardisation of back pain definitions for use in prevalence studies. Spine. 2008;33:95–103.
4) Krismer M, van Tulder M. Strategies for prevention and management of musculoskeletal conditions. Low back pain (non-specific). Best Pract Res Clin Rheumatol. 2007;21:77–91.
5) Dagenais S, Caro J, Haldeman S. A systematic review of low back pain cost of illness studies in the United States and internationally. Spine J. 2008;8:8–20.
6) Di Blasi Z, Harkness E, Ernst E, Georgiou A, Kleijnen J. Influence of context effects on health outcomes: a systematic review. Lancet. 2001;357:757–62.
7) Iles RA, Davidson M, Taylor NF. Psychosocial predictors of failure to return to work in non-chronic non-specific low back pain: a systematic review. Occup Environ Med. 2008;65:507–17.
8) Iles RA, Davidson M, Taylor NF, O'Halloran P. Systematic review of the ability of recovery expectations to predict outcomes in non-chronic non-specific low back pain. J Occup Rehabil. 2009;19:25–40.
9) van Tol-Geerdink JJ, Stalmeier PF, van Lin EN, Schimmel EC, Huizenga H, van Daal WA, et al. Do patients with localized prostate cancer treatment really want more aggressive treatment? J Clin Oncol. 2006;24:4581–6.
10) Bensing J. Bridging the gap. The separate worlds of evidence-based medicine and patient-centered medicine Patient Educ Couns. 2000;39:17–25.
11) Stalmeier PFM, van Tol-Geerdink JJ, van Lin ENJT, Schimmel E. Huizenga H, van Daal WA, et al. (the patient chooses for feasibility and effectiveness). NedTijdschrGeneeskd. 2009;153:600–6.
12) Salzburg Global Seminar. Salzburg statement on shared decision making. BMJ. 2011;342:d1745.
13) Edwards A, Elwyn G. Inside the black box of shared decision making: distinguishing between the process of involvement and who makes the decision. Health Expect. 2006;9:307–20.
14) Stiggelbout AM, Van der Weijden T, De Wit MP, et al. Shared decision making: really putting patients at the Centre of healthcare. BMJ. 2012;344:e256.
15) Elwyn G, Durand MA, Song J, et al. A three-talk model for shared decision making: multistage consultation process. BMJ. 2017;359:j4891.
16) Légaré F, Stacey D, Turcotte S, Cossi MJ, Kryworuchko J, Graham ID, et al. Interventions for improving the adoption of shared decision making by healthcare professionals. Cochrane Database Syst Rev. 2014;9:CD006732.
17) Stacey D, Légaré F, Lewis K, Barry MJ, Bennett CL, Eden KB, et al. Decision aids for people facing health treatment or screening decisions. Cochrane Database Syst Rev. 2011;10:CD001431.
18) Benedetti F, Amanzio M. Mechanisms of the placebo response. Pulm Pharmacol Ther. 2013;26:520–3.
19) Thomas KB. General practice consultations: is there any point in being positive? Br Med J (Clin Res Ed). 1987;294:1200–2.

20) Légaré F, Ratté S, Stacey D, Kryworuchko J, Gravel K, Graham ID. Interventions for improving the adoption of shared decision making by healthcare professionals. Cochrane Database Syst Rev. 2010;5:CD006732.
21) Saragiotto BT, Machado GC, Ferreira ML, Pinheiro MB, Abdel Shaheed C, Maher CG. Paracetamol for low back pain. Cochrane Database Syst Rev. 2016;6:CD012230.
22) Elwyn G, Hutchings H, Edwards A, Rapport F, Wensing M, Cheung WY, Grol R. The OPTION scale: measuring the extent that clinicians involve patients in decision-making tasks. Health Expect. 2005;8:34–42.
23) Sanders AR, Bensing JM, Essed MA, Magnée T, de Wit NJ, Verhaak PF. Does training general practitioners result in more shared decision making during consultations? Patient Educ Couns. 2017;100:563–74.
24) Elwyn G, O'Connor AM, Bennett C, Newcombe RG, Politi M, Durand MA, et al. Assessing the quality of decision support technologies using the international patient decision aid standards instrument (IPDASi). PLoS One. 2009;4:e4705.
25) Schers H, Wensing M, Huijsmans Z, van Tulder M, Grol R. Implementation barriers for general practice guidelines on low back pain a qualitative study. Spine (Phila Pa 1976). 2001;26:E348–53.
26) Brouwer S, Kuijer W, Dijkstra PU, Göeken LN, Groothoff JW, Geertzen JH. Reliability and stability of the Roland Morris disability questionnaire: intra class correlation and limits of agreement. Disabil Rehabil. 2004;26:162–5.
27) Ostelo RW, de Vet HC. Clinically important outcomes in low back pain. Best Pract Res Clin Rheumatol. 2005;19:593–607.
28) Foster NE, Bishop A, Thomas E, Main C, Horne R, Weinman J, et al. Illness perceptions of low back pain patients in primary care: what are they, do they change and are they associated with outcome? Pain. 2008;136:177–87.
29) Broadbent E, Petrie KJ, Main J, Weinman J. The brief illness perception questionnaire. J Psychosom Res. 2006;60:631–7.
30) Jellema P, van der Windt DA, van der Horst HE, Blankenstein AH, Bouter LM, Stalman WA. Why is a treatment aimed at psychosocial factors not effective in patients with (sub)acute low back pain? Pain. 2005;118:350–9.
31) Donders AR, van der Heijden GJ, Stijnen T, Moons KG. A gentle introduction to imputation of missing values. J Clin Epidemiol. 2006;59:1087–91.
32) Sterne JA, 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:b2393.
33) Rubin LH, Witkiewitz K, Andre JS, Reilly S. Methods for handling missing data in the behavioral neurosciences: don't throw the baby rat out with the bath water. J Undergrad Neurosci Educ. 2007;5:A71–7.
34) Kovacs FM, Abraira V, Royuela A, Corcoll J, Alegre L, Cano A, et al. Minimal clinically important change for pain intensity and disability in patients with nonspecific low back pain. Spine (Phila Pa 1976). 2007;32:2915–20.
35) Jellema P, van der Windt DA, van der Horst HE, Twisk JW, Stalman WA, Bouter LM. Should treatment of (sub)acute low back pain be aimed at psychosocial prognostic factors? Cluster randomized clinical trial in general practice. BMJ. 2005;331:84.
36) Gysels M, Richardson A, Higginson IJ. Communication training for health professionals who care for patients with cancer: a systematic review of effectiveness. Support Care Cancer. 2004;12:692–700.
37) Pengel HM, Maher CG, Refshauge KM. Systematic review of conservative interventions for subacute low back pain. Clin Rehabil. 2002;16:811–20.

38) Ramond-Roquin A, Bouton C, Gobin-Tempereau AS, Airagnes G, Richard I, Roquelaure Y, et al. Interventions focusing on psychosocial risk factors for patients with non-chronic low back pain in primary care–a systematic review. Fam Pract. 2014;31:379–88.
39) Schers H, Braspenning J, Drijver R, Wensing M, Grol R. Low back pain in general practice: reported management and reasons for not adhering to the guidelines in the Netherlands. Br J Gen Pract. 2000;50:640–4.
40) Couët N, Desroches S, Robitaille H, Vaillancourt H, Leblanc A, Turcotte S, et al. Assessments of the extent to which health-care providers involve patients in decision making: a systematic review of studies using the OPTION instrument. Health Expect. 2015;18:542–61.
41) Benedetti F, Lanotte M, Lopiano L, Colloca L. When words are painful: unraveling the mechanisms of the nocebo effect. Neuroscience. 2007; 147:260–71.
42) Patel S, Ngunjiri A, Hee SW, Yang Y, Brown S, Friede T, et al. Primum non nocere: shared informed decision making in low back pain–a pilot cluster randomized trial. BMC Musculoskelet Disord. 2014;15:282.
43) Braddock CH 3rd. The emerging importance and relevance of shared decision making to clinical practice. Med Decis Mak. 2010;30:5S–7S.

APPENDIX 1

Desktop tool original Dutch version and translation of the desktop tool

Prompt sheet

1. Voorkeuren van de patiënt voor de therapie gevraagd?
   a. *Wat zou u van mij willen?*
   b. *Wat had u er zelf van tevoren over bedacht?*
   c. *Wat denkt u zelf?*
   d. *Heeft u zelf een idee over de beste aanpak?*

   Were the patients asked about their therapy preferences?
   a. What do you expect of me?
   b. What are your expectations of the therapy?
   c. What are your own thoughts on the therapy?
   d. Do you have ideas regarding the most suitable approach?

2. Verwachtingen van patiënt over therapie bekend
   e. *Welke aanpak kent u?*
   f. *Wat is uw ervaring/verwachting daarover?*
   g. *Waar hoopt u op?*

   Are the patients' expectations about the therapy known?
   e. What approaches are you aware of?
   f. What are your experiences with those approaches?
   g. What are your hopes for this approach?

3. Zorgen van patiënt over probleem uitgevraagd
   h. *Wat dacht u toen u hier naartoe kwam?*
   i. *Waar ligt uw zorg?*
   j. *Bent u ergens bang voor?*

   Were the patients questioned about their worries regarding their medical problems?
   h. What were your thoughts on your way here?
   i. What are your main concerns?
   j. Are you afraid of anything specific regarding the therapy?

4. Keuzehulp uitgelegd en consequenties duidelijk aan patiënt
   k. *Kunt u zo uw keuze maken?*
   l. *Helpt dit u bij het maken van een keuze?*

Were the available options and resulting consequences explained clearly to the patients?

k. Are you able to make your choice with the given information?
l. Does this explanation help in making your choice?

5. Ruimte gegeven aan patiënt om vragen te stellen
   m. *Heeft u hier nog vragen over?*

Have the patients been given an opportunity to ask questions?

m. Do you have any questions?

6. Voorkeur rol patiënt m.b.t. maken van keuze helder
   n. *Zou u zelf een keuze willen maken?*
   o. *Of wilt u dat ik u daarbij help?*

Are the patients' decision-making preferences clear?

n. Would you like to make the decision on your own?
o. Would you like me to help you make a decision?

7. Besluitvorming in gezamenlijk overleg
   p. *Wat zou uw keuze zijn?*
   q. *Wat zou u beslissen?*
   r. *Waaraan geeft u de voorkeur?*

Shared decision-making

p. What would be your choice?
q. What have you decided?
r. Which option do you prefer?

8. Verwachtingen van patiënt positief bekrachtigd
   s. *Naar verwachting zal de pijn op korte termijn overgaan. Als u daarbij goed blijft bewegen zal dat zeker helpen.*
   t. *U zou graag …… willen, dat is een keus die bij u past. Dat zal u zeker helpen bij het herstel.*

Are the patients' expectations positively reinforced?

s. My expectation is that the pain will go away quickly. By moving regularly, you can facilitate recovery.
t. You would like to ……., which is a suitable option for you. That will definitely help your recovery.

APPENDIX 2

Baseline demographic and clinical characteristics of patients in imputed dataset. Continuous variable values are represented as means (standard deviation). Dichotomous variable values are represented as numbers (percentage).

	intervention group (n=112)	control group (n=114)
PATIENTS' CHARACTERISTICS		
mean age (years)	45.4 (13.2)	44.3 (14.4)
male	53 (47%)	56 (49%)
Dutch origins	102 (91%)	106 (93%)
educational level		
primary only	15 (13%)	19 (17%)
secondary	60 (54%)	53 (47%)
college. university	37 (33%)	42 (37%)
employed	79 (71%)	81 (71%)
BASELINE CLINICAL CHARACTERISTICS		
functional disability score (RMD range 0-24) (primary measure)	10.7 (5.0)	10.3 (5.1)
pain severity at baseline (VAS scale 0-100)	48.9 (15.4)	46.4 (15.8)
absenteeism (yes/no)	53 (47%)	36 (32%)
illness perception dimensions (IPQ range 0-10)		
consequences	6.2 (2.3)	6.0 (2.5)
timeline	4.2 (2.7)	3.5 (2.4)
personal control	5.0 (2.2)	5.5 (2.0)
treatment control	6.6 (1.9)	6.9 (1.8)
identity	6.9 (1.6)	7.2 (1.5)
concerns	4.5 (2.5)	4.8 (2.6)
illness comprehensibility	5.9 (2.3)	6.2 (2.2)
emotional response	5.0 (2.5)	5.2 (2.6)

RMD = Roland Morris Disability Questionnaire. VAS = Visual Analogue Scale combined score of low back pain, leg pain and both. IPQ = Illness Perception Questionnaire.

APPENDIX 3

Complete case analysis. Difference in mean scores between the control and intervention group during the six-month follow-up in complete cases (please note the univariate analyses are not influenced by imputation since only baseline variables were imputed in the univariate analysis).

Table 3a. Difference in mean scores between control and the intervention group during six months follow-up on complete cases.

	UNI-VARIATE ANALYSIS*			MULTI VARIATE ANALYSIS	
	mean difference/ rate ratio	confidence interval	p-value	mean difference/ rate ratio	p-value
ENDPOINT					
disability (RMD range 0-24) ¶	-0.233	-1.258-0.791	0.655	-0.495	0.280
Secondary outcomes					
pain (VAS scale 0-10) ¶	-1.120	-6.133-3.893	0.662	-2.297	0.312
adequate relief (yes/no) †	1.118	0.510-1.567	0.696	1.053	0.829
absenteeism (in days) ‡	1.032	0.927-1.338	0.249	0.889	0.769
healthcare consumption					
telephone consultations (number per patient) §	1.0142	1.001-1.018	0.845	1.020	0.780
practice consultations (number per patient) §	1.0143	0.880-1.169	0.845	0.999	0.989

RMD = Roland Morris Disability Questionnaire. VA S= Visual Analogue Scale. mean score of low back pain, leg pain and both. * = corrected for clustering effect. ¶ = mean difference between control and intervention group over 26 weeks. † = odds ratio no baseline correction. ‡ = rate ratio in the multilevel model corrected for dichotomous baseline value. § = rate ratio no baseline correction. Multilevel: corrected for gender, age, educational level and clustering on GP level.

CHAPTER 5

The effect of shared decision-making on recovery from non-chronic low back pain in primary care; a post-hoc analysis from the patient, physician and observer perspectives.

Ariëtte R.J. Sanders
Niek J. de Wit
Nicolaas P.A. Zuithoff
Sandra van Dulmen

BMC Primary Care. 2022 February 02;22-33.

ABSTRACT

Background. Although shared decision-making (SDM) is increasingly accepted in healthcare and has demonstrated merits for several psychological outcomes, the effect on recovery from somatic conditions is still subject to debate. The objective of this study is to measure the effect of SDM on recovery from non-chronic low back pain.

Methods. This study is a post-hoc analysis of data from a cluster-randomised trial that evaluated the effectiveness of SDM on recovery in patients with non-chronic low back pain. In this analysis, we re-evaluate the impact of SDM from three perspectives: that of external observers, participating GPs and participating patients. Recovery was measured with the Visual Analogue Scale (VAS) for pain and with the Roland Morris Disability questionnaire (RMD) and defined as a VAS<30 and an RMD <4. Logistic regression was used to analyse the effect of SDM on recovery at 6 and 26 weeks.

Results. At 26 weeks, 105 (74%) of all 176 included patients had recovered. No significant effect of SDM on recovery at 6 or 26 weeks after the consultation was found when considering SDM from an observer perspective or a patient perspective. From a GP perspective SDM had a significant effect on recovery, but at 26 weeks only, and with the lowest probability of recovery observed at a medium level of GP-perceived SDM.

Conclusions. We found no evidence that SDM as perceived by the patient or by external observation improves recovery from non-chronic low back pain. The long-term recovery may be better for patients in whom the GP perceives SDM during their consultations. Further research should highlight the hierarchy and the relation between the perspectives, which is needed to come to an integral effect evaluation of SDM.

Trial Registration. The Netherlands National Trial Register (NTR) number: NTR1960

BACKGROUND

Low back pain, i.e. back pain without a known specific somatic origin, is among the top ten most frequently presented complaints in primary care. (1) It subsides within two weeks in the majority of patients but can become chronic (>3 months) or frequently recurring (≥3 episodes a year). (1) Worldwide, it is one of the leading causes of disability, with a societal burden primarily incurred through costs related to losses in productivity (93% of total costs). (2,3)

Professional guidelines for low back pain commonly recommend assessing patients' perceptions and informing patients properly about the expected favourable course of the complaints. (1,4) In addition, it has been suggested that patient involvement in the medical decision-making process has a positive effect on the course of illnesses like low back pain. (5)

A process in which the professional and the patient share their perspectives and jointly decide on a treatment plan, is shared decision-making (SDM). (6)

Elwyn identified four key steps in SDM:

1. the professional informs the patient that a decision needs to be made and that the patient's opinion is important;
2. the professional explains the options and their pros and cons;
3. the professional and the patient discuss the patient's preferences, and the professional supports the patient in their deliberation;
4. the professional and patient discuss the patient's wish to make the decision, they make or defer the decision and discuss follow-up. (6)

Although the concept is increasingly accepted in healthcare, the implementation of SDM in clinical practice varies significantly, depending on the perspective of patients, providers or external observers. (7)

Although the benefit of SDM has been demonstrated for several psychological outcomes, such as the patient's emotional status, the effect on recovery from somatic conditions is still subject to debate. (8) In SDM, patients' concerns are explored, which might increase their feelings of being taken seriously and might improve trust in the professional. (9) Moreover, in SDM patients' preferences and outcome expectations are taken into account when jointly deciding on the treatment plan. In a symptom-based illness like low back pain, recovery seems associated with patients' outcome expectations, the attitude of the professional and the relationship with the professional. (10-13) Therefore theoretically, if the outcome expectations of the patient are reinforced in the context of a mutually agreed therapeutic plan within a patient-doctor relationship in which the patient feels supported, one would expect low back pain complaints to subside more quickly compared to traditional care which is professional driven and therapy focussed. (14-16)

We previously reported the results of a randomised controlled trial (RCT) in general practice examining the effects of training GPs in SDM in patients suffering from non-chronic low back pain, with SDM measured from the observer perspective only. (9,17,18) We could not detect a significant benefit from SDM on patient recovery, as objectified with the Roland Morris Disability scale or Visual Analogue Pain scale. (19,20) This could potentially be explained either by a lack of contrast between the intervention and control groups in the level of SDM in practice because of inadequate application of SDM in the intervention group, or by differences between patients and GPs in perceptions of SDM during the consultations in the two study arms. To assess this, further detailing is needed of the association between recovery independently of the allocation and the level of SDM during the consultations from a broader view than just the observer perspective ('treatment fidelity').

We therefore performed a post-hoc analysis of the RCT data and re-evaluated the effectiveness of SDM, as perceived from three perspectives: that of external observers, participating GPs and participating patients.

METHODS

Design

This is a post-hoc analysis of data from a clustered randomised trial that evaluated the effectiveness of SDM among patients with non-chronic low back pain. Details of the design, intervention and overall outcome are described elsewhere. (18,21)

For the current analysis, we merged all patients from the intervention and control groups into a prospective cohort, including follow-up measurements until 26 weeks after the initial consultation. We excluded patients who missed more than 5 outcomes for either restrictions or pain on all 19 time points (questionnaires at baseline, 2, 6, 12 and 26 weeks, and a diary for the first 14 days after the consultation). (Figure 1)

Participants

Adult patients (aged between 18-65 years) who contacted their GP because of a new episode of non-chronic low back pain, as defined by the guideline of the Dutch College of General Practice, were invited to participate in the trial between August 2009 and May 2011. (4) Exclusion criteria were: duration of low back pain longer than three months, recurring backache within three months of the primary episode, pregnancy and insufficient mastery of the Dutch language.

