

STUDY PROTOCOL

Open Access



# Evaluation of a risk-stratification strategy to improve primary care for low back pain: the MATCH cluster randomized trial protocol

Dan Cherkin<sup>1\*</sup>, Benjamin Balderson<sup>1</sup>, Georgie Brewer<sup>2</sup>, Andrea Cook<sup>1</sup>, Katherine Talbert Estlin<sup>3</sup>, Sarah C. Evers<sup>1</sup>, Nadine E. Foster<sup>4,5</sup>, Jonathan C. Hill<sup>6</sup>, Rene Hawkes<sup>1</sup>, Clarissa Hsu<sup>7</sup>, Mark Jensen<sup>2</sup>, Anne-Marie LaPorte<sup>2</sup>, Martin D. Levine<sup>8</sup>, Diane Piekara<sup>3</sup>, Pam Rock<sup>3</sup>, Karen Sherman<sup>1</sup>, Gail Sowden<sup>5,9</sup>, Rob Wellman<sup>1</sup> and John Yeoman<sup>2</sup>

## Abstract

**Background:** Despite numerous options for treating back pain and the increasing healthcare resources devoted to this problem, the prevalence and impact of back pain-related disability has not improved. It is now recognized that psychosocial factors, as well as physical factors, are important predictors of poor outcomes for back pain. A promising new approach that matches treatments to the physical and psychosocial obstacles to recovery, the STaRT Back risk stratification approach, improved patients' physical function while reducing costs of care in the United Kingdom (UK). This trial evaluates implementation of this strategy in a United States (US) healthcare setting.

**Methods:** Six large primary care clinics in an integrated healthcare system in Washington State were block-randomized, three to receive an intensive quality improvement intervention for back pain and three to serve as controls for secular trends. The intervention included 6 one-hour training sessions for physicians, 5 days of training for physical therapists, individualized and group coaching of clinicians, and integration of the STaRT Back tool into the electronic health record. This prognostic tool uses 9 questions to categorize patients at low, medium or high risk of persistent disabling pain with recommendations about evidence-based treatment options appropriate for each subgroup. Patients at least 18 years of age, receiving primary care for non-specific low back pain, were invited to provide data 1–3 weeks after their primary care visit and follow-up data 2 months and 6 months (primary endpoint) later. The primary outcomes are back-related physical function and pain severity. Using an intention to treat approach, intervention effects on patient outcomes will be estimated by comparing mean changes at the 2 and 6 month follow-up between the pre- and post-implementation periods. The inclusion of control clinics permits adjustment for secular trends. Differences in change scores by intervention group and time period will be estimated using linear mixed models with random effects. Secondary outcomes include healthcare utilization and adherence to clinical guidelines.

**Discussion:** This trial will provide the first randomized trial evidence of the clinical effectiveness of implementing risk stratification with matched treatment options for low back pain in a United States health care delivery system.

**Trial registration:** NCT02286141. Registered November 5, 2014.

**Keywords:** Back pain, Risk-stratification, STaRT back, Quality improvement, Guidelines, Primary care, Physical therapy, Patient outcomes, Utilization, Systems

(Continued on next page)

\* Correspondence: cherkin.d@ghc.org

<sup>1</sup>Group Health Research Institute, Seattle, USA

Full list of author information is available at the end of the article



(Continued from previous page)

**Abbreviations:** CAM, Complementary and alternative medicine; CATI, Computer-assisted telephone interview; CBT, Cognitive behavioral therapy; CME, Continuing medical education; DVD, Digital video disc; EHR, Electronic health record; GAD, Generalized anxiety disorder; GH, Group Health; GLMM, Generalized linear mixed models; LMM, Linear mixed models; MATCH, Matching appropriate treatments to consumers healthcare needs; NIH, National Institutes of Health; PC, Primary care; PCORI, Patient Centered Outcomes Research Institute; PCP, Primary care provider; PGIC, Patient global impression of change; PHQ, Patient health questionnaire; PSEQ, Pain self efficacy questionnaire; PT, Physical therapy/physical therapist; QI, Quality improvement; RMDQ, Roland-Morris disability questionnaire; RS, Risk stratification; STarT, Subgroups for targeted treatment” risk stratification; TSK, Tampa Scale of Kinesiophobia; UC, Usual care; UK, United Kingdom; US, United States; WPAI, Work productivity and activity impairment

## Background

Finding effective and affordable treatments for chronic pain is a major challenge for clinicians, researchers, payers, and patients in industrialized countries [1]. Back pain is the most prevalent and costly type of pain. More than 50 % of US adults are bothered by back pain each year and up to 80 % are afflicted by it at some time in their lives [2]. About \$26 billion is spent annually in the US on personal medical care for back pain [3], and in 2002, the costs of lost worker productivity due to back pain were almost \$20 billion [4]. Despite the growing number of options available for evaluating and treating back pain, and the increasing costs of medical care devoted to this problem, the health and functional status of people suffering from back pain in the US has deteriorated [5]. The current epidemic of opioid addiction and deaths among those treated for chronic pain illustrates the importance and urgency of finding safer and more effective approaches to relieving chronic pain [6, 7].

There has been an increasing appreciation among experts of the limitations of the traditional view of back pain as a largely biomedical problem [8]. More comprehensive biopsychosocial models acknowledge that chronic pain usually has an underlying biological basis, but recognize that psychosocial factors—such as pain beliefs/cognitions, coping behaviors and the social environment—also significantly influence the experience of pain and its effects on physical, psychological, and social functioning [9, 10]. Research has consistently found strong relationships between maladaptive beliefs/cognitions and negative pain-related outcomes [9, 11]. This research provides a clear rationale for incorporating cognitive behavioral principles into the management of the distressed and disabled (high risk) low back pain patients for the purposes of secondary prevention of pain-related disability.

### The “Subgroups for Targeted Treatment” (STarT) risk stratification approach

A promising approach to targeting treatments to subgroups of patients, the STarT Back approach [12], has

been found effective in improving patient physical functioning and satisfaction with care while reducing costs of healthcare in both physical therapy settings [13] and family physician practices in the UK [14, 15] ([www.keele.ac.uk/sbst/](http://www.keele.ac.uk/sbst/)). The STarT Back approach first uses a 9-item “STarT Back tool” patient questionnaire to categorize patients by their risk of persistent disabling back pain and then identifies the treatments deemed most appropriate for patients in each subgroup [12, 16, 17]. Patients found to have 4 or more “psychosocial” risk factors (pain bothersomeness, fear, worry, catastrophizing, depression) are considered high risk and those with relatively few (0–3) physical or psychosocial risk factors are considered low risk. The remaining patients, who have significant pain and/or activity limitations but few psychosocial risk factors, are considered moderate risk. This tool has been validated for use with primary care adults with non-specific low back pain [12].

