Sanders, A.R.J., Bensing, J.M., Essed, M.A.L.U., Magnée, T., Wit, N.J. de, Verhaak, P.F.M. Does training general practitioners result in more shared decision making during consultations? Patient Education and Counseling: 2017, 100(3), 563-574

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

A.R.J. SANDERS[A, ,], J.M. BENSING[B, C], M.A.L.U. ESSED[A], T. MAGNÉE[B], N.J. DE WIT[A], P.F.M VERHAAK[B, D]

[a] Julius Centre for Health Sciences and Primary Care, University Medical Centre Utrecht, Mailbox 85500, 3508 GA Utrecht, The Netherlands
[b] NIVEL (Netherlands Institute for Health Services Research), Mailbox 1568, 3500 BN Utrecht, The Netherlands
[c] Faculty of Social and Behavioural Science, Utrecht University, Postbus 80140, 3508 TC Utrecht, The Netherlands
[d] Groningen University, University Medical Centre Groningen, Dept. of General Practice Groningen A. Deusinglaan 1, 9712 CP Groningen, The Netherlands

HIGHLIGHTS

- Two training sessions of 2,5 h increases GP behaviour towards more shared decision making.
- GPs express a tension between guideline recommendations and the implementation of patient preferences.
- Taking the patient's perspective into account remains troublesome for GPs.

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.

Methods: 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.

1. INTRODUCTION

In medical decisions, little attention is devoted to the patient perspective, and patients' expectations often remain unnoticed [1], [2] and [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] and [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] and [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] and [9].
Despite impressive scientific efforts, effective methods of implementing this approach remain unclear [9], [10] and [11]. Further, current knowledge on effective methods of directing professional behaviour towards more patient-centred care and SDM is scarce and inconsistent [9] and [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] and [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 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 low back pain than untrained GPs.

2. METHODS

2.1. Design

This study was embedded in a clustered randomised trial that evaluated the effectiveness of SDM among patients with 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 low back pain and to videotape their consultations with those patients. Of the consultations completed with 226 recruited patients, 175 consultations were videotaped and used for this secondary analysis (Fig. 1).

[FIGURE 1]

2.2. Participants

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

2.3. 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 low back pain according to professional guidelines. The GPs were encouraged to discuss the favourable prognosis of 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] and [13]. In the training sessions, group discussion, theory, role-playing and reflections on personal behaviour were alternated [13]. A decision aid for non-chronic, non-specific 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).

[BOX 1]

In addition to receiving training sessions, the GPs in the intervention group received personalised feedback on each videotaped 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 videotape, 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.

2.4. 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.

2.5. 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, they completed one questionnaire before the consultation (baseline measurement). The GPs videotaped the consultation about back pain (Fig. 1).

2.6. 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.

2.7. Measurement instruments

2.7.1. 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/48 × 100. The scale exhibits adequate properties (Cronbach's alpha = 0.728).

2.7.2. 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))

2.7.3. 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.”

2.7.4. 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 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 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.)

2.7.5. Duration of the consultation

The total duration of the consultation was measured from the videotape. The taping started when the patient entered the room and ended when the GP stopped the videotape because the consultation about 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.

2.7.6. Coding reliability

Two blind observers (AL and IvdE) scored the videotapes using Observer (Noldus, 7th edition), a program designed to aid in the observation of videotapes [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).

2.8. 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, 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 the video rank 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.

2.9. 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 videotaped.

3. RESULTS

3.1. GPs and consultations

Twenty-three trained GPs videotaped 86 consultations, and 19 untrained GPs videotaped 89 consultations (Fig. 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 videotaped 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, videotaped 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).

[TABLE 1]

3.2. Intervention effects

3.2.1. 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 (Table 2). Nevertheless, the level of positive reinforcement was not assessed reliably and should be interpreted with caution.

[TABLE 2]

3.2.2. 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).

[TABLE 3]

3.2.3. 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.

3.2.4. 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 (Table 4).

[TABLE 4]
The duration of the consultation had an important and significant positive effect ($p = 0.000$) on the summed score of the process elements.

4. DISCUSSION AND CONCLUSION

4.1. 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] and [19]. The last item was primarily observed in the intake phase. The Dutch guidelines for low back pain recommend 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] and [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 videos 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] and [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? To what extent does SDM imply an equal relationship in decision making? 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 process score, and after

duration is controlled for, the association 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.

4.1.1. Strengths of the study

We developed a training and supportive tool that conforms to recent internationally accepted guidelines [9] and [11]. The duration of the training was sufficient considering recently gained knowledge [11] and [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] and [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 taken into account [10].