Participating GPs were part of the primary care network around Utrecht (the Netherlands) affiliated with the university. They completed a questionnaire about baseline information after they were recruited for the trial. (18) Intervention GPs were trained in SDM. (21) Control GPs delivered care-as-usual. During the study period, 68 GPs included 226 patients with non-chronic low back pain. (Figure 1) (18)

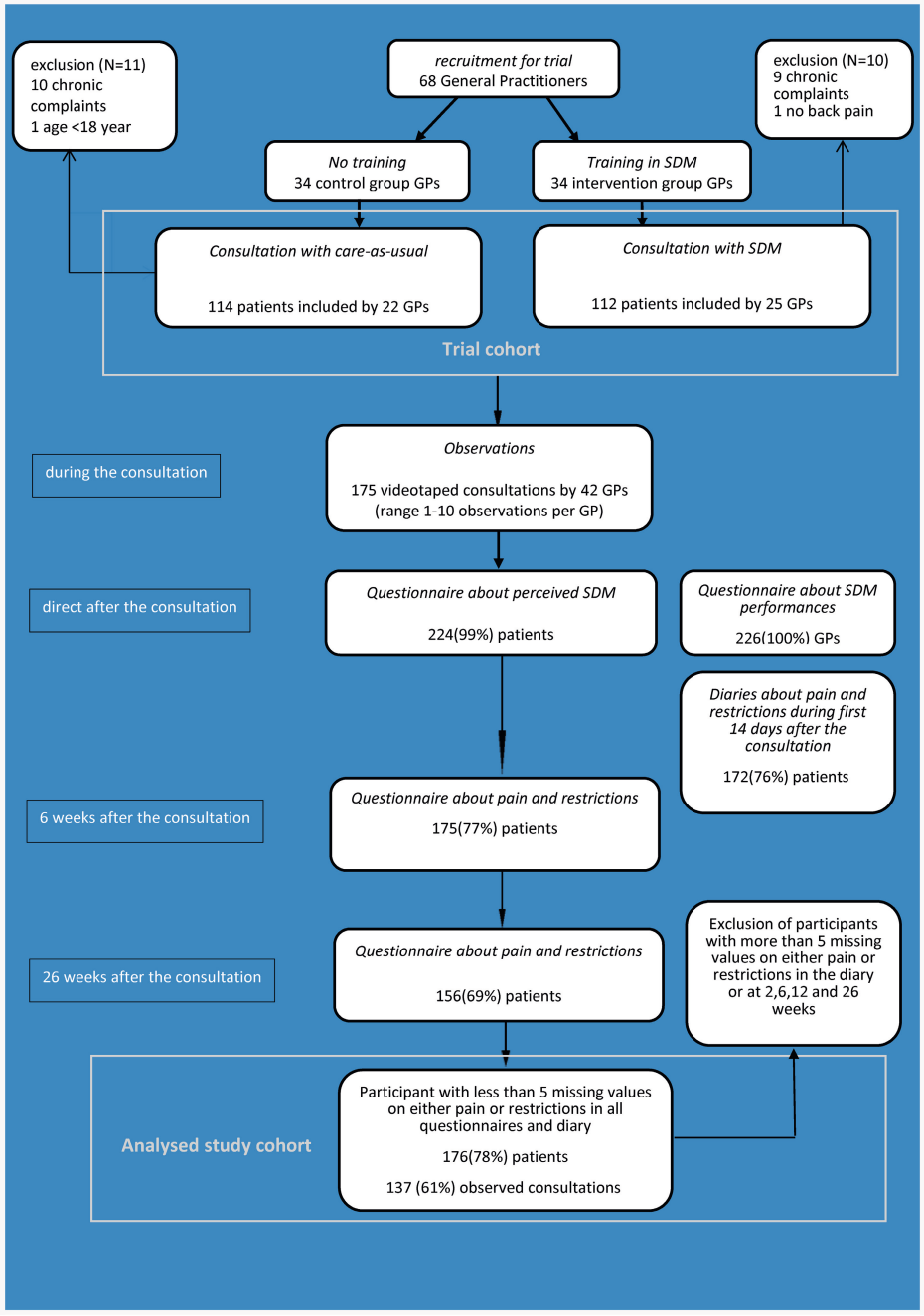


Figure 1. flowchart

Data collection

GPs were asked to reflect on their SDM performance directly after each individual consultation. Patients gave permission for the video-recording of their consultation and were asked to complete questionnaires before and after the consultation with the GP. (18) From all 226 patients in the trial, we had video-recordings of 86 consultations conducted by 23 GPs in the intervention group and 89 consultations conducted by 19 control-group GPs. (Figure 1) (21)

Measurements and instruments

Low back pain-related outcomes

Pain severity was quantified by the validated Visual Analogue Scale (VAS), referring to patients' self-reported level of pain during the past week, ranging from 0 to 10. (22) Perceived functional disability was assessed by the Dutch version of the Roland Morris Disability Questionnaire (RMD). This validated questionnaire contains 24 questions about restrictions in daily activities during the past day, which patients had to tick when applicable. Scores are summed as the number of positive answers. (23)

The primary outcome was recovery at 26 weeks, which was defined by a VAS of 30 mm (scale 0-100) or less combined with a maximum of three disabilities on the RMD questionnaire (range 0-24) at 26 weeks. (20) As a secondary outcome, we assessed recovery at 6 weeks, which was defined by a VAS of 30 mm or less (scale 0-100 mm) combined with a maximum of three disabilities on the RMD (range 0-24) at six weeks after the consultation. (20)

Recorded SDM

SDM was assessed from the three different perspectives.

The observer-reported SDM was assessed using the OPTION scale, a validated observation instrument for measuring the extent to which a healthcare provider involves a patient in SDM. (24) The scale distinguishes 12 process elements of SDM (ranging from 0-4), and these scores are summed to obtain one overall score. A score of 0 corresponds to 'no behaviour observed', and a score of 4 indicates that the behaviour is exhibited to a high degree. The sum scores of the 12 process elements are transformed into a scale from 0-100. (21)

Patient-reported SDM was evaluated by a single question on a 4-point scale directly after the consultation: "How much were you involved in decision-making" ('not at all', 'not really', 'on the whole, yes' and 'yes'). The answers were scored 1 to 4 where 'yes' = 4 and transformed into a scale from 0-100. This non-validated but easily applicable measure requires a patient to provide a global assessment rather than a reflection on the separate process steps. Single-item, generic patient-reported measurements are simple and easy to understand, and have demonstrated comparable validity and reliability to multi-items scales in other fields like quality-of-life measurements. (25)

To measure *GP-reported* SDM, we developed a GP questionnaire, transforming the description of each previously mentioned process element of the OPTION scale into one question about the GP's self-reflection on the level of performance of the corresponding SDM process element. We left out item three (inquiry into the preferred information format), because it might reveal the intervention to the control GPs. In the OPTION, this element is hardly ever scored and has hardly any influence on the overall scale. (26) Scores were on a 4-point scale ranging from 'not at all', 'not really', 'on the whole yes' to 'yes'. All questions were formulated by AS, checked for content by two research students (DA and ME) and tested on two non-participating GP colleagues. The sum score of the 11 process elements was transformed into the qualifying answers: 'not at all' =1, 'not really' =2, 'on the whole yes' =3, and 'yes' =4. The different process items were summed to give one overall score and transformed into a scale from 0-100.

Statistical analysis

Baseline characteristics (i.e. just before, during or directly after the consultation) were reported as means and standard deviations or N and percentages, as applicable. Recovery at 6 and 26 weeks was analysed using logistic regression analysis for each of the three SDM assessments separately. In a first step, we estimated the effect of SDM as reported by the patients and GPs or SDM as scored by independent observers without any adjustment. In a second step, we included the patient's age, sex, educational level, absenteeism and baseline measurements of both the VAS and RMD as potential confounders. (27)

The assumption of linearity of continuous variables, including the SDM scores, was assessed with restrictive cubic splines and tested with likelihood ratio tests. (Supplementary file 1)

Prior to preforming the analysis, we noted substantial missing values for multiple variables, including missing values for the VAS and the RMD at different time points during follow-up, including 6 and 26 weeks, where these scores are used to determine the outcomes for this study. Most of these missing values were due to patients not returning the diary used to assess VAS and RMD during the first 14 days of follow-up.

Fifty patients were excluded because they were missing at least five of the 18 outcomes for either VAS or RMD. In most of the cases this was due to diary measurements that were incomplete or not returned. (Figure 1) Baseline characteristics and recovery rates of all 50 excluded patients and their GPs are given in supplementary file 2, supplementary table 1. For the remaining 176 patients, we used multiple imputation techniques. We imputed missing values for the VAS and RMD measurements over time, SDM as reported by the patient and the GP and the SDM scored by independent observers, the patient's sex, age, absenteeism from work at baseline and scores for the Illness Perception Questionnaire (IPQ). Patient perceptions of low back pain can be confounders and were measured during the consultations as described in the original trial. (18) For patients with missing values for VAS or RMD at 6 or 26 weeks, outcomes were determined based on the imputed scores. The number of imputations was based

on the percentage of patients with one or more missing values. We imputed the data 67 times and performed all analyses on each imputed dataset. Results were pooled according to Rubin's rule. (18) Results were reported as odds ratios with 95% CIs and corresponding p-values. (28)

Spearman's correlations of the non-imputed data were calculated between observer-reported SDM, patient-reported SDM and GP-reported SDM.

RESULTS

Baseline characteristics

GP and patient characteristics, patient recovery rates and the numbers of missing data are provided in table 1 for the participants in this post-hoc analysis. Participating patients and GPs in the constructed database did not differ from the original trial cohort in any of the variables presented in table 1. The mean level of patient-rated pain at baseline was 48.90 (sd 15.70) on a scale from 0-100 and they perceived a disability in 10.03 (sd 5.75) of the 24 items on average. At 6 weeks after consultation 101 (66%) patients were recovered, and at 26 weeks 105 (74%) patients were recovered. (Table 1) The mean level of observer-rated SDM was 30.76 (sd 36.73), which is less than half the value recommended for best practice based on the maximum score of 100. The patient-perceived level of involvement in decision-making was 78.03 (sd 36.73), and GPs scored their SDM performance on average as 53.46 (sd 20.23), all on a scale from 0-100. (Table 1) Almost one quarter (24%) of all patients experienced no involvement at all. (Supplementary file 2, supplementary table 2) GPs indicated that in 22% of the cases they did not involve patients in at least one of the 11 steps of decision-making. (Supplementary file 2, supplementary table 2) Spearman's correlations between the three SDM measurement perspectives were low (<0.150) except for a moderate, significant correlation of 0.418 between observer-reported SDM and GP-reported SDM. (Supplementary file 2, supplementary table 3)

The effect of SDM on recovery at 26 weeks

In the unadjusted analysis the observer-reported SDM process steps measured using the OPTION scale were not significantly associated with recovery at 26 weeks after the consultation (OR 1.026 (95% CI: 0.986-1.068, p-value = 0.206). After adjustment for confounders, the OR was 1.033 (95% CI: 0.987-1.080, p-value = 0.156).

Patient-reported SDM was also not significantly associated with recovery when unadjusted (OR 1.002 (95% CI: 0.992-1.012, p-value = 0.745) and after adjustment (OR 0.998 (95% CI: 0.987-1.009, p-value = 0.723). GP-reported SDM had a non-linear association with recovery. To solve the problem of non-linearity, splines were introduced in the analysis of GP-scores with cut-off points below 40 (indicating almost no SDM according to the GP) or above 70 (indicating high levels of SDM according to the GP). Scores below 40 showed an odds ratio of 0.965 (95% CI: 0924-1.008, p-value = 0.111); medium-rated SDM (scores from 40-70) showed an odds ratio of 1.021 (95% CI 0.995-

1.047, p-value = 0.110); and a high level of SDM (scores above 70) showed an odds ratio of 1.069 (95% CI: 1.010-1.132, p-value = 0.021). After adjustment ORs were 0.985 (95% CI: 0.9140-1.008, p-value = 0.960) for scores below 40, 1.022 (95% CI: 0.993-1.051, p-value = 0.145) for scores between 40 – 70 and 1.076 (95% CI: 1.007-1.150, p-value = 0.031) for scores above 70. (Table 2)

Table 1. Baseline and recovery characteristics of all 176 patients, GP characteristics and the level of SDM.

	number or mean	percentage or standard deviation (sd)	number of missing values (percentage)
PATIENTS CHARACTERISTICS			
male	80	46.2 %	3 (0.02%)
mean age	46.77	13.16 sd	0
educational level			7 (0.04%)
primary school educational attainment only	25	14.8 %	
at least secondary school educational completion	84	49.7 %	
at least college, university completion	60	35.5 %	
absenteeism from work (yes/no)	70	30.3 %	24 (14%)
intervention group	91	51.7 %	0
DISEASE CHARACTERISTICS AT BASELINE			
pain severity (VAS; scale 0-100)	48.90	15.70 sd	27 (15%)
functional disability score (RMD; range 0-24)	10.03	5.75 sd	5 (0.03%)
GP CHARACTERISTICS			
male	27	57 %	0
mean age	51.38	7.029 sd	0
educator	33	70 %	0
years' experience as GP	18.02	7.820 sd	0
mean number of patients included	5.06	3.03 sd	0
SDM			
observer-reported SDM (OPTION scale; 0-100)	30.76	10.82 sd	41 (23%)
patient-reported SDM (scale 0-100)	78.03	36.73 sd	3 (0.02%)
GP-reported SDM (scale 0-100)	53.46	20.23 sd	8 (0.05%)
RECOVERY†			
recovered at 6 weeks	101	66.45 %	24 (14%)
recovered at 26 weeks#	105	73.94 %	34 (19%)

† = recovery defined by a VAS-score <30 mm and a RMD ≤ 3 restrictions; # = primary outcome

Table 2. Unadjusted and adjusted logistic regression of recovery as a function of the prognostic variables after multiple imputation.

	recovered at *26* weeks†						recovered at *6* weeks†					
	unadjusted logic regression			adjusted logistic regression			unadjusted logic regression			adjusted logistic regression		
	odds ratio	confidence interval	p-value	odds ratio	confidence interval	p-value	odds ratio	confidence interval	p-value	odds ratio	confidence interval	p-value
SDM												
observer-reported SDM (measured by OPTION)	1.026	0.986-1.068	0.206	1.033	0.987-1.080	0.156	1.007	0.975-1.041	0.663	1.016	0.979-1.055	0.399
patient-reported SDM	1.002	0.992-1.012	0.745	0.998	0.987-1.009	0.723	0.997	0.988-1.007	0.587	1.006	0.995-1.016	0.301
GP-reported SDM (hardly) no SDM at all (<40)	0.965	0.924-1.008	0.111	0.985	0.9140-1.008	0.960	1.006	0.990-1.023	0.458	1.006	0.988-1.025	0.509
GP-reported SDM intermediate level SDM (40-70)	1.021	0.995-1.047	0.110	1.022	0.993-1.051	0.145						
GP-reported SDM (relatively) high level SDM (>70)	1.069	1.010-1.132	0.021*	1.076	1.007-1.150	0.031*						
PATIENT CHARACTERISTICS												
male	1.230	0.846-1.789	0.278				1.249	0.894-1.743	0.193			
mean age	0.9777	0.950-1.006	0.123				1.012	0.987-1.038	0.339			

educational level secondary school educational completion related to primary school educational attainment only	0.490	0.260-0.924	0.0276*	0.520	0.291-0.930	0.028*
educational level college, university educational completion related to primary school educational attainment only	1.299	0.782-2.156	0.3126	1.232	0.776-1.055	0.376
DISEASE CHARACTERISTICS						
pain severity (VAS; scale 0-100)	0.969	0.943-0.994	0.017*	0.956	0.933-0.981	0.001*
functional disability score (RMD; range 0-24)	1.026	0.953-1.104	0.497	0.988	0.926-1.054	0.715

† = recovery defined by a Visual Analogue Scale <30 mm and a Roland Morris Disability questionnaire ≤ 3 restrictions after the consultation;
* = p-value<0.05

The effect of SDM on recovery at 6 weeks

In the adjusted analysis the observer-reported SDM process steps measured by the OPTION scale showed no effect on recovery at 6 weeks after the consultation, with an odds ratio of 1.016 (95% CI: 0.979-1.055, p-value = 0.399). Patient-reported SDM and GP-reported SDM also showed no effect on recovery, with odds ratios of 1.016 (95% CI: 0.995-1.1.016, p-value = 0.301) for patient-reported SDM and 1.006 (95% CI: 0.988-1.025, p-value = 0.509) for GP-reported SDM. (Table 2)

In the unadjusted analysis a higher baseline level of pain significantly decreased the likelihood of recovery at 6 and 26 weeks. The impact is modest (< 5% less chance of recovery per 10 mm on the VAS scale from 0-100 mm). (Table 2)

DISCUSSION

In this post-hoc analysis we assessed the impact of SDM (as assessed from the patient, GP and observer perspectives) on long-term and short-term recovery from non-chronic low back pain. From any of these three perspectives SDM did not improve recovery from low back pain at 6 or 26 weeks, except for the GP-reported level of SDM, which was associated with recovery at 26 weeks.

From a GP's point of view, there was a non-linear significant effect on recovery. The lowest probability of recovery was observed at a medium level of GP-reported SDM, where each increase in the level of SDM (on the 4-point scale) per single process step of the 11 steps of GP-reported SDM increased the patient's recovery chance at 26 weeks by 2.3 percent.

The results of this post-hoc analysis confirm the conclusion from our trial: there is no convincing evidence that SDM improves outcomes in patients with non-chronic low back pain, despite the fact that patients with low back pain indicate a need for patient-centred care and active involvement and the fact that recovery from low back pain is associated with patients' and GPs' recovery expectations. (9,17) The results of this study further strengthen this conclusion of absence of a detectable effect by considering different angles for the evaluation of SDM: that of patients, of GPs and of external observers. We could not identify an integral SDM effect from these different perspectives. Only one of the six SDM measurements tested was significantly associated with recovery. Moreover, a post-hoc analysis increases the risk that this significance may be caused by multiple testing rather than by a true effect. The fact that GP-perceived SDM was found to be associated with favourable long-term recovery may equally be explained by a professional perception that was not shared by patients, and not confirmed in observation. (17)

Since the introduction of SDM in 1982 by the President's Commission for the Study of Ethical Problems in Medicine and Biomedical and Behavioral Research in its report Making HealthCare Decisions, there has been an ongoing debate about the concept and how to measure optimal performances. *(*29-35) It is worthwhile returning to this original

report as while we do not wish to ignore the knowledge that has been accumulated since then, nor do we want to fall into the trap of restricting ourselves to an elaboration of the concepts that currently happen to receive most attention. The original report describes SDM clearly as 'a process based on mutual respect and partnership, that will usually consist of discussions between professional and patient that bring the knowledge, concerns, and perspective of each to the process of seeking agreement on a course of treatment.' Clearly, health professional and patient both have responsibility in this process: the health professional creates the opportunity for a dialogue and makes sure the patient understand the medical situation and available courses of action, the patient expresses prevailing concerns, needs and wishes. In addition, the practitioner offers alternative courses of action to allow the patient to make a decision according to his views on well-being. (29)

In 2019, after a systemic review of SDM models, Bomhof provided a map of 24 components of SDM elements. (30) Like Joseph-Williams et al., they indicated that exclusive or essential SDM behaviour should be separated from more general, context-related communication skills that serve as facilitators. (30,31) But these general skills play a crucial role in creating an environment to optimise the exclusive elements. (31) Since SDM is displayed in communicative behaviour and perceived in patients' minds, one might argue that even exclusive elements actually serve as facilitators to reach the outcome of 'a shared choice', defined as a mutually agreed plan of action preferred by the patient but achieved after a respectful dialogue where the knowledge, concerns and perspective of each are shared. (29,32,33)

The measurement of SDM should therefor at least include the patient and the outcome. Unfortunately, patients' evaluation of 'a shared choice' seems based solely on the level of mutual agreement without incorporating the quality of the information exchange or deliberation. (34) Patient-reported (and provider-reported) measurements of process elements suffer from ceiling effects, possibly due to the halo effect, defined as incorporating the whole encounter, the ongoing relationship with the clinician or clinician attributes. (7,35)

Observer-based coding schemes, requiring raters trained in the evaluation of SDM, usually apply stricter criteria and reveal lower levels of SDM compared to results based on self-report instruments. (26) But external raters cannot accurately determine the level of 'a shared choice' since this is predominantly perceived in patient's mind. (34) Although reflection ('stop-and-think') before rating did not mitigate ceiling/halo effects, training patients (or providers) as raters of observed or audio-recorded encounters might increase the performance of self-reported SDM measures. (35)

Methodologic considerations

Several limitations need to be addressed. Illness-related characteristics, like levels of experienced pain or disabilities, patient's characteristics, like their mood, behaviour or socio-economic status, or even GP characteristics and the interaction between GP and patients might also influence the prognosis of non-chronic low back pain. However,

adjustment for these variables, which we did in the second step of our analysis, did not change any results. (27) Even in a short time span of 6 weeks we did not find any evidence that SDM might have influenced recovery. But enhanced health outcomes are not the only aim of SDM. Besides ethical considerations, SDM aims to limit practice variation and thus decrease inequality, promote patient autonomy and ensure that treatment decisions reflect patient preferences. (6)

A possible explanation of the observation that GP's reflections on their moderate performance of SDM aligns with a higher chance of developing non-chronic complaints might be that GPs adapt their behaviour to the patient's characteristics associated with recovery rates. (12,36,37) Recently, Arnborg Lund described GPs' views on treating low back pain as an act of dialogue rather than a fragmented experience with different explanations and recommendations. (38)

The OPTION scale used by independent observers of SDM is an externally validated scale. For both patient-reported SDM and GP-reported SDM, easily applicable validated instruments were not available at the time of this study. Even in consultations for a relatively simple complaint like low back pain, patients might experience difficulty in recognising involvement in the decision being assessed. (39) Therefore we decided to use the single question measurement for patient SDM assessment because this is easily applicable and simple to understand. However, we do realise that this measurement is not validated. The overall level of observer-reported SDM of the cohort was low (less than 50% of the maximum score), although comparable to other studies. (14,26) Substantially higher levels of observer-reported SDM behaviour are rarely measured in controlled trials (26). It is unclear what effect substantial observer-reported SDM would have on recovery. (26)

An important methodological limitation is the high number of missing values. Fifty patients were considered lost to follow-up and excluded from the analysis, as the number of missing values was deemed too high. In additional analyses, we detected no clear association between loss to follow-up and SDM measurements or baseline measurements of the VAS or the RMD. For the remaining 176 patients, we used multiple imputation. In line with current recommendations, we based the number of imputations on the percentage of patients with one or more missing values. (40) When evaluating the percentage of patients with missing values, we considered two factors. First, a large proportion of the remaining patients had only a few missing values. (Supplementary file, supplementary table 2) The VAS (over time) was the variable with the highest percentage of missing values. Second, we incorporated all VAS and RMD scores over time, including measurements not used to define the outcome, as consecutive measurements of VAS and RMD scores showed correlations of 0.80. We incorporated these measurements to obtain the best possible imputation model, even though the number of patients with any missing values increased. Nevertheless, a bias due to either loss to follow-up or missing values cannot fully be excluded.