In the UK trial, low risk patients were offered analgesic medications, reassurance, and advice and education on self-care; medium risk patients were offered physical therapy (PT) treatment, and high risk patients were offered PT from therapists who had been specially trained to elicit and address key psychosocial obstacles to recovery using a psychologically informed physical therapy approach [18, 19]. The success of this risk stratification approach in the UK has generated high levels of interest, providing new hope that meaningful improvements in primary care for back pain are within reach [20–25].

### Will the STarT back risk stratification strategy work in the US?

The MATCH (Matching Appropriate Treatments to Consumers Healthcare needs) trial evaluates implementation of a version of the STarT Back risk stratification approach adapted for use in an integrated health care delivery system in the U.S. serving over 600,000 members—Group Health (GH) in Washington State. We obtained funding from the National Institutes of Health (NIH) to evaluate the implementation process and from the Patient-Centered Outcomes Research Institute (PCORI) to evaluate the

effectiveness of this approach. This is the first major evaluation of the risk stratification approach based on the STarT Back tool for back pain in the US.

Prior to the initiation of the MATCH trial, GH had communicated elements of the STarT Back approach to all clinicians via a one-time video-webcast continuing medical education (CME) presentation. This optional 45-min presentation briefly described the STarT Back tool and how to access it in the electronic health record (EHR) as well as GH's new clinical guidelines for low back pain. This initial version of the STarT Back tool was similar but less well-developed than the one used in the intervention. Neither primary care clinicians nor physical therapists received any training or support to implement the STarT Back approach. Not surprisingly, very few clinicians began using the STarT Back tool following this minimal and passive CME approach and it is not clear if those who did use it were recommending treatment options appropriate for the patients' risk level.

Funding for the MATCH study provided GH an opportunity to conduct a trial to evaluate full implementation of the STarT Back approach in the form of a mandatory quality improvement (QI) initiative. This manuscript describes the development and implementation of this QI strategy and the protocol for evaluating its effectiveness. The goal of this initiative was to give primary care providers (PCPs) and physical therapists (PTs) the knowledge, tools, and confidence they needed to provide their patients with a broader understanding of their back problem, reassurance that their condition would likely improve, and provide treatment options most likely to be helpful. We hypothesized that this QI intervention would improve patient outcomes by promoting the increased use of effective treatment options that address patients' needs.

## Methods/design

### Development and implementation of the QI strategy

Development and implementation of the risk stratification approach included assembly of the project team, elicitation of perspectives of key stakeholders, selection of the treatment options to be recommended for each risk subgroup, programming the EHR to support the intervention, training the primary care (PC) teams and PTs, and implementation in three PC clinics where both the PC teams and PTs are co-located.

### Assembled the project team

We developed the QI strategy between March 2013 and April 2014 with the support of key GH primary care and PT leaders, several of whom actively served on our project team (ML, DP, PR). To fully benefit from the expertise of the UK group that had developed, tested and implemented the STarT Back approach, we included

three key UK members (NF, JH, GS) on our team. We also invited four local individuals with extensive personal experience with chronic pain to serve on our project team. Three are actively involved in helping others with chronic pain through classes and support groups. These "Patient Partners" provided valuable perspectives on the implementation strategy, including identification of ways clinicians could more effectively communicate with patients and outcome measures to use in our evaluation.

### Elicited perspectives of key stakeholders

To ensure we understood the perspectives of GH patients and clinicians on key issues related to the design and implementation of the STarT Back approach we convened 3 focus groups with a total of 28 patients and 3 focus groups with a total of 15 clinicians (10 PCPs and 5 PTs). To maximize the potential for efficiently implementing our intervention strategy in other U.S. healthcare settings, we recruited a diverse group of national advisors representing patients, large employers, major governmental and independent healthcare systems, primary care practice networks, major government payers and insurers, and institutions that train complementary and alternative medical care (CAM) providers. Most advisors were in positions to influence policies within their organizations.

### Identified treatment options for each risk subgroup

We used the STarT Back tool without modification, but selected treatment options matched to each risk stratum that were deemed appropriate for a US, and specifically GH, setting. We used GH's new guidelines for assessing and treating patients with back pain (Additional file 1) to identify the evidence-based treatment options that could be matched to each patient subgroup. These guidelines adapted leading back pain guidelines in the US [26] and UK [27], for use at GH. For example, the US guidelines recommend a multi-disciplinary intensive rehabilitation program as an option for persistent back pain [26], but this was not readily available at GH. However, GH patients do have varying degrees of access to the remaining six treatment options recommended by the US guidelines [26]. CBT (available from GH's behavioral health service), exercise programs (available from GH PTs), spinal manipulation therapy (chiropractic)/acupuncture/massage (available to most GH members through a contracted network of external providers), and yoga (available in the community, but not covered by insurance). Given the similar clinical effectiveness observed in previous systematic reviews [28, 29], we encouraged clinicians to try to interest patients in active treatments (exercise, yoga) before recommending passive options (chiropractic, massage, acupuncture).

Clinicians in the intervention clinics were trained to routinely administer the STarT Back tool to determine each patient's risk of persistent disabling pain and to

tailor their treatment recommendations to the patient's risk level [12]. Patients could be offered treatment options appropriate for patients at lower risk levels. The treatments recommended for each risk level were:

- **Low Risk (~40 % of patients):** The PCP can manage most patients in this category through minimal interventions without referral to an additional provider [30]. For these patients the provider should conduct a brief assessment to rule out potentially serious causes of back pain (i.e., "red flags"), elicit and listen to patients' concerns, provide reassurance about the positive prognosis and self-care recommendations to relieve pain (i.e., appropriate physical activity, use of pain medications and avoiding bed rest). PCPs were trained and encouraged to recommend that their patients access online DVDs that reinforced information about acute or chronic back pain and the importance of self-care.
- **Moderate Risk (~40 % of patients):** Patients in this group should be offered additional guideline-recommended treatment options, particularly those involving exercise and activating treatments that could reduce fear of movement (i.e., PT and yoga). Patients not interested in activating treatments should be offered the more passive options (acupuncture, spinal manipulation, or massage therapy) in the hope these treatments will help decrease their pain and prepare them for transition to more active approaches, including the Living Well with Chronic Conditions self-management classes available at GH.
- **High Risk (~20 % of patients):** High risk patients are on average more complex (distressed and disabled) than patients in the low or medium risk subgroup, therefore, in addition to the treatment options for moderate risk patients, the best available treatment options for patients in this subgroup at GH was psychologically-informed practice delivered by PTs specially-trained for this initiative or referral to a psychologist for CBT. "Psychologically informed practice offers a systematic approach to the integration of physical and psychological approaches to treatment for the management of people with low back pain..." [19]. Here, 'psychological' refers to the beliefs/expectations, emotional and behavioral responses associated with low back. Unfortunately, access CBT from a psychologist was very limited. PCPs were also encouraged to proactively follow-up with high risk patients within 2 weeks.

#### **Incorporated STarT back tool, recommended treatments, and decision aids into the EHR**

GH's Clinical Informatics specialists had previously incorporated the STarT Back tool into its EHR (EpicCare),

but this produced only the total number of risk items endorsed by patients without determination of risk level or recommendation of appropriate treatments. To resemble the way the STarT Back tool was used in the UK, we enhanced the tool so that after a patient's responses had been entered, the EHR automatically calculated the patient's risk stratum and displayed the recommended treatment options for that stratum on a screen visible to both clinicians and patients. This provided an opportunity for clinicians to discuss treatment options with their patients. Providers had the option of including this information in the patient's after visit summary. To promote routine use of the STarT Back Tool, we let clinics and clinicians decide how to most consistently and efficiently collect and enter the STarT Back Tool data into the EHR. Although the process varied, most clinicians had their medical assistants collect the data either on a paper copy and then enter it into the EHR, or enter it directly into the EHR.

Shortcuts were incorporated into the EHR to help clinicians efficiently access these tools and the GH back pain guidelines and link patients with existing GH educational resources (DVDs about acute and chronic back pain, and when surgery might be indicated) including GH's self-management groups for persons with chronic conditions (Living Well With Chronic Conditions).

#### **Developed training modules for primary care teams**

We developed a package of 6 one-hour training sessions to be presented in each of the 3 primary care intervention clinics (Table 1). Each topic was presented on several occasions to ensure that all PCPs participated. Sessions were presented roughly monthly over a 6-month period (May – October, 2014). Training focused on the STarT Back tool and matched treatment options (emphasizing the importance of the biopsychosocial model), techniques and strategies for talking about chronic pain with patients, the special training GH PTs had received in incorporating simple CBT techniques into their PT practice to use with high risk patients, and understanding the role of evidence-based CAM therapies. Physical therapists and members of the nursing staff were invited to attend several of the sessions.

PCPs and staff received coaching on how to locate and correctly use the STarT Back and other related tools in the EHR. This coaching was usually done with one or individuals. Usually team members with the same role were paired. Each session lasted approximately 30 min and took place at work stations. Most PCPs participated in at least one coaching session.

Finally, to reduce knowledge barriers to recommending evidenced-based treatment options, we compiled a list of the names and contact information of meritorious CAM providers and made them available to clinic staff.



**Table 1** Overview of Training Sessions for Primary Care Providers (PCPs) in Intervention Clinics

Overview	
Session 1	<p><b>Introduction</b></p> <p>Physician leader of the back pain care quality improvement project introduced the researchers, the rationale and importance of reorganizing care for back pain and overview of the project aims and activities. Focused on getting clinicians to support the project and excited about the opportunity to improve back pain care in their clinics. Introduced the STarT Back tool and risk stratification strategy.</p>
Session 2	<p><b>Using the STarT Back Tool</b></p> <p>Focused on getting PCPs and staff comfortable with administering the STarT Back approach, including use in the EHR (EPIC), scoring of tool, and understanding matched treatment recommendations for each risk level. Discussions of how to find and use the GH Back Pain Guidelines, attach the STarT Back tool in secure messages to patients, enter patient responses into the EHR, view results, use tools within the EHR to enhance visits for back pain such as patient instructions and ordering patient centered back pain (acute and chronic) DVDs.</p>
Session 3	<p><b>Improving Diagnosis and Ruling out Red Flags</b></p> <p>Check-in on clinic's use of the STarT Back tool. Used patient case examples to provide a "refresher" on how to conduct differential diagnoses of common back pain problems (focused on lower back, L2-L4, L5, S1). Reviews included how to conduct an efficient exam, common errors in examinations, appropriate use and interpretation of diagnostic imaging, red flags for serious conditions, how to communicate to patients during the examination.</p>
Session 4	<p><b>Talking with Patients about Chronic Pain</b></p> <p>Training focused on ways to communicate more effectively with patients about chronic pain including: 1) preferred language in discussing pain, 2) ways to better communicate anatomical links to pain, 3) explaining what chronic pain is (pain centralization, gate theory, recurrence of pain – having continued pain with no injury), 4) talking about red flags and when to return to primary care, 5) focusing on improving function rather than reducing pain, and 6) how to discuss outcomes from the STarT Back tool and shared decision making around treatment options.</p>
Session 5	<p><b>Improving Partnership between PCPs and PTs</b></p> <p>This session brought together PCPs and PTs for an interactive discussion on how to improve team based care for patients. Topics included: how PT was providing improved care based on training, shared responsibility and roles of providers, how to integrate the STarT Back tool across the departments, and how to collaboratively work together to help patients not showing improvement.</p>
Session 6	<p><b>Complementary and Alternative Medicine Treatments for Chronic Back Pain</b></p> <p>Focused on building an understanding of the role that CAM providers (acupuncturists, chiropractors, massage therapists) and yoga classes can play for back pain patients. Emphasized practical information: brief description of each CAM modality, scientific evidence for their effectiveness, contraindications, dosing. CAM therapies referrals were linked to STarT Back tool risk category, emphasized the use of active over passive therapies, and how to conduct referrals within the healthcare system.</p>

We did this by asking clinicians in each intervention clinic to identify local acupuncturists, chiropractors, and massage therapists covered by GH insurance to whom they had become comfortable referring their patients

with back pain. We also asked clinicians to recommend local yoga classes (not covered by insurance).