4.1.2. 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 videotaped 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 videos, the observed behavioural differences were significant.
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 videotaped 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].

4.2. 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 clinical guidelines.

4.3. 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 van Dedem-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.

FUNDING

The trial was funded by the National Institute for Health Research and Development (Grant Numbers 42011009 and 510000005).

ROLE OF FUNDING

The funding source (National Institute for Health Research and Development) was not involved in the research process.

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 videotapes. We thank Amy Los (AL) and Ilse van de Ende (IvdE) for observing the videotapes, Jan Maessen for transporting the observation data to statistical programs and Peter Spreeuwenberg for performing the statistical analyses.

Front-side of plasticized A3 format decision aid card

APPENDIX A. DECISION AID (SUB)ACUTE NONSPECIFIC LOW BACK PAIN
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.) od 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 can not tell you which one will be best for you.

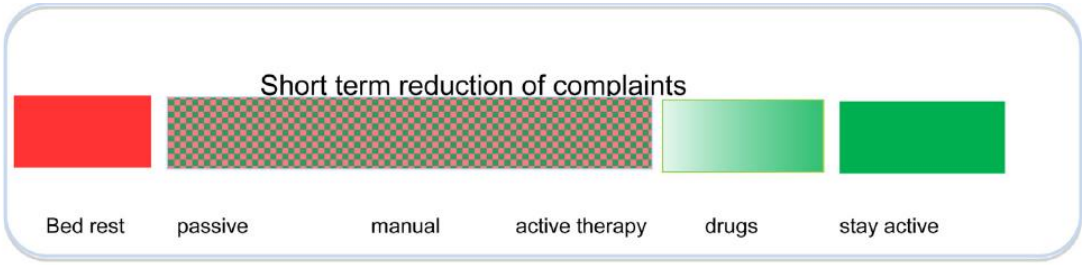
You can choose to wait and see whether pain will go away by itself. What we do know is that your recovery will be sooner if you stay active as much as possible and that staying in bed has negative results.

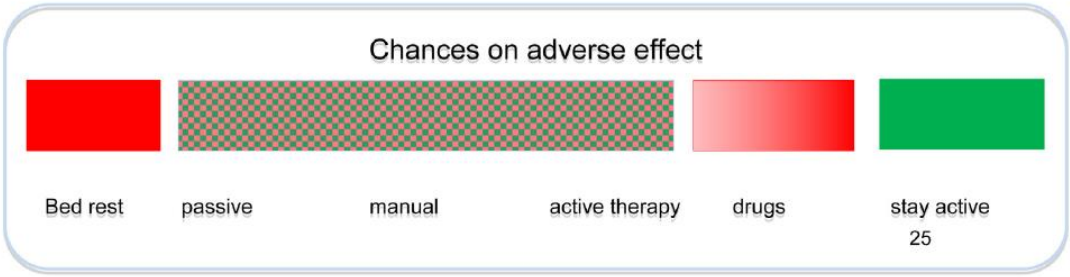
Back-side of plasticized A3 format decision aid card

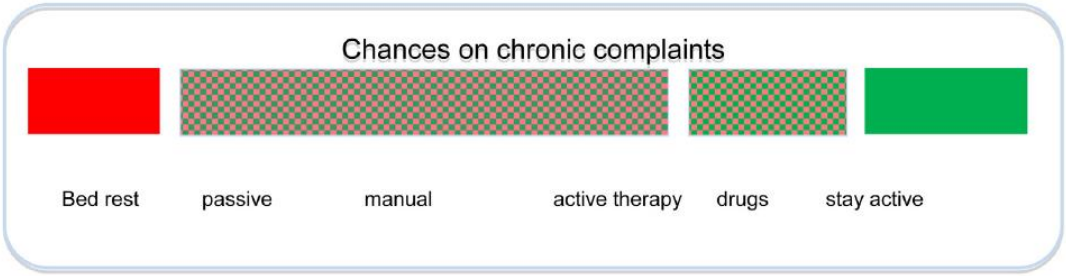
In the tables mentioned below you can see how effective different therapies are for the majority of people on short term. What are chances on adverse effects and on chronic complaints. The tables can help you to decide together with your physician how you want to cope with your back pain. You can also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

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 LBP*, 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 LBP*. 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 sub-acute 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 randomized 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 mobilization, 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 *sub-acute and chronic non-specific 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 LBP. Three studies had low risk of bias. Included studies evaluated a range of chiropractic procedures in a variety of sub-populations of people with LBP. No trials were located of combined chiropractic interventions compared to no treatment. For *acute and sub-acute LBP*, 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 LBP* and for studies that had a mixed population of LBP.