CONCLUSION

In a post-hoc-analysis of RCT data on primary-care patients suffering from non-chronic low back pain, we found no convincing evidence that SDM improves recovery from non-chronic low back pain, neither in the long term nor in the short term. These results were unaffected by the perspective from which SDM was measured (observer, patient perception or GP reflection). Further research should focus on the consistently high performance of SDM to determine whether SDM influences recovery at all.

LIST OF ABBREVIATIONS

CI = confidence interval
IPQ = illness perception questionnaire
GP = general practitioner
LBP = aspecific low back pain
NTR = Netherlands trial register
RCT= randomised controlled trial
RMD = Roland Morris disability questionnaire
sd = standard deviation
SDM = shared decision-making
VAS = visual analogue scale for pain

ETHICS APPROVAL AND CONSENT TO PARTICIPATE

The study protocol for the original trial was assessed by the Ethical Committee of the University Medical Centre of Utrecht and exempted from full assessment. Patients were individually asked to provide permission to be video-recorded by their GP; all patients signed informed consent forms.
All methods were carried out in accordance with relevant guidelines and regulations.

AVAILABILITY OF DATA AND MATERIALS

The datasets used and/or analysed during the current study are available from the corresponding author on reasonable request.

COMPETING INTERESTS

The authors declare that they have no competing interests.

AUTHORS' CONTRIBUTIONS

AS participated in choosing the design, outcomes, and analysis methods, participated in the statistical analysis and interpreted the output, and drafted the manuscript.

NW participated in choosing the design, outcome variables, and analysis methods, interpreted output and revised the final draft for important content, discussion, and conclusions.

SD participated in choosing the design, outcomes, and analysis methods, interpreted output and revised the final draft for important content, discussion, and conclusions.

NZ participated in choosing the design and outcome variables, performed statistical analysis and interpreted output, and participated in the draft of the statistical analysis, results and discussion sections.

All authors read and approved the final version of the manuscript.

ACKNOWLEDGEMENTS

The authors thank the patients and healthcare providers who participated in the study.

AUTHORS' INFORMATION

The corresponding author is a GP and worked as a GP vocational trainer for 12 years.

REFERENCES

1) Maher C, Underwood M, Buchbinder R. Non-specific low back pain. Lancet (London, England) 2017;389:736-747.
2) Murray CJ, Vos T, Lozano R, Naghavi M, Flaxman AD, Michaud C, et al. Disability-adjusted life years (DALYs) for 291 diseases and injuries in 21 regions, 1990-2010: a systematic analysis for the Global Burden of Disease Study 2010. Lancet 2012 Dec 15;380(9859):2197-2223.
3) van Tulder MW, Koes BW, Bouter LM. A cost-of-illness study of back pain in The Netherlands. Pain 1995 Aug;62(2):233-240.
4) Bons SCS, Borg, M. A. J. P., Van den Donk M, Koes BW, Kuijpers T, Ostelo, R. W. J. G., et al. NHG-Standaard Aspecifieke lagerugpijn (Tweede herziening). 2017; Accessed 10/21, 2021.
5) Stiggelbout AM, Van der Weijden T, De Wit MP, Frosch D, Legare F, Montori VM, et al. Shared decision making: really putting patients at the centre of healthcare. BMJ 2012 Jan 27;344:e256.
6) Stiggelbout AM, Pieterse AH, De Haes JC. Shared decision making: Concepts, evidence, and practice. Patient Educ Couns 2015 October 01;98(10):1172-1179.
7) Kasper J, Heesen C, Kopke S, Fulcher G, Geiger F. Patients' and observers' perceptions of involvement differ. Validation study on inter-relating measures for shared decision making. PLoS One 2011;6(10):e26255.
8) Legare F, Adekpedjou R, Stacey D, Turcotte S, Kryworuchko J, Graham ID, et al. Interventions for increasing the use of shared decision making by healthcare professionals. Cochrane Database Syst Rev 2018 Jul 19;7:CD006732.
9) Shay LA, Lafata JE. Where is the evidence? A systematic review of shared decision making and patient outcomes. Med Decis Making 2015 January 01;35(1):114-131.
10) van der Windt, D. A., Kuijpers T, Jellema P, van der Heijden, G. J., Bouter LM. Do psychological factors predict outcome in both low-back pain and shoulder pain? Ann Rheum Dis 2007 Mar;66(3):313-319.
11) Lewis M, Morley S, van der Windt, D. A., Hay E, Jellema P, Dziedzic K, et al. Measuring practitioner/therapist effects in randomised trials of low back pain and neck pain interventions in primary care settings. Eur J Pain 2010 Nov;14(10):1033-1039.
12) Darlow B, Fullen BM, Dean S, Hurley DA, Baxter GD, Dowell A. The association between health care professional attitudes and beliefs and the attitudes and beliefs, clinical management, and outcomes of patients with low back pain: a systematic review. Eur J Pain 2012 January 01;16(1):3-17.
13) Farin E, Gramm L, Schmidt E. The patient-physician relationship in patients with chronic low back pain as a predictor of outcomes after rehabilitation. J Behav Med 2012 Apr 4.
14) Meijers MC, Noordman J, Spreeuwenberg P, Olde Hartman TC, van Dulmen S. Shared decision-making in general practice: an observational study comparing 2007 with 2015. Fam Pract 2019 May 23;36(3):357-364.
15) Fassaert T, van Dulmen S, Schellevis F, van der Jagt L, Bensing J. Raising positive expectations helps patients with minor ailments: a cross-sectional study. BMC Fam Pract 2008 June 30;9:38-38.
16) Chou L, Ranger TA, Peiris W, Cicuttini FM, Urquhart DM, Sullivan K, et al. Patients' perceived needs of health care providers for low back pain management: a systematic scoping review. Spine J 2018 April 01;18(4):691-711.
17) Kapoor S, Shaw WS, Pransky G, Patterson W. Initial patient and clinician expectations of return to work after acute onset of work-related low back pain. J Occup Environ Med 2006 Nov;48(11):1173-1180.

18) Sanders ARJ, Bensing JM, Magnee T, Verhaak P, de Wit NJ. The effectiveness of shared decision-making followed by positive reinforcement on physical disability in the long-term follow-up of patients with nonspecific low back pain in primary care: a clustered randomised controlled trial. BMC Fam Pract 2018 Jun 28;19(1):102-8.
19) Kuijer W, Brouwer S, Dijkstra PU, Jorritsma W, Groothoff JW, Geertzen JH. Responsiveness of the Roland-Morris Disability Questionnaire: consequences of using different external criteria. Clin Rehabil 2005 Aug;19(5):488-495.
20) Mehling WE, Gopisetty V, Acree M, Pressman A, Carey T, Goldberg H, et al. Acute low back pain and primary care: how to define recovery and chronification? Spine (Phila Pa 1976) 2011 Dec 15;36(26):2316-2323.
21) Sanders AR, Bensing JM, Essed MA, Magnee T, de Wit NJ, Verhaak PF. Does training general practitioners result in more shared decision making during consultations? Patient Educ Couns 2017 Mar;100(3):563-574.
22) Kamper SJ, Ostelo RW, Knol DL, Maher CG, de Vet HC, Hancock MJ. Global Perceived Effect scales provided reliable assessments of health transition in people with musculoskeletal disorders, but ratings are strongly influenced by current status. J Clin Epidemiol 2010 Jul;63(7):760-766.e1.
23) Brouwer S, Kuijer W, Dijkstra PU, Goeken LN, Groothoff JW, Geertzen JH. Reliability and stability of the Roland Morris Disability Questionnaire: intra class correlation and limits of agreement. Disabil Rehabil 2004 Feb 4;26(3):162-165.
24) Pellerin MA, Elwyn G, Rousseau M, Stacey D, Robitaille H, Legare F. Toward shared decision making: using the OPTION scale to analyze resident-patient consultations in family medicine. Acad Med 2011 Aug;86(8):1010-1018.
25) de Boer AG, van Lanschot JJ, Stalmeier PF, van Sandick JW, Hulscher JB, de Haes JC, et al. Is a single-item visual analogue scale as valid, reliable and responsive as multi-item scales in measuring quality of life? Qual Life Res 2004 March 01;13(2):311-320.
26) Couet N, Desroches S, Robitaille H, Vaillancourt H, Leblanc A, Turcotte S, et al. Assessments of the extent to which health-care providers involve patients in decision making: a systematic review of studies using the OPTION instrument. Health Expect 2013 Mar 4.
27) Kent PM, Keating JL. Can we predict poor recovery from recent-onset nonspecific low back pain? A systematic review. Man Ther 2008 Feb;13(1):12-28.
28) Graham JW, Olchowski AE, Gilreath TD. How Many Imputations are Really Needed? Some Practical Clarifications of Multiple Imputation Theory. Prev Sci 2007 “june 5, “;8:206-213.
29) United States. President’s Commission for the Study of Ethical Problems in Medicine and Biomedical and Behavioral Research. U S Code Annot U S 1982;Title 42 Sect. 300v as added 1978:Unknown.
30) Bomhof-Roordink H, Gartner FR, Stiggelbout AM, Pieterse AH. Key components of shared decision making models: a systematic review. BMJ Open 2019 December 17;9(12):e031763-031763.
31) Joseph-Williams N, Elwyn G, Edwards A. Knowledge is not power for patients: a systematic review and thematic synthesis of patient-reported barriers and facilitators to shared decision making. Patient Educ Couns 2014 March 01;94(3):291-309.
32) Pieterse AH, Bomhof-Roordink H, Stiggelbout AM. On how to define and measure SDM. Patient Educ Couns 2018 August 01;101(8):1307-1309.
33) Crispin V, Bugge C, Stoddart K. Sufficiency and relevance of information for inpatients in general ward settings: A qualitative exploration of information exchange between patients and nurses. Int J Nurs Stud 2017 October 01;75:112-122.
34) Shay LA, Lafata JE. Understanding patient perceptions of shared decision making. Patient Educ Couns 2014 September 01;96(3):295-301.

35) Kunneman M, LaVecchia CM, Singh Ospina N, Abu Dabrh AM, Behnken EM, Wilson P, et al. Reflecting on shared decision making: A reflection-quantification study. Health Expect 2019 October 01;22(5):1165-1172.
36) Giroldi E, Veldhuijzen W, Leijten C, Welter D, van der Weijden T, Muris J, et al. 'No need to worry': an exploration of general practitioners' reassuring strategies. BMC Fam Pract 2014 July 07;15:133-133.
37) Kappen T, van Dulmen S. General practitioners' responses to the initial presentation of medically unexplained symptoms: a quantitative analysis. Biopsychosoc Med 2008 November 17;2:22-22.
38) Arnborg Lund R, Kongsted A, Backer Hansen E, Myburgh C. Communicating and diagnosing non-specific low back pain: A qualitative study of the healthcare practitioners perspectives using a social diagnosis framework. J Rehabil Med 2020 March 31;52(3):jrm00036-2656.
39) Barr PJ, Elwyn G. Measurement challenges in shared decision making: putting the 'patient' in patient-reported measures. Health Expect 2016 October 01;19(5):993-1001.
40) Graham JW, Olchowski AE, Gilreath TD. How many imputations are really needed? Some practical clarifications of multiple imputation theory. Prev Sci 2007 September 01;8(3):206-213.
41) Harrell FEJ. Regression Modeling Strategies: With Applications to Linear Models, Logistic and Ordinal Regression, and Survival Analysis . 2nd ed.: Springer Series in Statistics; 2015

SUPPLEMENTARY FILE 1 'ADDITIONAL INFORMATION REGARDING THE METHODS OF THE ANALYSIS'

Restrictive cubic splines in regression

Most often used regression methods (i.e. linear, logistic, Cox regression) rely on an assumption of linearity for continuous explanatory variables such as age. For age in a logistic model, this may be seen as an assumption that the odds ratio for an increase from 30 to 40 years is the same compared to an increase from 60 to 70 years of age. This, however, may not always be the case: the association may for example follow a parabolic form (to be modelled with a squared term) or any other form. This is sometimes referred to as the functional form of the association. Key problem is that a functional form is often unknown prior to a statistical analysis.

In this study, two of the three SDM assessments as well as age were included in the analysis as continuous explanatory variables. As a violation of this assumption may lead to biased results, we tested the assumption (for each variable and each outcome separately) with restrictive cubic splines.

In this document, we provide a somewhat simplified explanation of splines as well as additional information on how these were used in the analysis. Even though we made every effort to provide a valid, yet easy to understand explanation, this document should by no means be considered as a full explanation on the use of restrictive cubic splines. We refer interested readers to Harrell (2015, chapter 2) for more detailed explanation as well as additional references. (Harrell, 2015)

The basic method of a spline involves dividing the distribution of a continuous variable into parts, i.e. the lowest value to the 20^{th} percentile, 20^{th} to 40^{th} percentile, etc. The cut-points are referred to as knots. It is important to realise that this is not a form of categorisation (i.e. the lowest value to the 20^{th} percentile becomes category 1), the values of individual patients within a certain part is left continuous. Instead of estimating a single OR, separate ORs are estimated for each part the distribution was divided in. Hence in the example above, 5 separate ORs may be estimated, introducing more flexibility in the assumption of linearity, whatever the functional form of the association may be.

This somewhat simplified example of a spline has several downsides: within a regression model the number of coefficients (i.e. log odds ratio's) that need to be estimated may become large and stability of the statistical model may be problematic. A more advanced approach is the restrictive cubic spline (also called natural splines), which involves the inclusion of so called higher order (e.g. Age^3). This method provides similar advantages in terms of flexibility with more stable results while estimating less regression coefficients.

Application

During the first phase of the analysis, we tested restrictive cubic spline (RCS) and compared these with a likelihood ratio test. First tests were performed in a univariate model, any non-linear effect was subsequently evaluated in a multivariate model. When applying splines, a decision needs to be made on the number of knots to be used and where these knots need to be placed. Initially, we relied on the default settings of the software (SAS v9.4) and evaluated increasing number of knots. Additionally, we evaluated the location of knots over the multiple imputed datasets, as the above-described method allows for different knot locations in different imputed datasets. A non-linear effect was only observed for GP-reported SDM. Although some differences in the knot locations for this variable were observed, in general, knots located at values of about 40, 55 and 70 seemed generally acceptable. These knots were subsequently used in the analysis of all imputed data to allow for pooling of the results and estimation of pooled odds ratios with corresponding 95% confidence intervals and p-values.

SUPPLEMENTARY FILE 2 'SUPPLEMENTARY TABLES'

Supplementary table 1. Baseline characteristics and recovery rates of all 50 excluded patients and their GPs.

	number or mean	percentage or standard deviation (sd)	number of missing values
PATIENTS' CHARACTERISTICS			
male	27	54 %	1
mean age	38.14	14.07 sd	0
educational level			1
primary school educational attainment only	9	18 %	
at least secondary school educational completion	25	50 %	
at least college, university completion	15	30 %	
absenteeism from work (yes/no)	28	56%	5
intervention group	21	42 %	0
DISEASE CHARACTERISTICS AT BASELINE			
pain severity (VAS; 0-100)	43.923	17.93 sd	7
functional disability score (RMD; 0-24)	9.960	4.907	0
DISEASE CHARACTERISTICS AT 6 WEEKS			
pain severity (VAS; 0-100)	13.890	15.33 sd	43
functional disability score (RMD; 0-24)	4.143	5.76 sd	43
DISEASE CHARACTERISTICS AT 26 WEEKS			
pain severity (VAS; 0-100)	19.619	26.70 sd	43
functional disability score (RMD; 0-24)	2.833	3.37 sd	44
SDM			
observer-reported (OPTION scale; 0-100)	29.955	11.35 sd	13
patient-reported SDM (scale 0-100)	75.694	35.56 sd	2
GP-reported SDM (scale 0-100)	56.000	20.73 sd	0
RECOVERY†			
recovered at 6 weeks	3	42.9 %	43
recovered at 26 weeks#	4	57.1 %	43

Supplementary table 1. (Continued)

	number or mean	percentage or standard deviation (sd)	number of missing values
GP CHARACTERISTICS			
male	16	59.3 %	0
mean age	53.296	7.05 sd	0
educator	21	77.8 %	0
years' experience as GP	20.667	7.92 sd	0

† = recovery defined by a VAS-score <30 mm and a RMD ≤ 3 restrictions after the consultation;
# = Primary outcome

SUPPLEMENTARY TABLES

Supplementary table 2. Levels of SDM from all three perspectives of the analysed cohort of 176 cases. Absolute numbers (percentage)

	not observed	minimal attempt	minimum level	good level	very high level	not at all	not really	on the whole yes	yes	number of missing values
SDM										
OBSERVER-REPORTED SDM										
GP draws attention to a decision-making stage	44 (25%)	78 (44%)	9 (5%)	2 (1%)	4 (2%)					39 (22%)
equipoise	89 (51%)	14 (8%)	24 (14%)	2 (1%)	8 (5%)					39 (22%)
information format	83 (47%)	48 (27%)	2 (1%)	3 (2%)	1 (0%)					39 (22%)
lists options	20 (11%)	0 (0%)	32 (18%)	27 (15%)	58 (33%)					39 (22%)
explanation of pros and cons of options	31 (18%)	47 (27%)	52 (30%)	5 (3%)	2 (1%)					39 (22%)
exploration of the patient's expectations	5 (3%)	47 (27%)	79 (45%)	5 (3%)	1 (0%)					39 (22%)
exploration of the patient's concerns	129 (73%)	7 (4%)	1 (0%)	0 (0%)	0 (0%)					39 (22%)
check of patient's understanding	13 (7%)	105 (60%)	19 (11%)	0 (0%)	0 (0%)					39 (22%)
offering opportunities to ask questions	20 (11%)	66 (38%)	50 (28%)	0 (0%)	0 (0%)					40 (23%)
elicitation of patient's preferred level of involvement	25 (14%)	42 (24%)	48 (27%)	12 (7%)	10 (6%)					39 (22%)
indication of a decision-making stage	81 (46%)	54 (31%)	1 (0%)	0 (0%)	0 (0%)					40 (23%)

indication of the need to review the decision	32 (18%)	10 (6%)	7 (4%)	1 (0%)	87 (49%)					39 (22%)
PATIENT-REPORTED SDM										
						42 (24%)	11 (6%)	40 (23%)	76 (43%)	7 (4%)
GP-REPORTED SDM										
GP draws attention to a decision-making stage						31 (18%)	56 (32%)	41 (23%)	44 (25%)	4 (2%)
equipoise						42 (24%)	46 (26%)	43 (24%)	43 (24%)	2 (1%)
lists options						10 (6%)	37 (21%)	39 (21%)	87 (50%)	3 (2%)
explanation of pros and cons of options						47 (27%)	36 (20%)	49 (28%)	42 (24%)	2 (1%)
exploration of the patient's expectations						24 (14%)	38 (22%)	65 (37%)	46 (26%)	3 (2%)
exploration of the patient's concerns						19 (11%)	55 (31%)	56 (32%)	43 (24%)	1 (0%)
check of patient's understanding						16 (9%)	35 (20%)	58 (33%)	63 (36%)	4 (2%)
offering opportunities to ask questions						1 (0%)	27 (15%)	59 (34%)	86 (49%)	3 (2%)
elicitation of patient's preferred level of involvement						53 (30%)	54 (31%)	44 (25%)	23 (13%)	2 (1%)
indication of a decision-making stage						96 (55%)	36 (20%)	22 (13%)	19 (11%)	3 (2%)
indication of the need to review the decision						79 (45%)	35 (20%)	38 (22%)	21 (12%)	3 (2%)
						418 (22%)	399 (21%)	514 (27%)	517 (28%)	30 (2%)

Supplementary table 3. Spearman's correlations between observer-reported SDM, patient-reported SDM and GP-reported SDM all scaled 0-100 of the *non-imputed* post-hoc analysis cohort

		observer-reported SDM	*patient-reported SDM*	*GP-reported SDM*
observer-reported SDM	number	135	133	131
	correlation	1	0.123	0.418*
patient-reported SDM	number	133	173	165
	correlation	0.123	1	0.146
GP-reported SDM	number	131	165	168
	correlation	0.418*	0.146	1

* = correlation is significant ($p<0.05$) at the 0.01 level (2-tailed)

CHAPTER 6

A cost-effectiveness analysis of shared decision-making followed by positive reinforcement versus usual care in patients with non-chronic low back pain in primary care.