### Training for physical therapists

The PTs in the intervention clinics received five days of training from a UK instructor (GS) who had trained the PTs in the original STarT Back Trial and IMPaCT Back study [18]. This training built upon the established professional expertise of Physical Therapists by paying specific and systematic attention to the psychosocial factors that are associated with a poor treatment outcome in people low back pain. To maximise PT effectiveness, the training included the key psychosocial variables that can contribute to the development and maintenance of pain related disability, a focus on raising clinician awareness about why people in pain think, feel and behave the way they do and to helping them apply pain relevant psychological theories and practice to their assessment and management of low back pain patients (Table 2). Following the training, the PTs were expected to begin using the STarT approach in their practices.

A psychologist on the team with expertise in CBT and chronic pain (BB) helped facilitate the training and interacted regularly with the PTs after the training. Two months after the formal training he conducted consultation meetings at each of the intervention clinics. These meetings, which reviewed how to use the tool and how to talk with patients about pain, provided therapists support for continuing to use the new approaches they had learned. In later sessions, PTs discussed experiences with

**Table 2** Physical Therapy Training

	Planned Topics
Day 1	<ul style="list-style-type: none"> <li>• Description of STarT Back Trial, IMPACT study and other related research</li> <li>• Description of stratified care</li> <li>• Myths and facts about patients that have chronic pain</li> <li>• Research on pain models, the complexity of pain experiences, with special emphasis on moving away from seeing pain as an indication of tissue damage.</li> <li>• Research on neurophysiology of pain</li> </ul>
Day 2	<ul style="list-style-type: none"> <li>• Research on neurophysiology of pain (continued)</li> <li>• Review of key factors that contribute to development and maintenance of pain related disability</li> <li>• Communication skills for working with patients with disabling chronic pain</li> </ul>
Day 3	<ul style="list-style-type: none"> <li>• Assessment of high risk patients</li> <li>• Managing/treating high risk patients</li> <li>• Integrating the psychosocial approach into manual therapy</li> </ul>
Day 4	<ul style="list-style-type: none"> <li>• Explaining pain</li> <li>• Managing expectations</li> <li>• Facilitating behavioral change/goal setting</li> </ul>
Day 5	<ul style="list-style-type: none"> <li>• Managing disability</li> <li>• Vocational rehabilitation</li> <li>• Clinical decision making and treatment planning</li> <li>• Monitoring and modifying treatment plans</li> </ul>

specific patients to highlight effective strategies for incorporating what they had learned into their practice. BB facilitated at least 4 of these sessions per clinic over a 6-month period. In addition, one-on-one training was targeted at PTs with specific needs.

**Implemented risk-stratification strategy**

Prior to implementation, we met with the primary care and PT teams at the intervention clinics to orient them to the study and explain how it fit into GH’s broader clinical improvement initiative for back pain. Over the next six months (May 2014 – October 2014), we worked with primary care and PT leaders to implement training.

**Protocol for evaluating the effects of implementing the STarT back strategy**

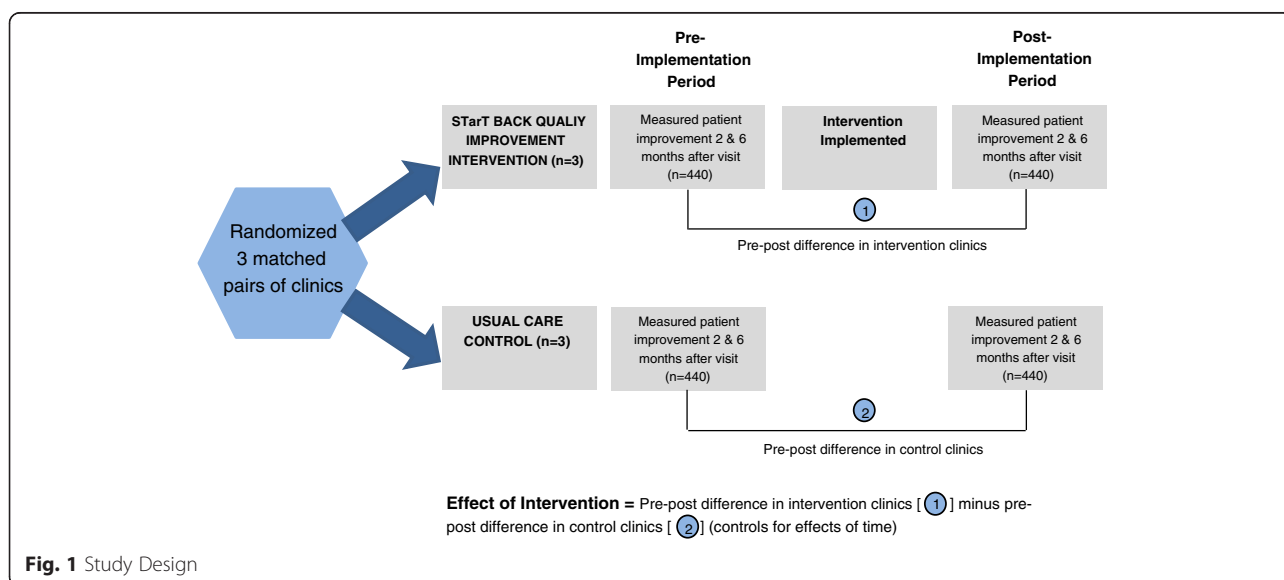
We evaluated the intervention’s effect on patients and healthcare utilization, including adherence to best practice guidelines. We also collected data that allowed us to evaluate the implementation process itself (reported separately).

**Effect on patient outcomes**

The primary focus of our evaluation is on patient outcomes. The basic study approach was to determine if patient outcomes measured 2 and 6 months after a visit for back pain were better after implementation of the quality improvement strategy than before (Fig. 1). We first randomly assigned one clinic from each of 3 pairs of large primary care clinics (matched on geographic region and sociodemographic characteristics of the patient populations) to receive the intervention. Implementation at the clinic level also minimizes differences in the

patient populations in the intervention and control groups. Implementation at the clinic level also minimizes risks of contamination and training costs and reflects the most practical approach for real-world healthcare systems. The inclusion of control clinics allows us to adjust the analyses for concurrent changes that may be occurring in GH primary care clinics during the intervention period. Randomization within matched pairs of clinics avoids biased allocation.