Exercise therapy for treatment of non-specific low back pain
February 2011
Sixty-one randomized controlled trials (6390 participants) met inclusion criteria: *acute (11), sub-acute (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 sub-acute 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 LBP with radiation versus LBP 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 LBP*. There is moderate evidence suggesting that back schools for *chronic LBP* 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 non-specific 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 LBP* on short-term pain relief. The pooled RR for non-benzodiazepines versus placebo after two to four days was 0.80 [95% CI; 0.71–0.89] for pain relief and 0.49 [95% CI; 0.25–0.95] for global efficacy. Adverse events, however, with a relative risk of 1.50 [95% CI; 1.14–1.98] were significantly more prevalent in patients receiving muscle relaxants and especially the central nervous system adverse effects (RR 2.04;

95% CI; 1.23–3.37). The various muscle relaxants were found to be similar in performance.

Antidepressants for non-specific 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); standardized 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 *sub-acute LBP*, 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 LBP* 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 0.48	0.89‡ 0.16‡	0.88 0.69	0.75 0.59
equipoise	0.44	0.35‡	0.21 0.36	0.29‡ 0.14‡	0.88 0.84	1.24 1.14
information format	0.86	0.86‡	0.90 0.99	0.89‡ 1.00‡	0.34 0.34	0.48 0.48
lists options	0.88	0.75‖	0.99 0.86	1.000† 0.88‖	2.53 1.16	1.27 0.85
explanation of pros and cons of options	0.70	0.55	0.74 0.87	0.78‡ 0.82‡	1.16 1.72	0.88 0.58
exploration of the patient's expectations	0.70	0.78†	0.60 0.27	0.59† 0.41†	1.72 1.59	0.58 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 0.84	n.a. ‡ n.a. ‡	0.97 1.00	0.54 0.51
offering opportunities to ask questions‡	0.79	0.71‡	0.92 0.68	0.88† 0.65†	1.16 1.13	0.72 0.76
elicitation of patient's preferred level of involvement*	0.83	0.75†	0.73 0.91	0.49† 1.00†	1.31 1.53*	1.06 1.05*
indication of a decision-making stage	0.13	0.13‡	0.24 0.98	0.22‡ 1.00‡	0.41 0.31	0.50 0.47
indication of the need to review the decision	0.79	0.93‡	0.72 1.00	0.70‡ 1.00‡	2.88 2.28	1.83 1.84
total score*	0.92	0.61††	0.92 0.93	0.67†† 0.75††	14.42 13.42	4.98 4.86
positive reinforcement of the chosen therapy	0.39	0.18	0.80 0.67	0.80‡ 0.64‡	0.63 0.66	0.79 0.83
level of autonomy (SDM = 2)	0.67	0.59†	0.84 0.81	0.80† 0.77†	1.19 1.16	0.82 0.77
discussion of the benign	0.89	1.00‖	0.853	0.80‖	2.09	1.20

	INTER-OBSERVER		INTRA_OBSERVER			
COMMUNICATIVE ELEMENT	**Spearman's correlation**	**Cohen's kappa**	**Spearman's correlation¶**	**Cohen's kappa ¶**	**Mean per observer¶**	**SD per observer¶**
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 scores1–13 = 1 and scores 14–27 = 2.