Cost-effectiveness of shared decision-making in non-chronic low back pain.

Ariëtte R.J. Sanders
Rabin Neslo
Jan Willem Uffen
Marieke van Noord
Niek J. de Wit
G. Ardine de Wit

2022 Leersum

ABSTRACT

Objective. To study the cost-effectiveness of shared decision-making followed by a positive reinforcement (SDM&PR), compared to usual care (UC) for patients suffering from non-chronic low back pain.

Design. Cost-effectiveness analysis from a societal perspective alongside a clustered randomised controlled trial.

Intervention. A consultation including SDM&PR delivered by a General Practitioner (GP) trained in 2 sessions of 2.5 hours each by a peer GP.

Setting and participants. Twenty-five trained Dutch GPs offered an SDM&PR consultation to 112 patients with non-chronic low back pain. The control group consisted of 22 GPs, who performed UC to 114 patients with non-chronic low back pain.

Primary and secondary outcome measures. The primary outcome was perceived functional status, measured with the Roland Morris Disability scale (RMD). Data on disability, freely accessible healthcare use and absenteeism were collected by patient questionnaires after 2, 6, 12 and 26 weeks. Medication use and use of other healthcare services were extracted from GP registration systems. Cost per perceived point functional improvement was compared over the 26 weeks follow-up period. Using a published mapping function, RMD outcomes were translated into utilities to estimate costs per Quality Adjusted Life Year (QALY). Probabilistic sensitivity analyses were performed.

Results. There was no clinically relevant nor significant difference between the groups in RMD-score over the 26 weeks period. SDM&PR increased consultation time with three minutes.
Healthcare costs were € 50.75 per patient lower after SDM&PR compared to the UC. After mapping RMD scores to QALYs, SDM&PR appeared cost-effective.

Conclusion. SDM&PR compared to usual care, did not lead to significant differences in disabilities or costs but increased consultation duration by 33%. However, mapping the disability effects of our trial to QALYs suggests that SDM&PR may be cost-effective at the Dutch threshold level for cost-effectiveness. Future research on effects of interventions to increase patient participation or on low back pain recovery improvement should include more general health outcomes like QALYs to facilitate economic evaluations.

Strengths and limitations of the study

- This cost-effectiveness study from a societal perspective is performed alongside a practice based clustered randomised controlled study.
- The estimation of the costs is based on real-life data and compares an intervention with care-as-usual.
- The complete cases analysis confirms the results of the imputed analysis.
- For the cost utility analysis, a validated mapping function was used.

INTRODUCTION

Low back pain, i.e. back pain without a known specific somatic origin is among the top ten of most frequently presented complaints in primary care. (1,2) Worldwide it is one of the leading causes of disability. Although low back pain subsides within two weeks in the majority of patients, low back pain can become chronic (>3 months) or frequently recurring (≥3 episodes per year). (4) However, even non-chronic low back pain, is a major economic burden due to its high prevalence. Low back pain was responsible for 83.063.000 years lived with disabilities worldwide in 2010. (3,4) The direct healthcare costs, like medication or use of healthcare facilities, are relatively low (7% of total costs). (4) The high societal burden of low back pain is primarily incurred through costs related to losses in productivity (93% of total costs).(4) Many interventions for low back pain in primary care were studied, but few were found effective in terms of improved recovery or prevention of chronicity. (5)

Recovery is highly influenced by contextual factors such as patients' expectations. (6) Positive expectations about a beneficial spontaneous course lead to better recovery and make patients return to work sooner. (6, 7) Shared decision-making (SDM) aims at involving patients' expectations and preferences in treatment decisions. (8) After SDM, patients experience less fear and adhere better to therapy. (9) A positive reinforcement (PR) by their general practitioner (GP), described as a firm statement about an expected fast recovery, may strengthen this effect. (6)

We performed a clustered randomised controlled trial comparing the effects of SDM followed by a positive reinforcement of the chosen therapy (SDM&PR) to usual care on the recovery of patients suffering from non-chronic low back pain in general practice in the Netherlands. In this trial we found a significant difference in patients' perceived level of involvement but no effects on recovery from functional disability, pain or absenteeism. (7,8)

Even if interventions demonstrate minor or no clinical effects on patient reported outcomes, they can still be cost-effective. (9-11) Small differences in clinical improvement may lead to less healthcare costs or earlier return to work, thus reducing direct and indirect disease related costs. (12) It has been postulated that SDM could be cost-effective because patients make more reasonable choices from a health care perspective. (13,14) However, this suggested economic advantage is also disputed. (15) Patel et al. even found negative effects on health status and costs of SDM with patients suffering from low back pain in primary care. (16)

The aim of this study was to assess the cost-effectiveness of introducing SDM&PR for patients with non-chronic low back pain in GP practice, in comparison with usual care.

MATERIALS AND METHODS

Design and randomisation

This study is an economic evaluation conducted alongside a clustered randomised controlled trial. The aim of the trial is to study effects of SDM&PR on functional recovery from disabilities in patients with non-chronic low back pain. (8) For this study, 68 Dutch general practitioners, recruited from the academic primary care network in Utrecht, the Netherlands, between August 2009 and May 2011, were randomly assigned to an intervention group (34 GPs) or control group (34 GPs). (8)

Intervention and control condition

The intervention GPs were trained in two sessions of two and a half hours, and received a decision aid to be used in patient contacts and a desktop card summarizing all consecutive process elements for SDM & PR. (7) The training was based on Kolbs' learning principles and the behavioural process elements as described by Elwyn. (17,18) GPs in the intervention arm introduced the decision-making, discussed the pros and cons of the therapeutic options, made inquiries about patients' expectations, concerns and preferred level of involvement in decision-making, came to an agreement with the patient on a therapeutic plan, positively reinforced this plan and ended with planning a follow-up to evaluate the plan with the included patients. The decision aid, specifically developed for this trial according to the International Patient Decision Aids Standards (IPDAS)-guidelines, was to be introduced during the consultation by the GP. (19) Patients in the control group received usual care. Although routine management was not predefined in the instructions for the study, GPs in the Netherlands are reported to follow the professional guidelines on low back pain in 70% of the patients. (20) Discussion of the favourable prognosis of non-chronic low back pain in general is part of the suggested management in the guideline, but SDM is not. (1)

Inclusion and blinding

The GP practice assistants recruited up to ten patients who contacted the practice with complaints of non-chronic low back pain. Inclusion criteria were age between 18-65 years, and consultation for a new episode of non-chronic low back pain, as defined by the guideline of the Dutch College of General Practitioners. (1) Exclusion criteria were: duration of low back pain longer than three months, any previous episode of low back pain within the three months prior to onset of the present episode, pregnancy and insufficient mastering of the Dutch language. All patients and practice assistants were blinded to the allocation. For a more detailed description see elsewhere. (7,8)

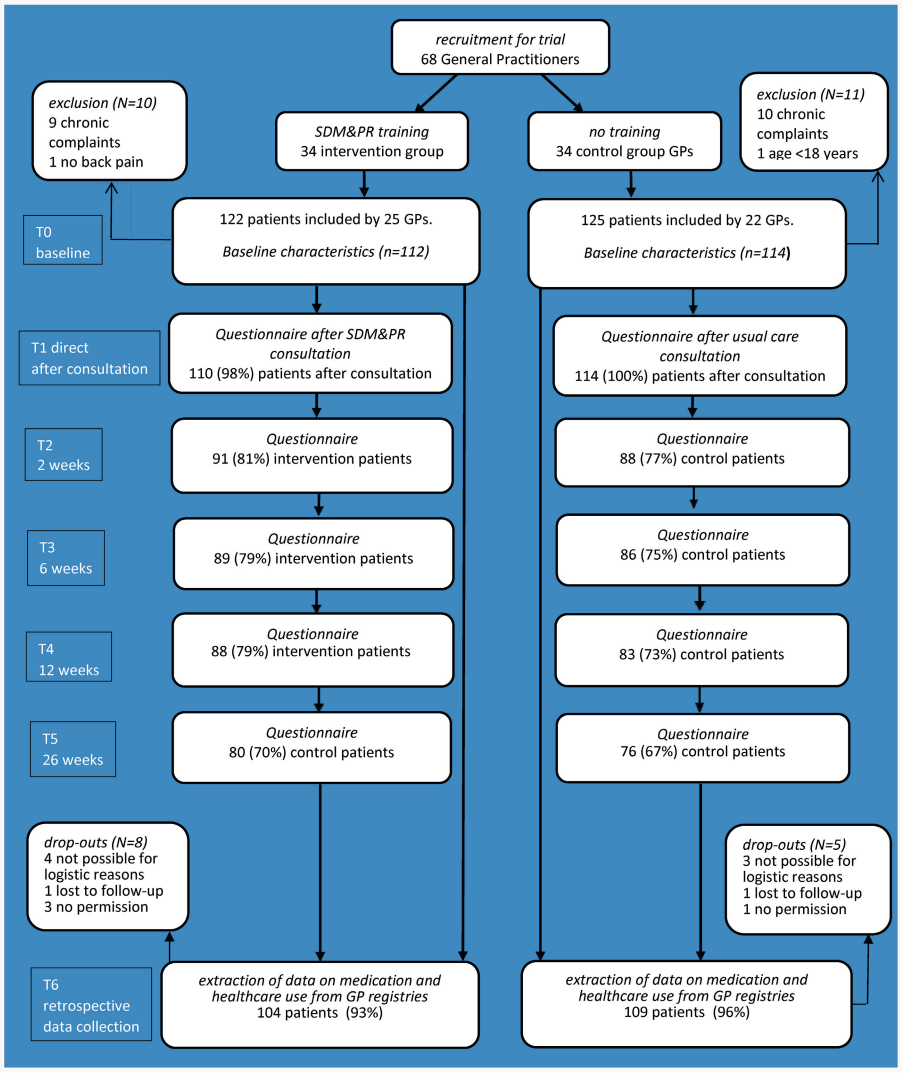


Figure 1. Flow of the participants in different phases through the trial

T0= directly *before* the consultation; T1= directly *after* the consultation; T2= 2 weeks after consultation; T3= 6 weeks after consultation; T4= 12 weeks after consultation; T5= 26 weeks after consultation; T6= retrospective data collection over the 26 weeks trial period from GP registries, GP=General Practitioner

Baseline data and clinical outcomes

Patients' baseline characteristics (age, sex, educational level, ethnic origin and non-chronic low back pain characteristics) were collected after written informed consent was

provided. Patients' illness perceptions, including recovery expectations, were recorded at baseline with the Dutch version of the abbreviated Illness Perception Questionnaire (IPQ). (21) Non-chronic low back pain impact was assessed with the Roland Morris Disability Questionnaire (RMD) at baseline and at 2, 6, 12 and 26 weeks after inclusion in the study. (22) (Figure 1) This questionnaire describes 24 potential functional restrictions due to low back pain. (22) Each functional restriction adds to the RMD-score (ranging from 0-24), higher scores reflecting more perceived functional restrictions. To compare the two groups regarding their RMD-score, we used an area-under-the-curve approach to capture the entire effect over the 26 weeks follow-up period.

Cost data

We conducted a cost effectiveness analysis from a healthcare and societal perspective and therefore included healthcare costs, patient and family costs (out-of-pocket costs) and productivity costs (productivity losses).

Information on the number of GP contacts, laboratory- and additional (diagnostic) investigations, prescribed medication and referral to secondary care were extracted from GP electronic medical records after completion of the 26 weeks follow-up period. Out-of-pocket payments of patients for physiotherapy, podiatry, massage or devices, like braces, were extracted from patients' questionnaires at 2, 6, 12 and 26 weeks after the baseline consultation. We excluded costs for over-the-counter medication, because patient recall of these purchases was expected to be unreliable over this relatively long period. Furthermore, these costs are expected to be relatively low compared to other costs. The patient questionnaires also included questions on ability to work, including sick leave because of low back pain. Patients had to indicate what their normal working pattern was, and whether they were absent from work as a consequence of low back pain. Each lost working day was valued at € 243.25, following Dutch guidelines for economic evaluation in healthcare. (23,24) The reference year for cost estimation was 2010. Unit prices of all cost categories are included. (Appendix 1)

Cost of the intervention

The 34 intervention GPs received two sessions of two and a half hours training and 23 of them recorded 86 consultations during the study period to receive feedback on SDM&PR behaviour. (16) We estimate that the invested time per trained GP on received written feedback was 3 minutes. On average the consultation duration of the control group took 13.1 minutes versus 15.8 minutes for intervention consultations. Consultations in general practice in the Netherlands take, on average, 10.2 min. (25) In 2010, GPs received an hourly rate of 53.20 euros.(26)

The costs of the GPs training were not included in the analysis, because the introduction of SDM&PR should lead to a consistent change in behaviour and thus these costs cannot be attributed to this intervention only.

Cost effectiveness analysis

All analyses were carried out on an intention-to-treat basis. Complete questionnaires were available for 156 patients (69%) and GP registration data for 213 (94%) patients. (Figure 1). Missing values for the questionnaire items were imputed using multiple imputation based on age, sex, treatment allocation and four illness perceptions from the abbreviated illness perception questionnaire (baseline consequences, timeline acute/chronic, personal control and treatment control scores). (21) These four perceptions were included in the imputation model because they correlate with a poor prognosis. (27)

Incremental cost-effectiveness ratios (ICERs) were calculated dividing the mean difference in costs between groups by the mean difference in Area Under the Curve (AUC) of RMD-scores over the 26-week follow-up period. Probabilistic uncertainty analysis was done by means of bootstrapping (5000 replications). A cost-effectiveness plane and a cost-effectiveness acceptability curve were plotted using the results from this bootstrap analysis.

Mapping disease-specific outcomes to Quality Adjusted Life Years and cost utility estimation

As quality of life was not specified as outcome in our trial and the societal willingness to pay for a better RMD outcome is unknown, we transferred RMD-scores to Quality Adjusted Life Year (QALYs), using a mapping function that was developed by Madan et al, using data collected from 701 UK primary care patients with sub-acute or chronic low back pain. (28) The estimated mapping algorithms were validated using external data from the UK Back Pain Exercise and Manipulation trial (n=1063). (28) We chose the best performing model for RMD to EQ-5D mapping from Madan et al. and transformed all observed RMD-scores to QALYs based on the utility mapping function. (28) Thereafter ICERs were estimated and probabilistic uncertainty analyses were performed as described above.

Sensitivity analysis

Cost-effectiveness and cost-utility was also estimated based on complete cases for the Roland Morris outcome and the mapped utilities only, including 72 control (63%) and 80 (71%) intervention cases without taking the consultation length into account. (Table 4)

Statistical Analysis

To test for baseline differences between patients with complete and missing questionnaires, patient characteristics and disease-related parameters were univariately tested with T-test for continuous variables and X^2-distributions for dichotomous variables. Differences in means between groups in primary outcomes (RMD and RMD AUC) and secondary outcomes (QALY's) were tested with two-tailed T-tests assuming heteroscedasticity. A significance level of 0.05 was used to reject or accept the null

hypothesis that there is no difference in means between the two groups. The same significance level was used to compute the confidence intervals around the means.

RESULTS

Baseline characteristics

Baseline characteristics of all included patients are shown in Table 1. Patients in the two groups were comparable regarding individual characteristics and disease-related parameters, except that the intervention group at baseline contained more patients that were temporarily unable to work. Furthermore, patients in the intervention group expected their complaints to last longer (timeline).

Table 1. Baseline demographic and clinical characteristics of patients (non-imputed data). Continuous variables values are means (standard deviation). Dichotomous variables values are numbers (percentages).

	SDM&PR group (n=112)	control group (n=114)
PATIENT CHARACTERISTICS		
mean age (years)	45.4 (13.2)	44.3 (14.4)
male†	52 (47%)	55 (49%)
Dutch origin‡	97 (91%)	103 (93%)
educational level‡		
primary only	15 (14%)	19 (17%)
secondary	56 (52%)	53 (48%)
college, university	36 (34%)	39 (35%)
employed§	73 (70%)	71 (70%)
BASELINE CLINICAL CHARACTERISTICS		
functional disability score (RMD, range 0-24) (primary outcome)¶	10.7 (5.0)	10.3 (5.2)
absent from work because of LBP*	39 (35%)	19 (20%)
illness perception dimensions (IPQ) (0-10)§		
consequences	6.3 (2.3)	6.1 (2.5)
timeline	4.2 (2.8)	3.5 (2.4)
personal control	5.0 (2.2)	5.4 (2.1)
treatment control	6.6 (1.9)	6.9 (1.9)

SDM&PR = Shared decision making followed by a Positive reinforcement, RMD = Roland Morris Disability Questionnaire, LBP = Low Back Pain, IPQ = Illness Perception Questionnaire, ¶ n = 3 missing, * n = 34 missing, † n = 4 missing, ‡ n = 8 missing, § n = 19 missing.

Primary outcomes

Table 2 shows that the RMD-scores, both the absolute difference score and in the AUC analysis, did not differ significantly between groups at 2, 6, 12 and 26 weeks.

Secondary outcomes

Table 2 shows that the mapped QALYs were significantly higher in the intervention group compared to the control group.

Table 2. Patient reported outcomes per phase and Area Under the Curve (AUC) expressed by the Roland Morris Disability Questionnaire (scale 0-24) and QALYs. Intention-to-treat analysis on imputed dataset and AUC.

	SDM&PR group Mean (SD; 95% CI)	Control group Mean (SD; 95% CI)	Difference Mean difference (SD; 95% CI)	p-value
Roland Morris-score				
baseline	10.78 (4.93;2.78-19.23)	10.41 (5.15 ;1-20.18)	-0.37 (0.68; -0.98 -1.71)	0.59
2 weeks	4.16 (5.35; 0 – 21.23)	4.32 (4.93; 0 – 17)	0.16 (0.68; -1.18 – 1.50)	0.81
6 weeks	2.47 (4.59; 0 – 21.23)	2.6 (4.03; 0 – 15.18)	0.12 (0.57; -1.00 – 1.25)	0.83
12 weeks	1.95 (3.83; 0 – 16.68)	2.32 (3.75; 0 – 15.18)	0.37 (0.50; -0.67 – 1.36)	0.46
26 weeks	1.79 (3.57; 0 – 15.68)	2.03 (3.41; 0 – 14.18)	0.24 (0.46; -0.67 – 1.15)	0.60
AUC	2.60 (3.25; 0.15 – 11.78)	2.84 (3.07; 0.11 – 12.00)	0.24 (0.42; 0 – 0.47)	0.58
QALY outcome				
baseline	0.505 (0.194; 0.103 – 0.774)	0.527 (0.192; 0.153-0.795)	-0.022 (0.0263; -0.073-0.029)	0.395
2 weeks	0.974 (0.050; 0.856 – 0.996)	0.928 (0.090; 0.538 – 0.992)	0.046 (0.005; 0.046 – 0.046)	0.0000
6 weeks	0.979 (0.045; 0.931 – 0.997)	0.951 (0.070; 0.853 – 0.994)	0.028 (0.004; 0.028-0.029)	0.0000
12 weeks	0.984 (0.015; 0.944 – 0.997)	0.957 (0.040; 0.811 – 0.992)	0.029 (0.001; -0.178 – 0.236)	0.0000
26 weeks	0.984 (0.017; 0.947 – 0.998)	0.957 (0.040; 0.857 – 0.992)	0.027 (0.001; 0.028 – 0.029)	0.0000
AUC	0.963 (0.021; 0.925 – 0.986)	0.935 (0.043; 0.822– 0.980)	0.028 (0.001; 0.028 – 0.02)	0.0000

Cost analysis

Table 3 shows the results of the cost analysis. The intervention group had lower mean total cost (€ 1165 (sd 3025) than the control group (€ 1304 (sd: 3659) per patient. Most of these costs were sick leave related: respectively € 771 (sd: 2951) per patient for the intervention group and € 893 (sd: 3544) per patient for the control group.