To estimate the effects of the intervention on patient outcomes, we compared the extent of improvement 2 and 6 months after a back pain visit during the pre-implementation period with that for the post-implementation period. For both periods, we identified patients within one week of a primary care visit for non-specific back pain, mailed letters inviting participation, collected baseline data within 3 weeks of the index visit, and then assessed outcomes 2 and 6 months later. This allowed estimation of the degree of improvement in dysfunction and pain intensity outcomes occurring prior to implementation of the risk stratification strategy in both the intervention and control clinics. Beginning about 10 weeks after the end of the 6-month implementation phase in the intervention clinics, we repeated the pre-implementation process of estimating the degree of improvement in the intervention and control clinics. Patients providing data during the pre-intervention period were excluded from the post-intervention period. We will estimate the effect of the intervention in the intervention clinics by comparing the pre-post differences in the intervention clinics with those in the control clinics and thus accounting for any secular trends over the course of the study.



**Fig. 1** Study Design

### Patient recruitment

We used the EHR to identify all adult patients (18+ years old) who received “primary” (i.e., first-listed) diagnoses consistent with non-specific back pain (e.g., lumbago, back pain not otherwise specified). To maintain broad applicability of the study population we only excluded patients with specific causes of their pain (such as pregnancy, disc herniation, vertebral fracture, or spinal stenosis) or whose pain resulted from a job injury since these patients are directed to the GH Occupational Medicine clinic. Otherwise, all patients visiting GH PCPs for non-specific back pain were eligible for the study.

Within one week of their visits for back pain, patients were mailed letters informing them that GH was surveying patients in an effort to improve its care for back pain. Patients receiving care from the intervention clinics were not aware that a QI activity was occurring. The letter stated that a research specialist would call them within a few days to invite their participation. A phone number was provided for those not wishing to be contacted. Research specialists called patients between 1 and 3 weeks after their back pain visits to explain the study, answer questions, confirm eligibility and obtain verbal informed consent for participation in a baseline and two follow-up interviews. To maximize participation and follow-up rates, we offered patients \$20 for completing each questionnaire. This recruitment process continued until we reached our required sample size for the pre-implementation and the post-implementation periods.

Data were collected by trained interviewers using a computer-assisted telephone interview (CATI) version of the questionnaires to minimize errors and missing data. We tracked the disposition of the consecutive patients seen in the primary care clinics, noting how many were successfully contacted by phone, agreed to participate, and completed each questionnaire.

### Patient outcome measures (Table 3)

#### Primary outcomes

During both the pre- and post-intervention periods we measured outcomes between about 2 weeks (range: 1 to 3 weeks) after the index clinician visit (baseline) and again two and six months later. We included two primary patient outcome measures. Back-related physical function in the previous week was measured with the modified Roland-Morris Disability Questionnaire (RMDQ) [31]. This instrument asks 23 yes/no questions selected for their relevance for patients with back problems. The RMDQ has been found reliable, valid, and sensitive to clinical changes [32, 33]. We also asked patients to rate their back pain severity during the previous week on a 0-to-10 scale where 0 represents “no pain” and 10 “pain as bad as it could be.” Such numerical rating scales have been found to be valid and reliable measures of pain [34].

**Table 3** Key Content of Baseline and Follow-up Assessments

Measures	Baseline	2-month	6-month
<b>Baseline Characteristics</b>			
Patient characteristics: education, income, employment). (Age, gender, race, and ethnicity extracted from EHR)	x		
Back pain problem (duration, pain elsewhere)	x		
Risk of Poor Outcome (STaT Back tool)	x	x	x
<b>Primary Outcomes</b>			
Back pain-related physical function (RMDQ)	x	x	x
Back pain severity (0-10 scale)	x	x	x
<b>Secondary Outcomes</b>			
Depression (PHQ-8)	x	x	x
Anxiety (GAD-7)	x	x	x
Fear of movement (TKS-10 item version)	x	x	x
Global improvement (PGIC)		x	x
Self-efficacy (PSEQ)	x	x	x
Patient satisfaction with caring, information, treatment effectiveness	x	x	x
Work loss in past 7 days due to LBP (hours); Effect on work productivity (0 to 10 scale) (2 items from WPAl)	x	x	x
<b>Treatment-Related Information</b>			
Helpfulness of treatments recommended by PCPs		x	x
Medications used in past week; change in medication use over time	x	x	x
Out-of-pocket expenses		x	x

**Abbreviations:** EHR, Electronic Medical Record; RMDQ, Roland Morris Disability Questionnaire; PHQ, Patient Health Questionnaire; GAD, Generalized Anxiety Disorder scale; TKS, Tampa Kinesiophobia Scale; PGIC, Patient Global Impression of Change scale; PSEQ, Patient Self Efficacy Questionnaire; WPAl, Work Productivity and Activity Impairment

#### Secondary outcomes

We measured the following secondary outcomes:

- *Depression* was measured with the Patient Health Questionnaire-8 (PHQ-8) [35], which is identical to the PHQ-9, but does not assess suicidal ideation. It has been found reliable, valid, and responsive to change [36, 37].
- *Anxiety* was measured with the Generalized Anxiety Disorder scale (GAD-7), which has good reliability and criterion, construct, factorial, and procedural validity in primary care populations [38, 39].
- *Fear of movement* was measured by an abbreviated version of the Tampa Scale for Kinesiophobia, a validated 17-item questionnaire that quantifies excessive fear of (re)injury due to movement in pain

patients [40]. The 10-item version was developed for another trial [41] had a measure of internal consistency of 0.66 [Michael von Korff, personal communication, May 1, 2016]. The total score of the 10 items was adjusted to yield a total score comparable to the 17-item version by multiplying the average item score for the items answered by 17.

- *Global improvement with treatment.* We used the Patient Global Impression of Change (PGIC) scale [42], a single question asking participants to rate improvement with treatment on a seven-point scale ranging from “very much improved” to “very much worse,” with “no change” as the mid-point.
- *Self-efficacy* was measured using the 10-item Pain Self Efficacy Questionnaire (PSEQ). It shows good internal consistency and construct validity [43].
- *Satisfaction with care* was measured using a 10-item instrument that has been validated and able to distinguish among three dimensions of satisfaction (caring, information and treatment effectiveness) [44].
- *The Work Productivity and Activity Impairment (WPAI)* is a self-administered instrument used to evaluate the impact of back pain on productivity. It assesses time missed from work (absenteeism), impairment while working/reduced on-the-job effectiveness (presenteeism), and overall work productivity loss (absenteeism v. presenteeism). It has been found reliable, valid, and responsive to change for several medical conditions [45].
- *Use of back-related medications and exercise* in the past week and treatments (e.g., yoga) paid for out of pocket (which are not captured in the GH database).