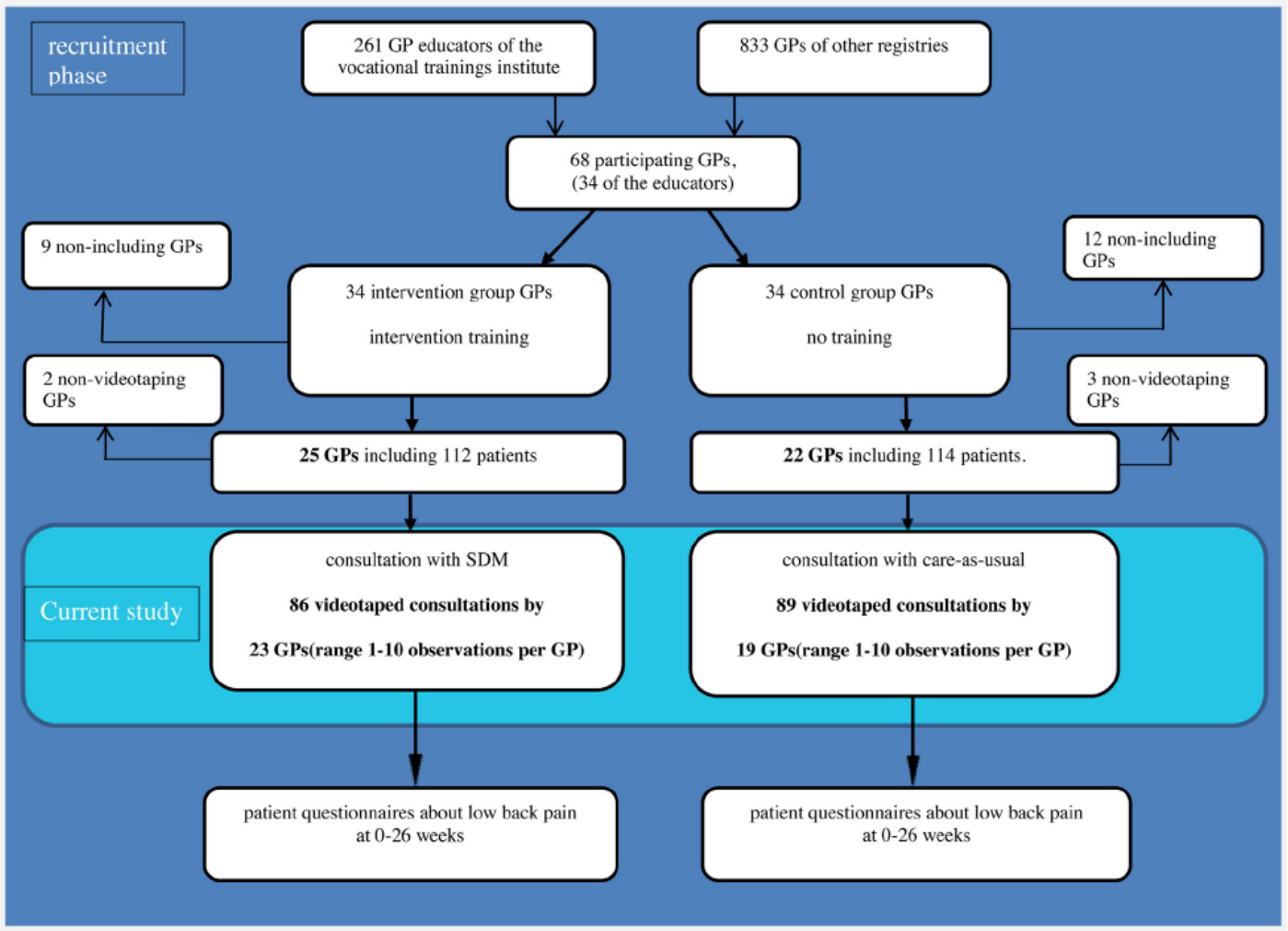
REFERENCES

[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.

FIGURES AND TABLES

Figure 1: Flowchart. GP = general practitioner, SDM = shared decision-making, PR = positive reinforcement of the chosen therapy, *these GPs did not include any patient.



Box 1.

Training steps, aim and format both training sessions.

	Step	Aim	Format
	FIRST AND SECOND SESSION		
1	Introduction	Create a safe environment	Introduction round
2	Inventory of attitudes	1. Focus on placebo-knowledge on patient recovery 2. Focus on learned skills of SDM and PR	Plenary short discussion
3	Reflection on daily practice	1. Create a shift from 'unaware unskilled' to 'aware unskilled' in SDM and PR 2. Create a situation of constructive friction in experienced (lack of) skills in SDM and PR	1. Group discussions of 2–3 persons 2. Group discussions of 2–3 persons 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 alternated with plenary discussion based on questions about 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 low back pain consultation 2. Role play with decision aid (obligatory) and desktop 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'
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.

[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.

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 number 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	45.48	14.0	44.53	14.1	0.655
male	47%(40)		52%(45)[a]		0.546
Dutch origin	96%(75)[c]		99%(79)[b]		0.926
educational level	[c]		[a]		0.969
low	18%(15)		17%(15)		
middle	49%(41)		49%(43)		
high	31%(26)		33%(29)		
VAS pain (0–100)	48.2[e]	15.6	46.9[d]	16.7	0.620

VAS = visual analogue scale.

* = $p < 0.05$.

A = 2 missing cases.
B = 3 missing cases.
C = 4 missing cases.
D = 14 missing cases.
E = 14 missing cases.

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 video rank 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.

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

		trained GP group (n = 86)		untrained GP group (n = 89)		
SECONDARY OUTCOMES		**mean**	**SE**	**mean**	**SE**	**difference**
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	−0.72

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

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.