Table 3. Total costs per group over the 26 weeks follow-up period, divided in healthcare costs and other costs (productivity losses and out-of-the-pockets costs). Means and standard deviations (SD) are given.

costs per perspective	SDM&PR group mean (SD;95%CI)	control group mean (SD;95%CI)	p-value
healthcare costs	374 (474; 0 – 1293)	393 (437; 0 – 1506)	0.75
other costs¶	791 (2953; 0 – 5141)	911 (3561; 0 – 9754)	0.81
total cost	1165 (3025; 4 – 5972)	1304 (3659; 0 – 11075)	0.76

¶ = patient and family costs and productivity costs.

Cost-effectiveness analysis

Bootstrapping of the costs and effects showed a difference in costs of € 139.44 per patient in favour of the intervention group (95% CI (€ -888.4 - 629.56), with a RMD difference of -0.23 (95% CI -0.52 - 0.98) per patient. (Table 4). This indicates that disability because of non-chronic low back pain was lower in the intervention group than in the control group . (Figure 2) Overall, 73% of the bootstrap replications indicated a better RMD outcome for the intervention group and 65% of the bootstrap replications pointed at lower costs for the intervention group.

Table 4. Incremental cost-effectiveness and cost-utility analysis on imputed dataset. Distributions of the bootstrap replications over the four quadrants of the cost-effectiveness plane, from a healthcare and societal perspective and the percentage of cost-effectiveness at a willingness to pay (WTP) of 20.000 €/QALY.

	Δ cost* (€) (95% CI)	Δ effect (95% CI)	distribution (%) cost effectiveness plane (quadrant)				percentage (%) cost-effective
			North east‖	South east†	South west¥	North west‡	
incremental cost-effectiveness							
healthcare costs	-18.95 (-121.71-89.14)	-0.23** (-0.47 - 0.95)	24%	48%	15%	13%	
total costs	-139.44 (-888.45-629.56)	-0.23** (-0.52 - 0.98)	21%	52%	13%	14%	

Table 4. (Continued)

	Δ cost* (€) (95% CI)	Δ effect (95% CI)	distribution (%) cost effectiveness plane (quadrant)				percentage (%) cost-effective
			North east‖	South east†	South west¥	North west‡	
incremental cost-utility							
healthcare costs	-18.95 (-121.71-89.14)	0.03*** (0.03-0.04)	35%	65%	0%	0%	100%
total costs	-139.44 (-888.45-629.56)	0.03*** (0.03-0.04)	36%	64%	0%	0%	95%

Δ = difference, * = average costs per patient, ** = effect measured in difference in Area Under the Curve (AUC) in Roland Morris (RMD) outcomes per patient over 26 weeks, *** = effect based on mapping the difference in AUC of RMD outcomes to QALYs, ‖ = North East accounts for higher costs and more effects, † = South East reflects more effects and costs savings (dominance), ‡ = North West reflects less effects and higher cost savings (inferiority), ¥ = South West represents less effects and costs savings.

Cost utility estimation

The mapping into QALYs resulted in a significant difference in favour of the SDM&PR group. (Table 4) Using a threshold value of € 20.000 per QALY (which is the general accepted threshold in the Netherlands), the probability that SDM&PR is cost-effective is 100% from a healthcare perspective and 95% from a societal perspective.

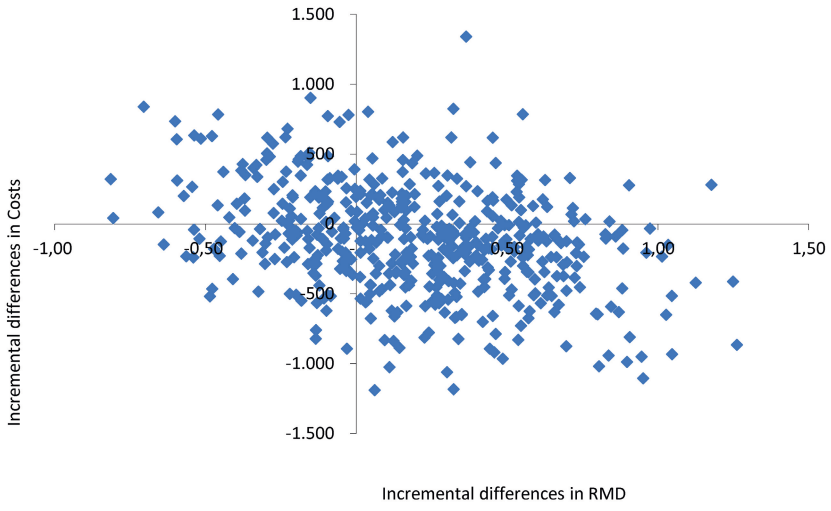


Figure 2. Cost effectiveness plane of the total cost in imputed dataset plotted against the effects on Roland Morris (RMD) outcome (Area Under the Curve (AUC) approach).

Sensitivity analysis

The sensitivity analysis with complete cases only, did not change our conclusions. The complete case analysis on RMD outcomes did not change the significant effect on QALYs (difference between groups of 0.0244). The ICER from a healthcare perspective was € 309.73/QALY and € 11 337.24/QALY from a societal perspective.

DISCUSSION

We studied the cost effectiveness of shared decision-making followed by positive reinforcement (SDM&PR) in a cluster randomised trial in 226 patients with non-chronic low back pain. We found small improvements in disability scores, as measured with the Roland Morris Disabilities Questionnaire (RMD), in favor of the intervention group after 26 weeks of follow-up. The bootstrap analysis pointed at better patient outcomes in all samples, and cost savings in 65% of the samples. As it is unknown what society is willing to pay for any improvement in RMD score, a formal cost-effectiveness analysis with this clinical outcome is non-informative to health policy makers. We therefore translated the RMD scores to Quality Adjusted Life Years (QALY), using a validated mapping function that translates RMD scores to QALY scores. Following this approach, we observed better QALY outcomes and lower costs for the intervention. At a societal threshold value of € 20.000 per QALY, 95% of the simulations indicated that the intervention was cost-effective.

Strengths

The results of this study can be considered well generalizable because we used a piggyback design alongside a cluster randomised controlled trial, collecting real-life data and comparing an intervention with care-as-usual. Data on healthcare consumption, either reimbursed by health insurance or out-of-pocket payments for complementary therapies, were derived from patients' questionnaires and GP electronic medical records, with the more substantial cost figures, such as those related to hospital visits and GP visits, based on the latter data source.

Limitations

Although the proportion of missing data is low for the data extracted from the GP-registration system, this was not the case for the data extracted from patients' questionnaires, especially absenteeism. For absenteeism the amount of missing data was significantly different between groups, which might have induced bias. Results of our cost analysis from the societal perspective should therefore be interpreted with caution although missing data were imputed using a model that used baseline characteristics including illness perceptions that are associated to recovery.

Furthermore, the cost-utility analysis was not based on patient derived utilities but on a mapping procedure, transferring RMD scores to QALYs. The significant difference in QALYs as observed should thus be interpreted with caution, because it is based on a

small difference in RMD score only. The mapping function used in this study is estimated on a solid basis and externally validated on populations that resemble our patient population with regard to main characteristics, but slight cultural differences cannot be excluded. On the other hand, Jellema found a similar difference in QALY's (p=0.02) in a trial on non-chronic low back pain patients who received a minimal psychosocial intervention in general practice in the Netherlands on data elicited directly from patients.(29) A plausible explanation for the significant positive QALY difference might be that QALY's reflect more domains of improvement for non-chronic low back pain than disabilities alone. We advise future researchers to implement not only disease-specific outcomes in trials estimating effects of patient centred interventions, but to also incorporate general health measures that facilitate the calculation of QALY's to be used in cost-utility analysis, such as the EQ-5D questionnaire.

Over the last decades, there is an increasing tendency to incorporate the patient perspective more explicitly in clinical decision-making. (30) The thrive to implement the patients' perspective in healthcare decision-making should not be based on the argument that patients' recovery benefits from SDM, or the argument that it saves costs for healthcare or society. Those claims are still not proven and this study does not provide conclusive evidence on this issue either. (15,31) In a review on cost-effectiveness of patient decision support interventions, none of the included seven studies found increased expenditure but significant effects were only found in studies with considerable risk of bias. (15) Hence, so far the argument to include patients in decision-making is merely ethically, i.e. that patients have a right to be equal participants in their care decisions involving their own health. (30) In general, studies have shown that both GPs and patients are satisfied with SDM. (32,33)

The societal savings of interventions for patients suffering from low back pain are merely based on productivity losses. (4) At an average consultation duration of 10.5 minutes, three minutes of additional consultation (28%), as we found in our study, would be an immense pressure on the time schedule of GPs. (34) In the Dutch healthcare system, the reimbursement of one consultation (€ 9.00) is fixed up to twenty minutes per consultation in addition to a fixed reimbursement of € 56.56 per patient per year in 2010. So, this study typically shows the difference between a societal and a healthcare perspective, with gains and losses affecting different parties.

GPs in our SDM&PR training mentioned during the training sessions that they felt responsible to prevent unnecessary societal spendings. This argument hindered them in agreeing on patients' preferred therapies (i.e., physiotherapy for acute complaints) that did not conform to the guideline, although one day loss of productivity (€ 243.25 in 2010) outweighed six sessions of physiotherapy (the average for non-chronic low back pain complaints, reimbursed for by health insurance at € 27.50 per session). (35) In the UK BEAM Trial, the addition of any therapy of the two offered above those recommended in clinical guidelines was cost-effective, from a healthcare and from a societal perspective. (36) Lin confirmed this conclusion in a GP cost effectiveness review, comparing other interventions for patients suffering from non-chronic low back

pain. (12) This knowledge could help GPs to not overestimate cost effects of 'guideline adherence' and still respect patients' preferences in decision-making.

CONCLUSION

Shared decision-making followed by a positive reinforcement of the chosen therapy (SDM&PR) did not lead to clinically relevant better outcomes or decreased costs but seemed cost-effective in the bootstrap analysis, after conversion of clinical outcome scores to utility scores. From a societal perspective, costs of a prolonged consultation duration due to increased patient involvement may offset reductions in productivity losses. Future research on effects of interventions to increase patient participation or on non-chronic low back pain recovery improvement, should include more general health outcomes like QALYs. The tentative results of mapping the disability effects of our trial to QALYs, suggest that SDM&PR may be cost-effective at acceptable threshold levels for cost-effectiveness.

REGISTRATION

The study protocol was registered in Dutch Trial Register under nr 1960.

ETHICAL APPROVAL

The study was exempted from full assessment by the Institutional Review Board of the Ethics Committee of the University Medical Centre of Utrecht.

CONFLICTS OF INTEREST

The authors declare that they have no competing interests.

ACKNOWLEDGEMENTS

The authors wish to thank all patients and GPs who voluntarily participated in the trial. Specific contributions Jozien Bensing and William Verheul: study design. Sample acquisition, Dinette Aghterhuis, Margan Essed, Emily Fick, Mijke van Gijn, Joost Janssen, Lisanne Louisse, Annemarie Schatsnabel, Maartje Vogelaar, Inge van Weeghel and Rhea Wilken.

CONTRIBUTORS

Data collecting Lisanne Louise, Annemarie Schachtschnabel, Rik van der Lans, Marcelle Uiterwijk and AS. Creating the pilot-study GW, JU, MN and AS. Data-analysis GW, Nanne Bos and RN. Drafting the manuscript: NW, GW and AS. All authors had full access to

all the study data and take responsibility for the integrity of the data and the accuracy of the data analysis. All authors were involved in critical review of the manuscript and have seen and approved the final version.

REFERENCES

1) Bons SCS, Borg, M. A. J. P., Van den Donk M, Koes BW, Kuijpers T, Ostelo, R. W. J. G., et al. NHG-Standaard Aspecifieke lagerugpijn (Tweede herziening). 2017; Available at: https://www.nhg.org/standaarden/samenvatting/aspecifieke-lagerugpijn. Accessed 05/19, 2017.
2) Maher C, Underwood M, Buchbinder R. Non-specific low back pain. Lancet (London, England) 2017;389:736-747.
3) Murray CJ, Vos T, Lozano R, Naghavi M, Flaxman AD, Michaud C, et al. Disability-adjusted life years (DALYs) for 291 diseases and injuries in 21 regions, 1990-2010: a systematic analysis for the Global Burden of Disease Study 2010. Lancet 2012 Dec 15;380(9859):2197-2223.
4) van Tulder MW, Koes BW, Bouter LM. A cost-of-illness study of back pain in The Netherlands. Pain 1995 Aug;62(2):233-240.
5) van Tulder MW, Suttorp M, Morton S, Bouter LM, Shekelle P. Empirical evidence of an association between internal validity and effect size in randomized controlled trials of low-back pain. Spine (Phila Pa 1976) 2009 Jul 15;34(16):1685-1692.
6) Thomas KB. General practice consultations: is there any point in being positive? Br Med J (Clin Res Ed) 1987 May 9;294(6581):1200-1202.
7) Sanders AR, Bensing JM, Essed MA, Magnee T, de Wit NJ, Verhaak PF. Does training general practitioners result in more shared decision making during consultations? Patient Educ Couns 2017 Mar;100(3):563-574.
8) Sanders ARJ, Bensing JM, Magnee T, Verhaak P, de Wit NJ. The effectiveness of shared decision-making followed by positive reinforcement on physical disability in the long-term follow-up of patients with nonspecific low back pain in primary care: a clustered randomised controlled trial. BMC Fam Pract 2018 Jun 28;19(1):102-8.
9) Dowie J. Why cost-effectiveness should trump (clinical) effectiveness: the ethical economics of the South West quadrant. Health Econ 2004 May;13(5):453-459.
10) Dowie J. Research implications of science-informed, value-based decision making. Int J Occup Med Environ Health 2004;17(1):83-90.
11) Foster NE, Dziedzic KS, van der Windt, D. A., Fritz JM, Hay EM. Research priorities for non-pharmacological therapies for common musculoskeletal problems: nationally and internationally agreed recommendations. BMC Musculoskelet Disord 2009 Jan 9;10:3-3.
12) Lin CW, Haas M, Maher CG, Machado LA, van Tulder MW. Cost-effectiveness of general practice care for low back pain: a systematic review. Eur Spine J 2011 Jul;20(7):1012-1023.
13) Veroff D, Marr A, Wennberg DE. Enhanced support for shared decision making reduced costs of care for patients with preference-sensitive conditions. Health Aff (Millwood) 2013 Feb;32(2):285-293.
14) O'Connor AM, Bennett CL, Stacey D, Barry M, Col NF, Eden KB, et al. Decision aids for people facing health treatment or screening decisions. Cochrane Database Syst Rev 2009 Jul 8;(3)(3):CD001431.
15) Walsh T, Barr PJ, Thompson R, Ozanne E, O'Neill C, Elwyn G. Undetermined impact of patient decision support interventions on healthcare costs and savings: systematic review. BMJ 2014 Jan 23;348:g188.
16) Patel S, Ngunjiri A, Hee SW, Yang Y, Brown S, Friede T, et al. Primum non nocere: shared informed decision making in low back pain--a pilot cluster randomised trial. BMC Musculoskelet Disord 2014 Aug 21;15:282-282.
17) Edwards A, Elwyn G. Inside the black box of shared decision making: distinguishing between the process of involvement and who makes the decision. Health Expect 2006 Dec;9(4):307-320.

18) Kolb AY, Kolb DA. Learning Styles and Learning Spaces: Enhancing Experiential Learning in Higher Education. Academy of Management Learning & Education 2005;Vol. 4(2):193-212.
19) Elwyn G, O'Connor AM, Bennett C, Newcombe RG, Politi M, Durand MA, et al. Assessing the quality of decision support technologies using the International Patient Decision Aid Standards instrument (IPDASi). PLoS ONE 2009;4(3):e4705.
20) Schers H, Wensing M, Huijsmans Z, van Tulder M, Grol R. Implementation barriers for general practice guidelines on low back pain a qualitative study. Spine (Phila Pa 1976) 2001 Aug 1;26(15):348.
21) Broadbent E, Petrie KJ, Main J, Weinman J. The brief illness perception questionnaire. J Psychosom Res 2006 Jun;60(6):631-637.
22) Brouwer S, Kuijer W, Dijkstra PU, Goeken LN, Groothoff JW, Geertzen JH. Reliability and stability of the Roland Morris Disability Questionnaire: intra class correlation and limits of agreement. Disabil Rehabil 2004 Feb 4;26(3):162-165.
23) Hakkaart-van Roijen L, Tan SS, Bouwmans CAM. Handleiding voor kostenonderzoek: methoden en referentieprijzen voor economische evaluaties in de gezondheidszorg. College voor zorgverzekeringen. Geactualiseerde versie 2010. 2011; Available at: https://www.zorginstituutnederland.nl/binaries/content/documents/zinl-www/documenten/publicaties/overige-publicaties/1007-handleiding-voor-kostenonderzoek/Handleiding+voor+kostenonderzoek.pdf.
24) Farmacotherapeutisch Kompas. 2014; Available at: https://www.farmacotherapeutischkompas.nl/. Accessed December 27, 2014.
25) Deveugele M, Derese A, De Bacquer D, van den Brink-Muinen A, Bensing J, De Maeseneer J. Consultation in general practice: a standard operating procedure? Patient Educ Couns 2004 Aug;54(2):227-233.
26) Schooten S, van Kemenade S. Tarievenmonitor geeft inzicht in uiteenlopende waarneemkosten huisartsen. Medisch Contact 2011;20:1281-1283.
27) Foster NE, Bishop A, Thomas E, Main C, Horne R, Weinman J, et al. Illness perceptions of low back pain patients in primary care: what are they, do they change and are they associated with outcome? Pain 2008 May;136(1-2):177-187.
28) Madan J, Khan KA, Petrou S, Lamb SE. Can Mapping Algorithms Based on Raw Scores Overestimate QALYs Gained by Treatment? A Comparison of Mappings Between the Roland-Morris Disability Questionnaire and the EQ-5D-3L Based on Raw and Differenced Score Data. Pharmacoeconomics 2017 May;35(5):549-559.
29) Jellema P, van der Roer N, van der Windt, D. A., van Tulder MW, van der Horst, H. E., Stalman WA, et al. Low back pain in general practice: cost-effectiveness of a minimal psychosocial intervention versus usual care. Eur Spine J 2007 Nov;16(11):1812-1821.
30) Salzburg Global Seminar. Salzburg statement on shared decision making. BMJ 2011 Mar 22;342:d1745.
31) Legare F, Adekpedjou R, Stacey D, Turcotte S, Kryworuchko J, Graham ID, et al. Interventions for increasing the use of shared decision making by healthcare professionals. Cochrane Database Syst Rev 2018 Jul 19;7:CD006732.
32) Edwards A, Elwyn G, Hood K, Atwell C, Robling M, Houston H, et al. Patient-based outcome results from a cluster randomized trial of shared decision making skill development and use of risk communication aids in general practice. Fam Pract 2004 Aug;21(4):347-354.
33) Edwards A, Elwyn G. Involving patients in decision making and communicating risk: a longitudinal evaluation of doctors' attitudes and confidence during a randomized trial. J Eval Clin Pract 2004 Aug;10(3):431-437.
34) Noordman J, Verhaak PF. Unpublished data: Mean duration of 808 GP consultations by 40 GPs is 10,5 minutes in 2007-2008 in Nivel library.

35) Nivel, Netherlands institute for health services and research. Kortdurende en langdurige klachten bij fysiotherapie, 2010. 2012; Available at: http://www.nivel.nl/kortdurende-en-langdurige-klachten-bij-fysiotherapie-2010, 2012.
36) UK BEAM Trial Team. United Kingdom back pain exercise and manipulation (UK BEAM) randomised trial: cost effectiveness of physical treatments for back pain in primary care. BMJ 2004 Dec 11;329(7479):1381.

APPENDIX

Appendix 1 Resource costs (€/unit), frequencies of usage, mean and standard deviation (SD) and mean total costs in Euro (€) during 26 weeks after the back pain consultation. Costs were valued for 2010.