**Statistical analysis plan and sample size calculations**

**Statistical analysis plan**

The primary purpose of this trial is to estimate the effectiveness of the risk stratification strategy for reducing patient dysfunction and pain severity related to back pain. We also are conducting pre-specified subgroup analyses to assess risk-strata-specific differences in physical function and pain severity in the intervention and

control group since we hypothesize there will be more benefit in the medium and high risk strata.

To evaluate the overall effectiveness of the proposed risk stratification intervention we estimate the difference in change scores between the control and intervention groups, while accounting for possible effects due solely to the difference in calendar time between the pre- and post-implementation periods (Table 4). Potential effects such as seasonal trends and health care system policy changes will be accounted for in our analyses by using information from concurrent control clinics during the same period of time. The primary analysis time point for the study is 6-months following baseline, though we will also evaluate 2-month changes.

Differences in change scores by intervention group and time period will be estimated using linear mixed models (LMM) [46] with random effects to control for correlation within provider, clinic, or both. Including random effects in the model will yield valid statistical inference while increasing the overall statistical efficiency of the pre- and post-treatment design. The general mean model framework will be the following:

$$E(Y) = \beta_0 + \beta_1 CompStrat + \beta_2 Post + \beta_3 CompStrat * Post + \vec{\beta}_4 \vec{Z},$$

where  $Y$  is the change score for the specified follow-up time ( $t = 2$  or  $6$  months),  $CompStrat$  indicates randomization to either a risk stratification or a control clinic,  $Post$  is 1 if the patient entered the study during the post-intervention period and 0 if they entered in the pre-intervention period, and  $\vec{Z}$  is a vector of potential baseline confounders to be controlled for in the analysis. Baseline variables that have previously been shown to be related to dysfunction and pain intensity outcomes include age, gender, RMDQ and pain severity scores, duration of pain, medication use, work-related exertion, overall health status, and psychological issues related to health [47–49]. To control for any additional confounding, we will also adjust for baseline variables that are associated with the changes in RMDQ and pain intensity

**Table 4** Estimates of 6-month change in primary outcomes estimated from proposed model, by period of time and randomization group

Period of time	Usual Care (UC) Clinics	Risk Stratification (RS) Clinics	Column Difference
Pre-implementation* of risk stratification (RS)	$\beta_0$ : Change score in UC clinics pre-implementation of RS	$\beta_0 + \beta_1$ : Change score in RS clinics pre-implementation of RS	$\beta_1$ : Difference in change scores due to RS in the pre-implementation period.
Post-implementation of risk stratification (RS)	$\beta_0 + \beta_2$ : Change score in UC clinics post-implementation of RS	$\beta_0 + \beta_1 + \beta_2 + \beta_3$ : Change score in RS clinics post-implementation of RS	$\beta_1 + \beta_3$ : Difference in change scores due to RS in the post-implementation period.
Row Difference	$\beta_2$ : Difference in change scores due solely to time (implementation period).	$\beta_2 + \beta_3$ : Difference in change scores due to time and RS.	$\beta_3$ : Difference in change scores due solely to the RS implementation.

Abbreviations: UC, usual care, RS, risk stratification

\*Clinics randomized to implement the risk stratification intervention will do so only in the latter of the two time periods



and are imbalanced between the intervention groups, or between pre and post time periods.

This model framework allows for the calculation of the confounder-adjusted difference between change scores in the risk stratification and control groups ( $\beta_3$ ) taking into account changes due solely to time period ( $\beta_2$ ). Table 4 summarizes quantities of interest and their respective parameter estimates from the model. The model presented here is a simplified version of what will actually be used for the final analysis. In the complete model we will include both follow-up times in the same model to more efficiently estimate the effects at each time point (efficiency gained from estimating confounder model parameters just once). We will use the same model framework to calculate the risk-stratum specific estimates and secondary outcomes. For binary secondary outcomes we use generalized linear mixed models (GLMM) [50] with logistic or log link functions to estimate odds ratios or relative risks instead of mean change scores.

We gave respondents \$20 per completed follow-up questionnaire to minimize missing data. Further, we are using a model framework of LMM and GLMM models that assume data are missing at random given baseline confounders. We will also conduct sensitivity analyses using imputation to assess the missing data assumptions.

#### **Sample size calculation**

We based our power on detecting clinically meaningful differences of 2.5 points on the RMDQ score and 1.5 points on the pain severity score reported in the literature [31]. We assumed that within the medium- and high- risk groups clinically meaningful differences will be observed in the change scores between intervention and control groups, but within the low- risk group we assumed that there will not be significant differences. We assumed that the baseline distribution of the risk stratification groups will be 40 % low- risk, 40 % medium- risk and 20 % high- risk [51]. Therefore the overall effect difference between the intervention and control groups would be 1.5 points ( $0 \cdot 0.4 + 2.5 \cdot 0.4 + 2.5 \cdot 0.2$ ) for RMDQ and 0.9 points for pain intensity. These differences in mean change scores correspond to standardized effect sizes of 0.30 and 0.36 for RMDQ and pain severity, respectively. We assume a standard deviation of 5 points for the RMDQ score and 2.5 for pain intensity based on previous studies [52–54]. For simplicity we have determined the sample size assuming no correlation of outcomes within provider or clinic, yielding conservative estimates of sample size (larger sample size than may be required). Most cluster randomized designs require an increase in sample size relative to independent randomization, but because we have included a cluster with cross-over design (pre versus post) it is a more efficient design than independent randomization. The same model specified in the previous statistical

analysis section was used for the sample size and power calculations. For simplification we have set  $\beta_0$ ,  $\beta_1$ , and  $\beta_2$  all to 0 since their values do not affect the power of  $\beta_3$ , the estimate of interest. When estimating  $\beta_3$  we use the full model including estimates for  $\beta_0$ ,  $\beta_1$ , and  $\beta_2$  to arrive at the correct power estimates. Power calculations assumed two-sided tests at the 0.05  $\alpha$ -level and were performed using R version 2.14.0 for Windows XP Professional.

A total sample size of 1,410 participants provides 80 % power to detect a between-group, overall, standardized effect size of 0.30 on the RMDQ. This sample size provides approximately 92 % power to detect a standardized effect of 0.36 for pain intensity. We assumed that 20 % of participants would drop out of the study and thus inflated the estimated sample size of 1,410 to 1,763. The inflated sample size of 1,763 was divided into four equal groups corresponding to intervention and control clinics in both the pre- and post-intervention. The result was rounded down to a final sample size of 1,760 participants (4 groups of 440). Within risk strata groups we estimate >90 % power to detect an effect size of 0.50 (2.5 points on the RMDQ score) for the medium- and high- risk groups combined, and ~80 % power to detect an effect size of 0.68 (3.4 points on RMDQ) for just the high- risk group. Similarly, for pain intensity, the study will have >95 % power to detect an effect size of 0.60 (1.5 points) for the pooled medium- and high- risk groups, and ~80 % power to detect an effect size of 0.68 (1.7 points) in the high- risk group. Therefore we have adequate power for all comparisons of interest. Additionally, though there may not be sufficient power to detect the noted effect sizes, we will also estimate the effect of the intervention separately in each of the risk strata.

#### **Effect on utilization and quality of care**

The same analytic approach used to evaluate patient outcomes will be used to evaluate the effect of the intervention on utilization of health care services for back pain. We will use EHR data to determine the percentage of patients with non-specific back pain in the intervention clinics who had a STarT Back risk score recorded and the effect of the intervention on the frequencies with which specific tests and treatments for back pain were used. We will also determine if the observed changes in utilization following the intervention indicated improvements in the quality of care. Specifically, we will examine if the use of guideline-recommended treatments for medium and high risk patients (i.e., PT which includes exercise, CAM, CBT) increased and the use of treatments that are generally not recommended for non-specific back pain (e.g., advanced imaging, opioids, neurosurgery visits, and epidural steroid injections) decreased. These analyses will be done for all patients and within each risk stratum. We will conduct two sets

of analyses, one including only patients who agreed to participate in the follow-up study and the other including all patients who were eligible to participate. The primary analysis will include all patients since this will best capture the overall effect of the intervention. Comparison of the two analytic populations will allow us to determine how representative patients agreeing to participate were of all patients in terms of utilization.

## Discussion

This trial will provide the first randomized trial evidence of the effectiveness of implementing risk stratification in a US setting based on the STarT Back approach. Major strengths of our trial are the use of an unbiased method of allocation (randomization) of PCP clinics to the intervention or control groups, inclusion of a pre-intervention period to adjust the analyses for baseline differences between the intervention and control clinics, and the use of a multi-faceted and comprehensive approach conducted in the context of a system-supported quality improvement strategy. The main limitations are the restriction to clinics in one health care system, the inclusion of a small number of clinics, and the collection of “baseline” data about 2 weeks following the back pain visit. The findings of this trial will be disseminated through publications in targeted journals, presentation to the diverse group of national advisors representing key stakeholders assembled for this study, presentation at conferences, and reports by the study sponsors. Given the effectiveness of the STarT Back approach in the UK’s National Health Service setting and the hope of health care systems in many other settings and countries that implementing a similar approach will lead to improvements in care for back pain, the results of this trial will be of great interest.

## Additional file

**Additional file 1:** Non-specific Back Pain Assessment, Management, and Follow-up Guideline. (DOCX 176 kb)

## Acknowledgments

NIHR Research Professorship (NIHR-RP-011-015): This 5 year award funds Dr. Foster’s time so she can collaborate on projects without the need for each project to cost her into their costs. It also funded the collaborative visit for Dr. Foster and Dr. Hill to travel to Seattle for an early collaborative meeting a few years ago. It had no role in the decision to publish.

## Funding

Funding for this trial was provided by the Patient Centered Care Research Institute (“Evaluation of a Patient-Centered Risk Stratification Method for Improving Primary Care for Back Pain”: Contract #398) and by the National Center for Complementary and Integrative Health/NIH (“Implementing Evidence-Based Treatments for Persistent Back Pain into Primary Care”: Grant #R21AT0007326). Martin Levine, Diane Piekara, and Pam Rock received support to participate in the quality improvement activities from Group Health. Nadine Foster was supported through an NIHR Research Professorship (NIHR-RP-011-015). The views expressed are those of the authors and not necessarily those of the NHS, the NIHR or the Department of Health.

## Availability of data and material

Not applicable as this is a protocol for a study.

## Authors’ contributions

The major contributors to development of each the critical components for this intervention and evaluation plan are described below. Development of the content and implementation strategy of the quality improvement intervention: BB, DC, KE, RH, ML, DP, PR, KS, and GS. Development of strategies for recruiting patients and measuring patient outcomes: BB, GB, DC, NF, RH, JH, MJ, AL, KS, and JY. Study design and quantitative evaluation methods: BB, DC, AC, NF, JH, RH, KS, and RW. Development and conduct of qualitative evaluations prior to and following implementation: SE and CH. All authors have read and approved the final version.

## Authors’ information

Our research team included 4 “patient partners” (GB, MJ, AL, JY) with chronic back pain who provided their unique and valuable expert advice on the study materials and processes.

## Competing interests

The authors declare that they have no competing interests.

## Consent for publication

Not applicable.

## Ethics approval and consent to participate

All research activities and materials were reviewed and approved by the GH Human Subjects Review Committee (Ref: 2013 MATCH). Because this intervention was delivered in the context of a quality improvement activity in 3 primary care clinics, consent was required only for collecting outcomes data. Informed consent was obtained from the subjects.

## Author details

<sup>1</sup>Group Health Research Institute, Seattle, USA. <sup>2</sup>Patient Partner, Seattle, USA. <sup>3</sup>Group Health Cooperative, Seattle, USA. <sup>4</sup>NIHR Professor of Musculoskeletal Health in Primary Care, Keele University, Keele, UK. <sup>5</sup>Arthritis Research UK Primary Care Centre, Research Institute for Primary Care and Health Sciences, Keele, UK. <sup>6</sup>Primary Care & Health Sciences, Keele University, Keele, UK. <sup>7</sup>Group Health Research Institute, Center for Community Health and Evaluation, Seattle, USA. <sup>8</sup>Iora Health, Shoreline, USA. <sup>9</sup>IMPACT Service, Stoke-on-Trent, UK.