		Intervention Group (N=112)			Control Group (N=114)		
Cost item	**unit cost (€/unit)**	**number¶**	**mean (SD)¶**	**total costs in € mean (SD)†**	**number¶**	**mean (SD)¶**	**total costs in € mean (SD)†**
HEALTHCARE COSTS							
Medication costs prescribed by the GP	-	74	0.66 (0.48)	105.99 (373.89)	83	0.73 (0.45)	123.82 (315.52)
General Practice costs							
Email contact	4.5	1	0.01 (0.09)	0.04 (0.43)	0	0.00 (0.00)	0.00 (0.00)
Phone Contact	14.18	39	0.63 (1.12)	8.99 (15.92)	47	0.96 (1.50)	13.56 (2.29)
Practice visit	28.36	76	1.74 (1.99)	49.38 (56.43)	79	1.74 (2.10)	49.26 (5.66)
Double practice visit	56.72	20	0.31 (0.77)	17.73 (43.76)	21	0.26 (0.65)	14.93 (37.01)
Visit at patient's home address	43.55	3	0.04 (0.23)	1.56 (10.00)	5	0.06 (0.31)	2.67 (13.32)
Double visit at patient's home address	87.11	0	0.00 (0.00)	0.00 (0.00)	1	0.01 (0.09)	0.76 (0.16)
Other costs (GP)	-	15	0.13 (0.34)	4.49 (15.11)	11	0.10 (0.30)	4.07 (16.27)
General Practice costs in out of office hours							
Phone contact	24.23	1	0.01 (0.09)	0.22 (2.29)	3	0.04 (0.31)	1.06 (7.49)
Practice visit	100.16	1	0.01 (0.09)	0.89 (9.46)	2	0.02 (0.13)	1.76 (13.21)
Visit at patient's home address	150.25	0	0.00 (0.00)	0.00 (0.00)	3	0.03 (0.16)	3.95 (24.16)
Outpatient clinic costs							

		Intervention Group (N=112)			Control Group (N=114)		
Cost item	**unit cost (€/unit)**	**number¶**	**mean (SD)¶**	**total costs in € mean (SD)†**	**number¶**	**mean (SD)¶**	**total costs in € mean (SD)†**
Hospitalisation (days)	462.88	6	0.05 (0.20)	21.07 (91.86)	5	0.05 (0.22)	21.35 (100.63)
Emergency	152.94	5	0.07 (0.35)	10.92 (53.20)	7	0.07 (0.29)	10.73 (44.20)
Interventions in outpatient clinic	-	4	0.04 (0.19)	7.12 (3.44)	4	0.04 (0.18)	7.11 (44.96)
Outpatient clinic visits	72.93	28	0.45 (0.98)	32.56 (71.83)	24	0.3 (0.87)	28.79 (63.34)
Radiology	-	18	-	10.31 (31.77)	18	-	9.25 (26.27)
Laboratory costs							
Biochemistry	17.56	22	0.25 (0.55)	4.39 (9.58)	20	0.25 (0.62)	4.31 (10.85)
Microbiological	32	4	0.04 (0.19)	1.14 (5.97)	7	0.06 (0.24)	1.96 (7.72)
Pathological	60.45	16	0.23 (0.79)	14.03 (39.56)	10	0.11 (0.41)	6.89 (25.01)
Other healthcare costs							
Other specialist care	-	0	0.00 (0.00)	0.00 (0.00)	2	0.02 (0.13)	4.60 (34.55)
Other radiology	-	10	0.09 (0.29)	6.55 (29.51)	6	0.05 (0.22)	2.24 (11.08)
Other caregivers	-	4	0.04 (0.19)	1.63 (8.70)	6	0.05 (0.22)	1.81 (7.89)
medical device	-	1	0.01 (0.09)	1.07 (11.34)	6	0.05 (0.22)	3.12 (16.71)
Basic psychologic treatment	-	1	-	2.17 (22.87)	2	-	4.80 (43.27)
Subtotal healthcare costs				383.89 (475.69)			401.03 (440.48)
OTHER COSTS							
Productivity costs							
Sick leave (days)	243.25	27	3.17 (12.13)	770.98 (2951.28)	23	3.67 (14.57)	892.97 (3543.95)

		Intervention Group (N=112)			Control Group (N=114)		
Cost item	**unit cost (€/unit)**	**number¶**	**mean (SD)¶**	**total costs in € mean (SD)†**	**number¶**	**mean (SD)¶**	**total costs in € mean (SD)†**
PATIENT AND FAMILY COSTS							
Physiotherapeutic treatment	163.21	47	-	79.92 (107.08)	52	-	83.29 (104.36)
Postural therapy	35.45	8	0.07 (0.26)	2.53 (9.17)	5	0.04 (0.18)	1.26 (6.55)
Acupuncture	56.47	1	0.01 (0.09)	0.50 (5.34)	1	0.01 (0.09)	0.00 (0.02)
Other free accessible therapies*	-	15	-	9.54 (25.68)	7	-	5.56 (29.24)
Subtotal patient and family costs				10.04 (26.15)			10.36 (51.94)
Total costs				1164.92 (3025)			1304.36 (3659)

¶ = non-imputed data, † = imputed data, * = other free accessible therapies are i.e. specialised psychological therapy, massages, podiatry, haptonomy, herbal therapy, Shiatsu and therapeutic sports activities.

CHAPTER 7

General Discussion

DISCUSSION

In this thesis, we studied how to maximise placebo effects in doctor-patient communication in general practice in order to increase patients' recovery from symptom-based illnesses. These are defined as commonly-accepted composites of symptoms that are treated as such, without a known pathophysiological substrate. To operationalise this aim, shared decision-making followed by a positive reinforcement (SDM&PR) by the GP was chosen as the communicative instrument for managing aspecific non-chronic low back pain.

In our clustered randomised trial, training in SDM&PR for GPs did not increase recovery from non-chronic low back pain compared to usual care, despite the fact that we detected more SDM behaviour in GPs who had been trained. Even in cases of enhanced SDM, limitations related to non-chronic low back pain did not decrease significantly compared to less or no SDM. However, despite the observed improvement, our trained GPs did not attain a sufficient level of SDM. As a result, patients' beliefs – including positive outcome expectations regarding preferred therapeutic options – were not reinforced and may even have been ignored. This low intervention fidelity might be the dominant explanation for the lack of effect on patients' recovery.

Despite the lack of clinical superiority, costs from both the healthcare and social perspectives seemed to be lower after SDM&PR, although non-significant. After mapping restrictions to QALYs, SDM&PR was found likely to be cost-effective in patients with non-chronic low back pain in general practice.

Reflections on our study questions

The main aim of this thesis was to assess to what extent patients' recovery from a symptom-based illness like non-chronic low back pain can benefit from training GPs in optimum use of SDM&PR.

This was operationalised in several study questions

1. What is known from literature about the relationship between enhanced patient participation and recovery from symptom-based illnesses in primary care, either from the perspective of the patient or in terms of disease outcomes?

Before deciding on the final intervention, we explored existing literature on which interventions on symptom-based illnesses in primary care might be effective in increasing patient participation. This was expected to be an intermediate and necessary step for an intervention to stimulate placebo effects. (1) Because the concept of 'patient participation' was poorly defined and because of the low quality of evidence in the studies, we did not identify convincing evidence for answering this question. The conclusion from 2013 that these concepts can be deemed 'fuzzy' (i.e. everyone understands what is meant by the concept generally, but the operationalisation is not defined precisely) has not changed. In 2006, Entwistle suggested a broader conceptual framework that acknowledges that patients can be involved not only because of what

they say and do to influence a decision, but also by virtue of what they think and feel about their roles, efforts and contributions to decision-making and their relationships with their clinicians. (2) Although this concept has been receiving more attention recently, a scoping review in 2021 of reviews on person-centred care (a comparable but not identical concept that focuses on increasing the patients' perspective in care) concluded that there is still no clear definition or model and that more personalised outcomes measured by mixed methods are suggested. (3)

We incorporated two pillars of placebo effects in the intervention: manipulating patients' expectations (patients' positive expectations and reinforcement of positive expectations) and reducing stress in the patient-professional relationship (by active listening, asking open-ended questions and asking patients about their preferred level of participation during the decision-making process). (4)

In 2009, the year in which the trial was developed, the knowledge that placebo effects are elicited by positive expectations had not yet developed. Even though it was generally accepted that raising positive expectations could benefit patients, it had not yet been proven. (5) Nowadays, more than 10 years later, there is evidence that placebo effects can be increased not only by positive expectations in the patients but also by statements by the healthcare professional that the therapy will benefit the patient, resulting in a medium to large estimated effect size on patients' non-chronic pain relief. (6,7)

We expected that using open-ended questions, a behaviour included as part of SDM training, is usually interpreted by patients as interest and increases trust. (8) Trust is an important pillar of the patient-professional relationship and can thus increase patient participation in decision-making and induce placebo effects, as we intended to prove in this study. (2,9,10) Nowadays SDM is even promoted in general practice for increasing placebo effects. (9) The argumentation is surprisingly not based on formulating open-ended questions and increasing trust but rather as a tool for raising expectations. (9)

2. What effects are there on communicative behaviour (SDM, positive reinforcing) by GPs after training them in SDM&PR compared to untrained GPs?

In our study, untrained GPs achieved an average score of 23.7 on the OPTION scale (0-100) and trained GPs scored 38.5. (11,12). Although the difference was statistically significant, 38.5 does not yet reflect half of the intended SDM behaviour. The effect of our training programme was comparable to training effects in trials investigated in a systematic review in 2013 by Couët *et al.* This was despite the fact we included mainly GPs in our trial from the academic network, who might be assumed to perform better in communication. (12,13) Similarly to Couët *et al.* and others, we observed that informative behaviour increased but actual patient participation in decision-making did not, despite a significant increase in the level of autonomy in decision-making. (12-15) However, this subjective autonomy was still insufficient to reach a *shared* decision. As a result, a substantial placebo effect based on increased trust, participation and implementation of patient preferred treatment choices was not achieved. (12)

Although exploring patients' experiences is explicitly advised in the Dutch guideline on low back pain and integrated in our training, we observed that patients' expectations and preferences were not taken into account in the decision. (12,16). On the contrary, patients' positive expectations were sometimes rather downgraded and interpreted as misconceptions when therapeutic options were not recommended in the guideline due to being 'only' a placebo effect. We confirmed an earlier conclusion that healthcare professionals are willing to incorporate patients' preferences in SDM, as long as there are no medical contraindications. (17) This raises the interesting question of whether a placebo effect can be a medical contraindication to reinforce patients' positive expectations.

When patients' positive expectations are not implemented in the therapeutic plan, GPs might even reduce patients' autonomy, hampering not only the *cognitive* component of the intended placebo intervention but also an *affective* element. (18) The increased level of autonomy in decision-making, as we found in our study, was insufficient to deem this behaviour to be SDM and suggests that patients' preferred therapeutic plans were not implemented in treatment decisions. (12) The low intervention fidelity within our trial can be considered a major cause of lack of additional effect of the intervention beyond that of usual care.

3. What are the effects of SDM and positive reinforcement by the GP (SDM&PR) compared to usual care on recovery rates in patients suffering from a new episode of non-chronic low back pain?

Our studies did not demonstrate convincing evidence that patients' health benefited from our intervention; neither in the proof of concept study nor in the trial. (19,20). This is in line with the conclusion of a systematic review and meta-analysis, which found insufficient evidence for beneficial effects in patients suffering from non-chronic low back pain. (21) However, a Cochrane review in 2019 on recovery of low back pain demonstrated (with evidence of moderate quality) that future work participation was strongly related to *patients' individual expectations.* For clinical recovery outcomes, however, such as pain or restrictions due to low back pain, the evidence was low. (22) Finally the previously mentioned systematic review by Burgers *et al.* concluded that although process outcomes of interventions aiming to increase *patient participation* are usually positive, the results on health outcomes are not convincing. (3)

In summary: in line with earlier reports, we could not detect any beneficial effects on recovery from non-chronic low back pain after training GPs in SDM&PR in comparison with usual care. We did notice positive effects at various points in time, but these effects were neither significant nor clinically relevant. As discussed in the introduction, the placebo effect is better described by the term 'contextual effects' and the mechanisms behind it are multi-factorial. This might explain why it seems easier to prove an effect on a single process outcome than on a complex outcome such as clinical recovery, which is influenced by many process effects. It might be plausible that a positive process effect shifts a delicate balance of all contextual effects resulting in less placebo effects or even

nocebo effects. (9,21) The effects we found were not substantial enough to warrant a role for placebo effects and might even have introduced nocebo effects, leaving too little room for recovery.

4. What are the effects of patient participation and SDM on recovery from non-chronic low back pain and do these effects differ when perceived from the perspectives of the patients, the professionals or external observers?

As mentioned above, we could not detect any effect in the proof of concept study, not even from the perspective of patients who actually experience all context effects. Although there was uncertainty as to what extent patients want to play an active role in healthcare decisions in 2010, nowadays it is beyond doubt that patients want to play a proactive role in most healthcare decisions. (17,23). However, the effects of SDM as experienced by patients on recovery are still not convincing. (24) In our study, patients achieved a high level of involvement in decision-making (78.03; sd 36.73, on a scale from 0 to 100). This might have hampered any discriminative effect on recovery because of a ceiling effect. This is a problem frequently encountered in studies when patients are questioned about participation. (25,26) Nonetheless, other explanations are possible too. As mentioned above, placebo effects are multifactorial and Shay *et al.* stated that when patients are asked to define SDM they refer to a *personalised recommendation by the physician* as well as sharing information, an open-minded and respectful atmosphere and patient self-advocacy. (23) In our attempt to foster patients' preferences based on positive expectations, we advised the GP trainees not to formulate their recommendation. It is unclear whether this might have introduced nocebo effects but it is conceivable.

Finally, as discussed in Chapter 5, GPs' reflections on the level of SDM as implemented were related to recovery. This might suggest that many contextual effects influence recovery from symptom-based illnesses like non-chronic low back pain. It is known from the literature that GPs' expectations are related to patients' recovery. (27) These expectations might have influenced GPs' behaviour and thereby introduced unwarranted communicative effects that affected the intervention.

5. How cost-effective is SDM&PR in patients suffering from non-chronic low back pain?

Although our trial suggested that the intervention might be cost-effective, the relatively high amount of missing data about the 'sick leave' variable and the estimation of QALYs derived from restrictions due to non-chronic low back pain do not allow a firm conclusion (Chapter 6). The most recent review of cost-effectiveness of low back pain in general practice (2011) demonstrates that *from a healthcare perspective* most treatments were less cost-effective than the usual GP care. (28) However, *from a social perspective,* GPs can improve cost-effectiveness by referring their patients for additional services, such as advice and exercise or by providing the services themselves. (28) More recent cost-effectiveness studies focus on single interventions, such as on effects of chiropractic care or acupuncture in non-chronic low back pain (29,30). Although these

studies suggest these are cost-effective, the level of evidence is insufficient to draw convincing conclusions. This ambiguity is also reflected in the recommendations in the 'Choose Wisely' programme of the Dutch government. (31) Despite the absence of firm evidence for cost-effectiveness, the programme suggests that GPs should refrain from routine referral to physiotherapy or X-rays in patients with low back pain. Remarkably, the motivation of professionals not to follow these *illness-specific* recommendations is driven by *contextual factors* like investing in the doctor-patient relationship. (32) In line with the conclusions of the above-mentioned review and the knowledge that placebo effects are multifactorial, this approach might be less effective than healthcare policy makers assume. The conclusion of our study – that cost-effectiveness should be investigated in interventions aiming to reduce costs due to low back pain – remains important.

METHODOLOGICAL CONSIDERATIONS

We aimed to examine a causal relationship between the deployment of placebo-intended behaviour of a GP and health-related outcomes on symptom-based illnesses at a patient level. We therefore developed an RCT. We learned from literature that training physicians demands extra focus on blinding the results to prevent reduced reliability of the results. (1) We therefore chose a cluster design at the GP level (to prevent contamination in the care-as-usual group) and did not tell the control GPs about the presence and type of intervention. To increase placebo-intended behaviour of GPs for the benefit of patients, we focused on training GPs and developed additional material (decision aid and prompt sheet) to help them implement the intended behaviour.

We cannot rule out that some control GPs might have guessed the aim of the intervention or that we selected GPs based on their interest in communication, which might have caused a selection bias. We found higher scores for SDM skills in the control GPs than was common in the Netherlands the year before the trial. (33) Nonetheless, we do not think that familiarity with the intervention of control GPs was an influential issue as the control GPs scored significantly lower than intervention GPs on the OPTION outcomes and all scores were below a moderate level of SDM. We experienced difficulties in motivating GPs to participate. The most frequently mentioned reason for GPs not to participate in the study was time restraints, but we might have recruited more GPs if we had emphasised the experience of healthcare or positive effects on working life more explicitly. (34,35)

We decided to use a widely promoted communication approach (SDM) that was broadly embraced by healthcare professionals but not yet widely adopted in healthcare. (36-38) Healthcare professionals acknowledged the need for increased patient participation, which motivated our participants to participate in the trial and uptake the intervention. (39,40) The lack of scientific knowledge about a not yet widely adopted concept naturally affects the development and performance of the intervention and measurements. (41,42)

Despite considerable efforts (Chapter 4) we only reached half of the predefined sample size (426 inclusions). However, even if we had reached the intended sample size, it is unlikely this would have changed the outcome statistically or clinically. (20) Although the effect was in favour of the intervention-group patients, the effect size was that small that we would have needed thousands of patients to achieve a statistically significant difference, and even then this effect would be not clinically relevant.

The coding system for positive reinforcement was unreliable, leaving us with uncertainty about a behavioural change after training. Scores were low anyhow and from observations of the process steps in the OPTION-12 instrument, we expect that at most the therapeutic choice of the GP and not of the patient was reinforced. (12) The Dutch guideline for GPs on low back pain advises discussing the beneficial natural course of the illness. (16) This was done to a moderate extent, which shows that the positive expectations of patients in particular and positive expectations overall were not fully used to increase placebo effects.

We explored the level to which patients felt they participated in decision-making and the reflections of GPs on how well they used SDM. (19)We did not use a measurement instrument that had been validated by 2010 for the level of patient participation experienced by the patients like e.g. the decisional conflict scale. (25) As far as we know, an instrument to measure the reflections of GPs on SDM behaviour did not exist in 2009. (12,25)

We performed a piggybag design (inclusion of data on expenses in an intervention trial) to measure the cost-effectiveness of the intervention. The possibility of including cost-effectiveness in our study became available after the RCT for clinical effectivity of the intervention was designed. We inquired about costs shortly after the intervention (except for the questionnaire at 26 weeks), which let us retrieve more accurate cost data than inquiries after a more prolonged delay after the intervention.

It is not known how many people want to pay for a decrease in one disability out of the 24 measured in the Roland Morris Disability Questionnaire (RMD). QALYs are used to allow comparison of the costs people are prepared to spend for health gains. In the Netherlands in 2010, it was estimated that people were prepared to pay €20,000 per QALY. We therefor used a validated British method to calculate QALYs from restrictions measured by the Roland Morris Disabilities Questionnaire (RMD). (43) Although we used a method to calculate QALYs from raw RMD-scores that were validated in 2014, a more reliable method has been developed since then using differences from repeated RMD observations, as mappings based on raw RMD data overestimate the EQ-5D-3L health utility gains from interventions that reduce RMD scores. (43,44)

PRACTICAL IMPLICATIONS

Appropriate SDM in clinical practice requires changing the professional paradigm from "I am the doctor so I *have the knowledge and expertise* how to achieve the *best possible health outcome* for the patient" to "I am a consultant with medical expertise and I *help* patients to achieve *the (health-related) goals in life that are important to them*". This is difficult for many doctors, as they experience the medical profession more as an identity than a job. (45) The predominantly medical focus of doctors might explain why they repeatedly state that SDM is only appropriate in 'situations where there is equipoise'. (35,45,46) In a recent study among residents in the Netherlands, their decision-making behaviour appeared strongly affected by their conviction that they *are responsible* for arriving at the correct diagnosis and providing the best evidence-based treatment. (35)

To change *feelings of a professional identity* into *facilitating patients to achieve health-related goals*, more elaborated training of professionals is necessary. Barnhorn *et al.* state that changing the professional identity of GPs demands a process of repetitive encouragement of reflection on behaviour and attitude as one aspect of experienced professionalism. (47) The country-wide campaign in the Netherlands in 2019 in which patients and professionals were encouraged to ask three important questions in healthcare contacts with their GP, proves that this process demands patience and considerable investment. (48) These three questions were not widely adopted yet, as of 2022. (48) On the other hand, increasing amounts of easily accessible decision aids have been introduced in general practice in the Netherlands, e.g. through thuisarts.nl (a widely-used independent online website for health information, developed and maintained by the Dutch College of GPs). (49) These instruments inform patients better and, maybe even more importantly, provide the opportunity to consider the therapeutic options in their personal context before making the decision with their professional. GP trainees and their trainers are repeatedly told about the importance of patient participation in decision-making. (49,50) Nonetheless, the quadruple aim of improving the experience of care, the health of populations, the per capita cost of healthcare and the working lives of HCPs might be even more effective in motivating GPs to change their professional identity. (51-53)

Although there still is debate on the time investments needed to perform SDM, training and acquiring experience demand time and patience. To increase the efficiency of SDM, it would save time if doctors do not focus strongly on transferring information but rather on informing after patients' needs to make a well-considered decision together.

SUGGESTIONS FOR FURTHER RESEARCH

More evidence-based knowledge on the placebo effect and communication could help medical professionals see these effects as instruments with added value in their clinical practice. Future research to increase SDM and patient participation should be based on the fact that behavioural change of physicians cannot be achieved by simple training programmes when the motivation for this stems from altering their professional identity.

CONCLUSION

In the studies in this thesis, we could not confirm that increased patient participation or positive reinforcement by a GP benefited recovery from non-chronic low back pain. GPs increase informative behaviour after training SDM but neglect to inquire about patients' preferences. Professionals and patients might benefit if healthcare professionals were to accept evidence-based placebo mechanisms as instruments that are part of their professional toolkit. Further research should focus on a quadruple aim when developing these interventions.

REFERENCES

1) Sanders AR, van Weeghel I, Vogelaar M, Verheul W, Pieters RH, de Wit NJ, et al. Effects of improved patient participation in primary care on health-related outcomes: a systematic review. Fam Pract 2013 May 22.
2) Entwistle VA, Watt IS. Patient involvement in treatment decision-making: the case for a broader conceptual framework. Patient Educ Couns 2006 Nov;63(3):268-278.
3) Burgers JS, van der Weijden T, Bischoff EWMA. Challenges of Research on Person-Centered Care in General Practice: A Scoping Review. Front Med (Lausanne) 2021 June 24;8:669491.
4) van Vliet LM, van Dulmen S, Mistiaen P, Bensing JM. The placebo effects of good communication. Ned Tijdschr Geneeskd 2016;160:D251.
5) Crow R, Gage H, Hampson S, Hart J, Kimber A, Thomas H. The role of expectancies in the placebo effect and their use in the delivery of health care: a systematic review. Health Technol Assess 1999;3(3):1-96.
6) Peerdeman KJ, van Laarhoven AIM, Keij SM, Vase L, Rovers MM, Peters ML, et al. Relieving patients' pain with expectation interventions: a meta-analysis. Pain 2016 June 01;157(6):1179-1191.
7) Hedderson WC, Dover GC, George SZ, Crow JA, Borsa PA. Expectancy Reduces Symptoms but not Functional Impairment Following Exercise-induced Musculoskeletal Injury. Clin J Pain 2018 January 01;34(1):1-7.
8) Mechanic D, Meyer S. Concepts of trust among patients with serious illness. Soc Sci Med 2000 Sep;51(5):657-668.
9) Lucassen P, Olesen F. Context as a drug: some consequences of placebo research for primary care. Scand J Prim Health Care 2016 December 01;34(4):428-433.
10) Rossettini G, Carlino E, Testa M. Clinical relevance of contextual factors as triggers of placebo and nocebo effects in musculoskeletal pain. BMC Musculoskelet Disord 2018 January 22;19(1):27-8.
11) Elwyn G, Hutchings H, Edwards A, Rapport F, Wensing M, Cheung WY, et al. The OPTION scale: measuring the extent that clinicians involve patients in decision-making tasks. Health Expect 2005 Mar;8(1):34-42.
12) Sanders AR, Bensing JM, Essed MA, Magnee T, de Wit NJ, Verhaak PF. Does training general practitioners result in more shared decision making during consultations? Patient Educ Couns 2017 Mar;100(3):563-574.
13) Couet N, Desroches S, Robitaille H, Vaillancourt H, Leblanc A, Turcotte S, et al. Assessments of the extent to which health-care providers involve patients in decision making: a systematic review of studies using the OPTION instrument. Health Expect 2013 Mar 4.
14) Braddock CH,3rd. The emerging importance and relevance of shared decision making to clinical practice. Med Decis Making 2010;30(5 Suppl):5S-7S.
15) de Mik SML, Stubenrouch FE, Balm R, Ubbink DT. Systematic review of shared decision-making in surgery. Br J Surg 2018 December 01;105(13):1721-1730.
16) Bons SCS, Borg MAJP, Van den Donk M, Koes BW, Kuijpers T, Ostelo RWJG, et al. NHG-Standaard Aspecifieke lagerugpijn (Tweede herziening). 2017; Available at: https://www.nhg.org/standaarden/samenvatting/aspecifieke-lagerugpijn. Accessed 05/19, 2017.
17) van Dulmen S, Roodbeen R, Schulze L, Prantl K, Rookmaaker M, van Jaarsveld B, et al. Practices and perspectives of patients and healthcare professionals on shared decision-making in nephrology. BMC Nephrol 2022 July 21;23(1):258-4.
18) Patel S, Ngunjiri A, Hee SW, Yang Y, Brown S, Friede T, et al. Primum non nocere: shared informed decision making in low back pain--a pilot cluster randomised trial. BMC Musculoskelet Disord 2014 Aug 21;15:282-282.

19) Sanders ARJ, de Wit NJ, Zuithoff NPA, van Dulmen S. The effect of shared decision-making on recovery from non-chronic aspecific low back pain in primary care; a post-hoc analysis from the patient, physician and observer perspectives. BMC Prim Care 2022 February 02;23(1):22-y.
20) Sanders ARJ, Bensing JM, Magnee T, Verhaak P, de Wit NJ. The effectiveness of shared decision-making followed by positive reinforcement on physical disability in the long-term follow-up of patients with nonspecific low back pain in primary care: a clustered randomised controlled trial. BMC Fam Pract 2018 Jun 28;19(1):102-8.
21) van Lennep, J H P A, Trossel F, Perez RSGM, Otten RHJ, van Middendorp H, Evers AWM, et al. Placebo effects in low back pain: A systematic review and meta-analysis of the literature. Eur J Pain 2021 May 29.
22) Hayden JA, Wilson MN, Riley RD, Iles R, Pincus T, Ogilvie R. Individual recovery expectations and prognosis of outcomes in non-specific low back pain: prognostic factor review. Cochrane Database Syst Rev 2019 November 25;2019(11):10.1002/14651858.CD011284.pub2.
23) Shay LA, Lafata JE. Understanding patient perceptions of shared decision making. Patient Educ Couns 2014 September 01;96(3):295-301.
24) Shay LA, Lafata JE. Where is the evidence? A systematic review of shared decision making and patient outcomes. Med Decis Making 2015 January 01;35(1):114-131.
25) Simon D, Loh A, Harter M. Measuring (shared) decision-making--a review of psychometric instruments. Z Arztl Fortbild Qualitatssich 2007;101(4):259-267.
26) Kasper J, Heesen C, Kopke S, Fulcher G, Geiger F. Patients' and observers' perceptions of involvement differ. Validation study on inter-relating measures for shared decision making. PLoS One 2011;6(10):e26255.
27) Kapoor S, Shaw WS, Pransky G, Patterson W. Initial patient and clinician expectations of return to work after acute onset of work-related low back pain. J Occup Environ Med 2006 Nov;48(11):1173-1180.
28) Lin CW, Haas M, Maher CG, Machado LA, van Tulder MW. Cost-effectiveness of general practice care for low back pain: a systematic review. Eur Spine J 2011 Jul;20(7):1012-1023.
29) Blanchette MA, Stochkendahl MJ, Borges Da Silva R, Boruff J, Harrison P, Bussieres A. Effectiveness and Economic Evaluation of Chiropractic Care for the Treatment of Low Back Pain: A Systematic Review of Pragmatic Studies. PLoS One 2016 August 03;11(8):e0160037.
30) Skonnord T, Fetveit A, Skjeie H, Brekke M, Grotle M, Klovning A, et al. Cost-effectiveness analysis of acupuncture compared with usual care for acute non-specific low back pain: secondary analysis of a randomised controlled trial. Acupunct Med 2022 April 01;40(2):123-132.
31) Flinterman, Linda, Schers H, Schermer T. Behandeling van klachten vanuit de lage rug in de eerste lijn. NIHES/Radboud University Medical Center 2019 "Sept, ":1-56.
32) Kool RB, Verkerk EW, Winnemuller LJ, Wiersma T, Westert GP, Burgers JS, et al. Identifying and de-implementing low-value care in primary care: the GP's perspective-a cross-sectional survey. BMJ Open 2020 June 03;10(6):e037019-037019.
33) Meijers MC, Noordman J, Spreeuwenberg P, Olde Hartman TC, van Dulmen S. Shared decision-making in general practice: an observational study comparing 2007 with 2015. Fam Pract 2019 May 23;36(3):357-364.
34) Driever EM, Brand PLP. Education makes people take their medication: myth or maxim? Breathe (Sheff) 2020 March 01;16(1):190338-2019.
35) Driever EM, Tolhuizen IM, Duvivier RJ, Stiggelbout AM, Brand PLP. Why do medical residents prefer paternalistic decision making? An interview study. BMC Med Educ 2022 March 08;22(1):155-2.
36) Legare F, Ratte S, Stacey D, Kryworuchko J, Gravel K, Graham ID, et al. Interventions for improving the adoption of shared decision making by healthcare professionals. Cochrane Database Syst Rev 2010 May 12;5:CD006732.

37) Stacey D, Legare F, Pouliot S, Kryworuchko J, Dunn S. Shared decision making models to inform an interprofessional perspective on decision making: a theory analysis. Patient Educ Couns 2010 Aug;80(2):164-172.
38) Makoul G, Clayman ML. An integrative model of shared decision making in medical encounters. Patient Educ Couns 2006 Mar;60(3):301-312.
39) de Wit NJ, Quartero AO, Zuithoff AP, Numans ME. Participation and successful patient recruitment in primary care. J Fam Pract 2001 Nov;50(11):976.
40) Allaire AS, Labrecque M, Giguere A, Gagnon MP, Legare F. What motivates family physicians to participate in training programs in shared decision making? J Contin Educ Health Prof 2012;32(2):98-107.
41) Scholl I, Loon MK, Sepucha K, Elwyn G, Legare F, Harter M, et al. Measurement of shared decision making - a review of instruments. Z Evid Fortbild Qual Gesundhwes 2011;105(4):313-324.
42) Legare F, Ratte S, Gravel K, Graham ID. Barriers and facilitators to implementing shared decision-making in clinical practice: update of a systematic review of health professionals' perceptions. Patient Educ Couns 2008 Dec;73(3):526-535.
43) Khan KA, Madan J, Petrou S, Lamb SE. Mapping between the Roland Morris Questionnaire and Generic Preference-Based Measures. Value Health 2014 Sep;17(6):686-695.
44) Madan J, Khan KA, Petrou S, Lamb SE. Can Mapping Algorithms Based on Raw Scores Overestimate QALYs Gained by Treatment? A Comparison of Mappings Between the Roland-Morris Disability Questionnaire and the EQ-5D-3L Based on Raw and Differenced Score Data. Pharmacoeconomics 2017 May;35(5):549-559.
45) Barnhoorn PC. Shared decision making seen through the lens of professional identity formation. Patient Educ Couns 2020 July 01;103(7):1446-1447.
46) Elwyn G, Edwards A, Wensing M, Hood K, Atwell C, Grol R. Shared decision making: developing the OPTION scale for measuring patient involvement. Qual Saf Health Care 2003 Apr;12(2):93-99.
47) Barnhoorn PC, Houtlosser M, Ottenhoff-de Jonge MW, Essers GTJM, Numans ME, Kramer AWM. A practical framework for remediating unprofessional behavior and for developing professionalism competencies and a professional identity. Med Teach 2019 March 01;41(3):303-308.
48) Springvloet L, Bos N, Jong de J, Friele R, Boer de D. De 3 goede vragen. Hoe dragen de 3 goede vragen bij aan transparantie en de bestpassende zorg voor patiënten? Nivel/Nihes 2019 july:1-25.
49) van der Weele GM, Baghus A. Welke rol spelen keuzekaarten bij samen beslissen? Huisarts Wet 2022 Nov 11,.
50) Baghus A, Timmerman A, Giroldi E, Van der Kraan J, Van der Weele G, Van der Weijden T. Hoe gebruik je keuzekaarten en keuzehulpen in de spreekkamer? . 2022 Nov 14,;65.
51) Bodenheimer T, Sinsky C. From triple to quadruple aim: care of the patient requires care of the provider. Ann Fam Med 2014 December 01;12(6):573-576.
52) Muller E, Strukava A, Scholl I, Harter M, Diouf NT, Legare F, et al. Strategies to evaluate healthcare provider trainings in shared decision-making (SDM): a systematic review of evaluation studies. BMJ Open 2019 June 21;9(6):e026488-026488.
53) van Veenendaal H, Voogdt-Pruis HR, Ubbink DT, van Weele E, Koco L, Schuurman M, et al. Evaluation of a multilevel implementation program for timeout and shared decision making in breast cancer care: a mixed methods study among 11 hospital teams. Patient Educ Couns 2022 January 01;105(1):114-127.

CHAPTER 8

Nederlandse samenvatting/ English summary

Dankwoord

About the author

List of publications

‘RUGGESPRAAK’: SAMEN BESLISSEN EN POSITIEVE BEKRACHTIGING TER VERSTERKING VAN PLACEBO-EFFECTEN BIJ NIET-LANGDURIGE LAGERUGPIJN

Placebo-effecten, ook wel omschreven als contexteffecten, zijn voor een deel een resultante van de verwachtingen van arts en patiënt, en van de mate van onderling vertrouwen. Dergelijke effecten spelen in de huisartsenpraktijk een aanzienlijke rol bij het herstellen van aandoeningen waarvoor geen afdoende pathofysiologische verklaring aanwezig is, zoals lagerugpijn. Voor huisartsen kan ‘samen beslissen’ een instrument zijn om placebofactoren te introduceren in de consultvoering. Om samen te beslissen is immers vertrouwen tussen arts en patiënt nodig. Daarbij zullen patiënten, als ze zeggenschap hebben over de beleidskeuze, positieve verwachtingen hebben over het gekozen plan van aanpak.

Onze algemene onderzoeksvraag was: in welke mate kan het herstel van patiënten die lijden aan een symptoom-gebaseerde aandoening, zoals niet-langdurige lagerugpijn, gebaat zijn bij de inzet van placebo-effecten?

Hoofdstuk 2 beschrijft een systematisch literatuuronderzoek naar gecontroleerde trials, gepubliceerd voor oktober 2011, die betrekking hebben op interventies in de eerste lijn die patiëntparticipatie tijdens fysieke arts-patiëntcontacten vergroten, en waarvan de uitkomstmaten gerelateerd zijn aan de ziekte of aan de patiënt. We vonden zeven gecontroleerde trials, waarvan twee met symptoom-gebaseerde klachten en vijf met patiënt-gerapporteerde uitkomsten. De studies naar de effecten op noch de ziekte-, noch de patiënt-gerapporteerde uitkomsten waren conclusief. Maar vooral ook het gebrek aan kwaliteit van de trials was opvallend. Mogelijk is dit te wijten aan de complexiteit van het concept ‘patiëntparticipatie’.

Vervolgens is in de hierop volgende studies onderzocht of de zorg voor patiënten met niet-langdurige lagerugpijn verbeterd kan worden door hun ideeën en verwachtingen mee te nemen in het plan van aanpak. In dit kader trainden we 34 huisartsen om samen met de patiënt te komen tot een therapiekeuze die beter aansluit bij de eigen ideeën van de patiënt. Zij konden daarbij gebruik maken van een eenvoudige keuzekaart voor therapeutische opties bij niet-langdurige lagerugpijn die voor dit onderzoek was ontwikkeld. Daarnaast vroegen we de artsen om de patiënt te bevestigen in de gekozen aanpak. De training bestond uit twee sessies van tweeënhalf uur met onder andere rollenspelen en het ontwikkelen van een kattenbelletje met ondersteunende vragen. Alle huisartsen werd gevraagd tien consulten op video op te nemen met patiënten die hen bezochten in verband met een nieuwe episode van niet-langdurige lagerugpijn. Aan de hand van deze video-opnames kregen de artsen feedback over het getrainde gedrag. Ter vergelijking vroegen we 34 andere huisartsen (controlegroep) om eveneens tien consulten op te nemen van patiënten met niet-langdurige lagerugpijn. Deze artsen ontvingen geen training, geen keuzehulp en geen feedback, en deden het consult zoals voor hen gebruikelijk was.

In **hoofdstuk 3** wordt de controle op de kwaliteit van de interventie in de trial beschreven. Van de 34 getrainde huisartsen hebben er 23 video-opnames gemaakt van in totaal 86 consulten met patiënten met niet-langdurige lagerugpijn. Deze opnames hebben we vergeleken met 89 opnames gemaakt door 19 huisartsen uit de ongetrainde groep. Gemiddeld scoorden de getrainde huisartsen zowel significant beter op 'samen beslissen' (gescoord met de OPTION-schaal, een gevalideerd observatie-instrument) als op positief bekrachtigen van de therapiekeuze. De gemiddelde score op de OPTION-schaal (0-100) was in de getrainde groep 38,53 punten (95%-betrouwbaarheidsinterval 35,31 tot 41,74), versus 23,66 punten (95%-betrouwbaarheidsinterval 20,25 tot 27,08) in de controlegroep. Helaas bleek het meetinstrument voor positieve bekrachtiging niet valide. We kwantificeerden de mate van autonomie in de besluitvorming door de impressie van de observator weer te geven op een schaal van 0 (volledig paternalistisch) tot 4 (volledig eigen keuze), waarbij 2 stond voor 'samen beslissen'. In de getrainde groep gaf de observator gemiddeld een score van 1,74 (standaard fout 0,11) aan de patiëntparticipatie, versus 0,86 (standaard fout 0,11) in de controlegroep. De gemiddelde consultduur was 15,8 minuten (standaard deviatie 6,0) in de getrainde groep versus 13,1 minuten (standaard deviatie 4,5) in de controlegroep.

Hoofdstuk 4 beschrijft de resultaten van de geclusterde gerandomiseerde trial. Van de 34 getrainde huisartsen includeerden er 25 in totaal 112 patiënten met niet-langdurige lagerugpijn (de interventiegroep) voor de trial. De controlegroep bestond uit 116 patiënten, geïncludeerd door 22 van de 34 ongetrainde artsen. Alle geïncludeerde patiënten vulden gedurende de eerste 14 dagen na het opgenomen consult een dagboekje in en na 2, 6, 12 en 26 weken een vragenlijst over de mate van pijn, beperkingen en andere relevante informatie. Beide groepen waren vergelijkbaar.

De meeste patiënten herstelden binnen 14 dagen. Op de meeste meetmomenten leek er een klein verschil te zijn in het voordeel van de getrainde groep, maar dit verschil bleek nergens relevant of significant.

Hoofdstuk 5 beschrijft de *proof-of-concept* studie waarin is gekeken of optimale toepassing van de interventie heeft geleid tot beter herstel. Naast de observaties hebben we patiënten en artsen gevraagd in welke mate ze hadden ervaren dat er sprake was van 'samen beslissen'. Van de 176 geïncludeerde patiënten waren er 26 weken na het consult 105 (74%) hersteld. Voor dit onderzoek definieerden we 'herstel' als een pijnscore < 30 (op een visuele analoge schaal van 0-100) en maximaal 3 beperkingen op de Roland-Morris Disability Questionnaire (RMD), een gevalideerde vragenlijst van 24 beperkingen die het gevolg kunnen zijn van lagerugpijn. Noch na 6, noch na 26 weken vonden we een significant verband tussen het herstel en de mate waarin observatoren of patiënten 'samen beslissen' hadden ervaren. Als huisartsen aangaven dat er amper sprake was geweest van samen beslissen, of dat er juist uitgebreider 'samen beslissen' had plaats gevonden, bleek er wel een significante samenhang met een beter herstel na 26 weken. Mogelijk passen huisartsen hun communicatie aan wanneer ze een moeizamer herstel verwachten.

In **hoofdstuk 6** hebben we de kosteneffectiviteit van de interventie ingeschat. In de vragenlijsten op 2, 6, 12 en 26 weken zijn de kosten, waaronder de kosten van ziekteverzuim of vrij-toegankelijke zorg, geïnventariseerd. Daarnaast hebben we kosten berekend van medicatie of zorggebruik zoals geregistreerd in de huisartsinformatiesystemen. Over de gehele periode van 26 weken waren de gezondheidszorgkosten per patiënt in de interventiegroep €50,75 lager dan in de controlegroep.

Het effect op de verandering van de kwaliteit van leven werd berekend met behulp van de RMD gemeten over de gehele periode van 26 weken met een formule die afkomstig is uit een Engels onderzoek. Deze formule berekent de effecten op het aantal levensjaren, gecorrigeerd voor mate van kwaliteit van leven (Quality-Adjusted Life Years of QALYs) vanuit de RMD uitkomsten. Op basis van deze berekening leek investering in interventie (de gehele training van de artsen inclusief de keuzehulp) effectief wat betreft de maatschappelijke kosten voor niet-langdurige lagerugpijn (hoofdzakelijk kosten als gevolg van werkverzuim).

Hoofdstuk 7 vat de uitkomsten van alle onderzoeken samen. We concluderen dat training van huisartsen in samen beslissen en positief bekrachtigen van het plan van aanpak niet leidt tot sneller herstel van patiënten met niet-langdurige lagerugpijn dan begeleiding door niet-getrainde huisartsen. We veronderstellen dat dit komt doordat de getrainde artsen weliswaar meer stappen van het 'samen beslissen'-proces uitvoerden en patiënten meer invloed hadden op het plan van aanpak, maar dat dit nog altijd onvoldoende was om de eigen positieve verwachtingen van de patiënt daadwerkelijk mee te nemen in de gekozen aanpak. Zelfs wanneer observatoren of patiënten van mening waren dat er meer samen beslissen had plaatsgevonden, namen wij geen sneller herstel waar. Ook bij getrainde huisartsen was er nog steeds geen sprake van samen beslissen. We mogen dus niet veronderstellen dat er voldoende positieve bekrachtiging heeft plaatsgevonden.

Mogelijk beschouwen artsen het als minder professioneel om placebo-effecten in te zetten ter bestrijding van klachten bij een patiënt. Dit zou betekenen dat training in samen beslissen vraagt om een discussie met artsen over de professionele invulling van hun vak.

BACK-TO-BACK CONSULTATIONS: SHARED DECISION-MAKING AND POSITIVE REINFORCEMENT TO INCREASE PLACEBO EFFECTS IN NON-CHRONIC LOW BACK PAIN.

Placebo effects, also known as context effects, can be partly explained by expectations and trust between patient and professional. Context effects play an important role in recovery from illnesses that are characterised by the symptoms but do not have a specific pathophysiological substrate. Patient participation in general practice could play a key role in optimising these context effects in for instance low back pain complaints. Trust between patient and professional is after all essential if the former is to participate. Increased participation in deciding on the treatment plan might lead to patients making choices for which they have higher outcome expectations.

We reviewed the existing literature about the effects of increased participation in primary care on recovery in **Chapter 2**. We could not draw firm conclusions because there were not enough studies of acceptable quality.

We then investigated in subsequent studies whether the care for patients suffering from non-chronic low back pain can be improved by incorporating the ideas and expectations of patients in the management plan for their complaint. Thirty-four general practitioners (GPs) were trained to put together a therapeutic plan jointly with the patient (shared decision-making abbreviated into SDM), in which the patients will see more of their own ideas reflected. They were given a simple decision aid about therapeutic options for patients suffering from non-chronic low back pain, developed especially for this trial. After deciding on the treatment plan, GPs were asked to confirm that the therapeutic choice was a good one (positive reinforcement). Overall, the training consisted of two sessions of two and a half hours each including role play and the development of a prompt sheet with supportive questions. GPs were each asked to video-record 10 patients suffering from a new episode of non-chronic low back pain, performing SDM and positive reinforcement (SDM&PR). They received feedback on each video-recorded consultation. Another 34 GPs (the control group) were asked to video-record consultations with patients suffering from non-chronic low back pain for whom they provided the usual care. They did not receive any training, decision aid or feedback and simply provided care as usual.

All participating patients completed questionnaires immediately and thereafter daily during the 14 days after the consultation (diary) and at 2, 6, 12 and 26 weeks after the consultation. In those questionnaires, they were asked how they perceived the consultation had gone (patient perspective) and how fast they recovered from pain or disability due to non-chronic low back pain. GPs all completed questionnaires about the personal characteristics and the level of SDM (GP reflections).

Chapter 3 describes the check on the quality of the intervention. Of all 34 trained GPs, 23 video-recorded a total of 86 consultations with patients suffering from non-chronic ow back pain. These observations were compared against 89 video-recorded consultations by 19 GPs providing the usual care. The trained GPs scored significantly

better on the OPTION scale (0-100), a validated instrument for measuring shared decision-making (SDM), with an average of 38.53 (95%-confidence interval 35.31–41.74). The control group scored an average of 23.66 (95%-confidence interval 20.25–27.08). The trained group scored significantly higher in terms of the positive reinforcement of the chosen therapeutic plan. However, the instrument for measuring positive reinforcement performance was not reliable. The level of autonomy in deciding was quantified by an impression of the observer on a 5-point Likert scale (from 0 = paternalistic to 2 = SDM and 4 = leaving full responsibility to the patient). The trained group showed significantly more patient participation (an average score of 1.74 (standard error 0.11) compared to an average score of 0.86 (standard error 0.11) in the untrained group. The duration of the consultation for trained GPs was 15.8 minutes (standard deviation 6.0), which was significantly longer than when giving the usual care, at 13.1 minutes (standard deviation 4.5).

Chapter 4 describes the clustered randomised trial of the thesis. Of the 34 GPs who were included and trained, 25 included 112 patients suffering from a new episode of non-chronic low back pain for the trial. The patients had similar characteristics to the 116 included by the 22 untrained GPs, although fewer of these patients were absent from work (11% versus 35%).

Most patients recovered with 14 days. Although we saw small differences in favour of the intervention group at most of the measurement moments, these differences were not only non-significant but, even more importantly, non-relevant as well.

In **Chapter 5**, as a proof of concept, the optimal SDM performance (as perceived by patients, GPs and observers) is related to recovery at 6 and 26 weeks after the consultation. At 26 weeks after the consultation, 105 (74%) of all 176 included patients had recovered, where recovery was defined as a pain score of <30 (on a Visual Analogue Scale from 0-100) and fewer than 4 disabilities on a list of 24 disabilities due to low back pain (Roland Morris Disability questionnaire). We could not detect a significant correlation between recovery and level of performance of SDM perceived from the patient's perspective or the observations at either 6 or 26 weeks. High and low scores on the extent to which GPs used SDM were significantly correlated to increased recovery at 26 weeks compared to intermediate level of SDM according to the GPs' reflections. Possibly GPs adapt their communication when they do not expect a quick recovery.

In **Chapter 6**, we estimate cost-effectiveness of the intervention. Costs related to absenteeism from work and freely accessible healthcare services were derived from questionnaires at baseline, 2, 6, 12 and 26 weeks. Other healthcare costs were derived from the GP registration systems. Healthcare costs were €50.75 per patient lower in the intervention group than in the control group over the whole study period of 26 weeks.

RMD outcomes were translated into utilities using a published mapping function, to estimate the costs per Quality-Adjusted Life Year (QALY). Investment in SDM&PR

training for patients suffering from non-chronic low back pain seemed cost-effective from the social perspective.

Chapter 7 summarises all study outcomes and concludes that training GPs in SDM&PR did not increase the recovery of patients suffering from non-chronic low back pain. We hypothesise that although trained GPS performed better in the process steps of SDM and patients' autonomy increased, the performance was still insufficient to have patients' positive expectations about the treatment options that were implemented in the treatment decision. Even when more SDM behaviour was involved (according to the patient or observer), recovery did not increase. However, observers did not detect at least an average level of SDM at all. We therefore do not think that patients' positive treatment expectations were reinforced.

It might be the case that healthcare providers consider it less professional to use placebo effects to alleviate patients' suffering. If this is true, it would imply that SDM training needs a discussion about the professional's role, considering patient participation and placebo effects.


Jockie Senders
Alma van der Pol
Rene Roelofs
Anne de Groote
Gerard Mientjes
Bert Pothaar
Kim van Wijck
Wim Alsem
Marielle de Wilt
Geurt Essers
Emmelien ten Hoor
Jan Simon Bakker
Nynke Wildervanck
Marian van den Brink
(Ma)rinus Grijzenhout
Maud de Bruijn
Erik en Romkje
Niek de Wit
Sandra van Dulmen
Rabin Neslo
Leontine Groen-van de Ven
Annelies van Randwijk
Leo Vromen
Adri de Heij
Joost Nijssen
Sabine Uijl
Anne Margreet Stiggelbout
Caroline Smits
Marijke Span
Janny Oskam
Liesbeth Kuhlemaijer
Alle patienten
Haske van Veenendaal
Marjo Broos
Sjoerd Siwart
Peep Stalmeier
Bart Schudel
Anouk Baghus
Frans Schormans
Durien Zwart
Anne Denteneer
Hans Burggraaf
Annemarie Schachtschabel
Mark van Dijk
Marloes Minnaard
Marleen Buijs
Apeldoornse deernen
Maartje Vogelaar
Margriet Bouma
Peter Verhaak
Yvonne van der Helm
Monique den Hartog
Betteke Mulder
Rikste de Hartog
Martina Frankena-van Arnebaij
Margan Essed
Anouk van Delf
Vincent Goenen
Lijan Mutsaers
Jan Cees Voskamp
Karin de Clercq
William Verheul
Dominique van Steenis
Roger Damoiseaux
Charles Helsper
Inge Groenewegen
Ernest Mutsaerts
Joop Raams
Ilse Boas-Jansen
Annemarije Kruis
Amy Los
Michiel Hagemans
Marieke Schuurmans
Jaap van der Laan
Jozien Bensing
Annemarie van Deursen
Henriette van Waaij
Bert-Jan de Boer
Tessa Magnee
Mijndert Pon
John Oosterink
Liesbeth van Vliet
Marieke Scholten
Hans, Suus en Sarah
Dinette Asterhuis
Erica van Maanen
Gijs Bos
Marlous Kortekaas
Merel van Wassenaar
Sofieke van Schaik
Janneke Noordman
Irene Drubbel
Otto Quartero
Froukje Runia
Marjam Djojodiningrat
Marieke Balkenbaas
Mijke van Gijn
Familie van Lennep
Christine Weenink
Familie Sanders
Trudy van der Weijden
Angelique Timmermans
Nienke Bleijenberg
Marieke den Hartog- van Noord
Maarten en Marie Louise Trampetter
Sandy van der Hoeven
Jan -Willem Van Uffen
Emily Fick
Richard van Kruysdijk
Peter Spreeuwenberg
Geessien Hutten-van Rooijen
Erik van Bruggen
Hans Bakker
Peter Jan de Wilde
Nielske Martens-Arnolds
Peter Zuithoff
Margit Vermeulen
Inge van Weeghel
Mark Valkenburg
Maarten Philipsen
Bertha Maat
Ardine de Wit
Tamara Plateel
Helene Voogdt
Alexander
Liesbeth Rodeke
Ab Speelman
Tennislifenleedjes
Henny en Marleen
Paul en Margot

DANKWOORD

Na meer dan tien jaar werken aan dit project, kan ik onmogelijk iedereen benoemen die ik dankbaar ben voor zijn bijdrage. Dus laat me beginnen met jullie allen te bedanken voor de vele steun en hulp die jullie me hebben gegeven. Zonder jullie was dit werk nooit afgerond. Ik heb me al die jaren omringd gevoeld door heel veel lieve, slimme, betrokken en hardwerkende mensen die op allerlei gebied (kennis, gevoel en middelen) me vooruit hielpen. Soms heb ik jullie niet de waardering terug gegeven die ik voelde maar laat deze eerste alinea de erkenning vormen voor jullie onvergetelijke bijdrage. Dank jullie wel!

Ik kan jullie onmogelijk allen bij name noemen en heb jullie daarom in groepen ingedeeld en velen van jullie op de linker pagina persoonlijk genoemd.

In de eerste plaats het promotieteam. Jozien, William en Ron maakten de start mogelijk van een onderzoek waarin ik interesse had. Niek bleef van bijna het begin tot het eind de continue factor. Ik ben me zeer bewust dat dat niet vanzelfsprekend was en het hem veel kruim heeft gekost! Peter Verhaak trok een vastgelopen traject uit het slob. Sandra gaf me met haar positieve energie de broodnodige impuls voor de eindsprint. In hun kielzog werd ik geholpen door de vele deskundigen van het Nivel en het Julius Instituut. De Peters, Zuithoff en Spreeuwenberg, hielpen me met veel geduld en uitleg bij diverse analyses. Ardine en Rabin staken veel energie in het hoofdstuk over de kosteneffectiviteit. Ook buiten het Julius Instituut en het Nivel, leerde ik vele wetenschappers kennen via het (inter)nationale platform Samen beslissen/SDM. Wat een bereidheid om kennis en ervaring te delen om zo de wetenschap vooruit te helpen! Ik ben blij dat Trudy en Anne mijn opponenten zijn bij de verdediging.

Voor het onderzoek hebben ruim tien studenten me geholpen met de dataverzameling. Zij stimuleerden de participerende huisartsen met hun maandblad “Over de rug gesproken” en bezochten met groot enthousiasme praktijken voor ondersteuning en dataextractie. De vele huisartsen, die bereid waren in het onderzoek te investeren, kregen daarvoor slechts woorden van dank en een cursus. Via hun praktijkteam betrokken zij de patiënten, die vele vragenlijsten voor dit doel invulden en bereid waren het consult op te laten nemen. Met als tegenprestatie slechts een zeer bescheiden tegoedbon. Wat bijzonder dat mensen, patiënten en zorgverleners, dit willen doen!

Tijdens de master epidemiologie leerde ik vele studenten en docenten kennen met een enorme wetenschappelijke nieuwsgierigheid en deskundigheid. Wat een vertrouwen heb ik gekregen *van,* maar ook *in* deze (jonge) professionals! De master gaf me inzicht in de deskundigheid van de Utrechtse hoogleraren waarvan er 3 in mijn beoordelingscommissie wilden plaats nemen.

Binnen de huisartsenopleiding Utrecht en de daaraan gekoppelde overlegstructuren, hebben velen mij ondersteund met raad en daad, waar ik regelmatig worstelde om het onderzoek met het onderwijs te combineren. Ook binnen het NHG en de redactie van

Huisarts en Wetenschap, heb ik een erg betrokken groep mensen om mij heen geweten met oog voor de mens en de wetenschap.

In onze huisartsenpraktijk hebben de patiënten en het personeel mijn bijdrage bij tijd en wijle moeten missen, ook als dat minder gelegen kwam. Het blijft een voorrecht om van zovelen het vertrouwen als huisarts en werkgever te krijgen.

Naast deze werk-gerelateerde kringen, kan ik natuurlijk mijn vriendinnen niet vergeten. Trouw zonder direct te oordelen maar ook correctief waar nodig. Lief-en-leed, heten 2 appgroepen, en niet voor niets! Maar ook de 'Apeldoornse Deernen', zouden gemakkelijk zo kunnen heten. Emily, mijn paranimf, heeft me niet alleen geholpen met enorm veel werk voor de dataverzameling dat ze aanvankelijk om niet en later betaald deed maar ze is vooral een ongelooflijk dierbare, kritische vriendin, steun en toeverlaat. Berthe steunde me al die jaren met wijze raad.

Tot slot de inner-circle van de familie. Van mijn ouders leerde ik in een veilige omgeving om nieuwsgierig te zijn en te volharden met zelfspot en humor. "Hoe moet ik nu de wereld weten, als pappa steeds grapjes vertelt?". Mijn broers prikkelen mijn geest met, door mij zeer gewaardeerde, kritiek vanuit het basale vertrouwen dat we in elkaar hebben. Een taak die Alexander en mijn zonen met hun partners en kind(eren) nu delen. Wat ben ik intens gelukkig dat ik jullie om me heen heb!

Lieve allemaal, dit proefschrift heeft mijn vermogen om jullie de tijd en energie te geven, die je verdient soms beperkt. Wat ben ik jullie dankbaar dat jullie levenspad het mijne kruiste en jullie mij zoveel hebben willen geven. Dank, Dank, Dank.

ABOUT THE AUTHOR

Ariëtte Sanders was born in 1962. She is the daughter of Wim and Reina Sanders, who both inspired her to be broadly interested and a critical scholar. They called her "Jettepret" because she was mostly joyful. They were both secondary school teachers. She was raised in Apeldoorn together with her two brothers. She passed her A-levels examination in 1981. After completion of her foundation course at the Agricultural University in Wageningen, she achieved her medical degree in 1990 at the Leiden University. She started her career as IVF-medical assistant at the Erasmus Academic Medical Center. In 1992 she became a mother. She started working in a nursery home and gave birth to a second son. In 1996 she was registered as a GP. A year later her third and last son was born.

At the beginning of the second millennium, Ariëtte and her husband established their GP-practice in Driebergen. Their practice has grown since then from two employees to ten nowadays. She still enjoys working as a GP and manager of the practice. She enriched her daily GP-work from 2006-2018 by being a vocational trainer for GPs at the Utrecht Medical University for two days a week. This position gave her the opportunity to start a research project in collaboration with the Netherlands Institute for Health Sciences (Nihes) in 2009 but also to obtain her master degree in Health Sciences in2014.

After 2018 she worked one year as staff member of The Dutch College of General Practitioners. In 2020, she started a position as editor for Huisarts en Wetenschap, a scientific journal for Dutch GPs. She also became a member of the Scientific Advisory Council of the Geneesmiddelenbulletin Foundation.

During her career she was especially interested in placebo-effects, communication, education, elderly, pharmacy and the emergency setting in general practice. But most of all she was interested in people.

Besides her medical career she loves to be a (grand)mother and friend.

After the completion of this thesis, she hopes to continue to spend time in doing things that give her energy and joy, like family, friends, sports, nature (including dogs), patients, practice-staff members and transferring knowledge to professionals and patients. If there still is some spare time left, she might refresh her Spanish and travel around Europe with her husband with their Kip-caravan or even further worldwide. However, above all she would like to celebrate life (with all of you) because life is so beautiful.

LIST OF PUBLICATIONS

This thesis

- **Sanders**, **A.R.J.**, van Weeghel, I., Vogelaar, M., Verheul, W., Pieters, R.H.M., de Wit, N.J., Bensing, J.M. 'Effects of improved patient participation in primary care on health-related outcomes: a systematic review.' Family Practice 2013;30:365 – 378
- **Sanders, A.R.J.**, Bensing, J.M., Essed, M.A.L.U., Magnee, T., de Wit, N. J. & Verhaak, P.F.M. 'Does training general practitioners result in more shared decision making during consultations.' Patient Educ. Couns. 2017 ;100:563-574
- **Sanders, A.R.J.,** Bensing J. M., Magnee T, Verhaak P, de Wit N.J. 'The effectiveness of shared decision-making followed by positive reinforcement on physical disability in the long-term follow-up of patients with nonspecific low back pain in primary care: a clustered randomised controlled trial. 'BMC Fam Pract 2018 June 28;19(1):102-018.
- **Sanders A.R.J.,** de Wit N.J., Zuithoff N.P.A., van Dulmen A.M.. 'The effect of shared decision-making on recovery from non-chronic aspecific low back pain in primary care; a post-hoc analysis from the patient, physician and observer perspectives.' BMC Primary Care 2022 23:22 https://doi.org/10.1186/s12875-022-01624-y

Other publications

- Speelman P., de Jongh B.M.,Th.F.W. Wolfs Th.F.W., Wittenberg J., (**Sanders A.R.J.** as a member of the Werkgroep 'Lyme-Borreliose'). CBO richtlijn voor de diagnose en behandeling van Lyme; Ned. Tijdschrift voor Geneeskd. 2004;148:659-63
- Verheul W., **Sanders A.R.J.**, Bensing J.M. 'The effects of physicians' affect-oriented communication style and raising expectations on analogue patients' anxiety, affect and expectancies.' Patient Educ. Couns. 2010;80:300-6
- van der Weijden T., **Sanders A.R.J.** 'Keuzehulpen voor de patiënt.' Huisarts en Wetenschap 2012;50:516-521
- **Sanders, A.R.J.** 'Leidt het trainen van huisartsen tot meer gedeelde besluitvorming tijdens het consult?' Huisarts en Wetenschap 2017;11: 564-567
- Voogdt H.R. , Kortleve V., Boss G., Raats I., **Sanders A.R.J.** 'Content and evaluation of a training programme on shared decision making for health care professionals on a stroke rehabilitation unit.' Free communication session; International Journal of Stroke 16(2S) 2021, Vol. 16(2S) 171–197.
- Westendorp, J., Peerdeman K.J., van Vliet, L.M., Meeuwis, S.H., olde Hartman, T.C., **Sanders, A.R.J.**, Godfried, M.B., Thiel, B., Evers, A.W.M. Virtual Reality communication training tool for optimizing placebo- and minimizing nocebo-effects. EACH conference: International Conference on Communication in Healthcare, Glasgow (Scotland), September 5-9 2022.

- Baghus A., Giroldi E., van Geel J., Leferink A., van de Pol M., **Sanders A.**, Dielissen P., Bisschop I., Pieterse A.H., Muris J., Timmerman A., van der Weijden T. Shared decision-making performance of general practice residents: an observational study combining observer, resident and patient perspectives. *In submission*
- Multiple news items and an interview for Huisarts and Wetenschap about divers topics with a focus on communication, elderly and emergency care. (2019-now)