Received: 9 July 2016 Accepted: 13 August 2016

Published online: 24 August 2016

## References

1. Institute of Medicine. *Relieving Pain in America: A Blueprint for Transforming Prevention, Care, Education, and Research*. Washington, DC: The National Academies Press; 2011.
2. Walker BF. The prevalence of low back pain: a systematic review of the literature from 1966 to 1998. *J Spinal Disord*. 2000;13(3):205–17. PubMed PMID: 10872758. Review.
3. Luo X, Pietrobon R, Sun SX, Liu GG, Hey L. Estimates and patterns of direct health care expenditures among individuals with back pain in the United States. *Spine*. 2004;29:79–86.
4. Stewart WF, Ricci JA, Chee E, Morganstein D, Lipton R. Lost productive time and cost due to common pain conditions in the US workforce. *JAMA*. 2003; 290:2443–54.
5. Martin BI, Deyo RA, Mirza SK, Turner JA, Comstock BA, Hollingworth W, Sullivan SD. Expenditures and health status among adults with back and neck problems. *JAMA*. 2008;299:656–64.
6. Von Korff M, Kolodny A, Deyo RA, Chou R. Long-term opioid therapy reconsidered. *Ann Intern Med*. 2011;155:325–8. PMID: PMC3280085.
7. Von Korff MR, Franklin G. Responding to America’s iatrogenic Epidemic of Prescription Opioid Addiction and Overdose. *Med Care*. 2016;54(5):426–9.
8. Moore JE. Chronic low back pain and psychosocial issues. *Phys Med Rehabil Clin N Am*. 2010;21(4):801–15. doi:10.1016/j.pmr.2010.06.005. Review.
9. Boothby JL, Thorn BE, Stroud MW, Jensen MP. Coping with pain. In: Gatchel RJ, Turk DC, editors. *Psychosocial factors in pain: critical perspectives*. New York: Guilford Press; 1999. p. 343–59.

10. Jensen MP, Karoly P. Control beliefs, coping efforts, and adjustment to chronic pain. *J Consult Clin Psychol*. 1991;59:431–8.
11. Keefe FJ, Rumble ME, Scipio CD, Giordano LA, Perri LM. Psychological aspects of persistent pain: current state of the science. *J Pain*. 2004;5:195–211.
12. Hill JC, Dunn KM, Lewis M, Mullis R, Main CJ, Foster NE, Hay EM. A primary care back pain screening tool: identifying patient subgroups for initial treatment. *Arthritis Rheum*. 2008;59:632–41.
13. Hill JC, Whitehurst DG, Lewis M, Bryan S, Dunn KM, Foster NE, Konstantinou K, Main CJ, Mason E, Somerville S, Sowden G, Vohora K, Hay EM. Comparison of stratified primary care management for low back pain with current best practice (STarT Back): a randomised controlled trial. *Lancet*. 2011;378:1560–71. PMID: 3208163.
14. Foster NE, Mullis R, Hill JC, Lewis M, Whitehurst DG, Doyle C, Konstantinou K, Main C, Somerville S, Sowden G, Wathall S, Young J, Hay EM, team IMBS. Effect of stratified care for low back pain in family practice (IMPACT Back): a prospective population-based sequential comparison. *Ann Fam Med*. 2014; 12:102–11. PMID: 3948756.
15. Whitehurst DG, Bryan S, Lewis M, Hay EM, Mullis R, Foster NE. Implementing stratified primary care management for low back pain: cost-utility analysis alongside a prospective, population-based, sequential comparison study. *Spine (Phila Pa 1976)*. 2015;40:405–14.
16. Hay EM, Dunn KM, Hill JC, Lewis M, Mason EE, Konstantinou K, Sowden G, Somerville S, Vohora K, Whitehurst D, Main CJ. A randomised clinical trial of subgrouping and targeted treatment for low back pain compared with best current care. The STarT Back Trial Study Protocol. *BMC Musculoskelet Disord*. 2008;9:58. PMID: 2377248.
17. Hill JC, Vohora K, Dunn KM, Main CJ, Hay EM. Comparing the STarT back screening tool's subgroup allocation of individual patients with that of independent clinical experts. *Clin J Pain*. 2010;26:783–7.
18. Sowden G, Hill J, Konstantinou K, Khanna M, Main C, Salmon P, Somerville S, Wathall S, Foster NE. Subgrouping for targeted treatment in primary care for low back pain: the treatment system and clinical training programmes used in the IMPACT Back study (ISRCTN 55174281). *Fam Pract*. 2012;29(1):50–62.
19. Main CJ, Sowden G, Hill JC, Watson PJ, Hay EM. Integrating physical and psychological approaches to treatment in low back pain: the development and content of the STarT Back trial's 'high-risk' intervention (StarT Back; ISRCTN 37113406). *Physiotherapy*. 2012;98(2):110–7.
20. Carey TS, Frebarger J. Physical therapy for low back pain: what is it, and when do we offer it to patients? *Ann Fam Med*. 2014;12:99–101. PMID: PMC3948755.
21. Murphy SE, Blake C, Power CK, Fullen BM. Comparison of a Stratified Group Intervention (STarT Back) with Usual Group Care in Patients with Low Back Pain: A Non-randomised Controlled Trial. *Spine (Phila Pa 1976)*. 2015.
22. Karlen E, McCathie B. Implementation of a quality improvement process aimed to deliver higher-value physical therapy for patients with Low back pain: case report. *Phys Ther*. 2015;95(12):1712–21.
23. Karstens S, Krug K, Hill JC, Stock C, Steinhäuser J, Szecsenyi J, et al. Validation of the German version of the STarT-Back Tool (STarT-G): a cohort study with patients from primary care practices. *BMC Musculoskelet Disord*. 2015;16:346.
24. Karstens S, Joos S, Hill JC, Krug K, Szecsenyi J, Steinhäuser J. General practitioners views of implementing a stratified treatment approach for Low back pain in Germany: a qualitative study. *PLoS One*. 2015;10(8):e0136119.
25. Matsudaira K, Oka H, Kikuchi N, Haga Y, Sawada T, Tanaka S. Psychometric properties of the Japanese version of the STarT back tool in patients with low back pain. *PLoS One*. 2016;11(3):e0152019.
26. Chou R, Qaseem A, Snow V, Casey D, Cross Jr JT, Shekelle P, Owens DK. 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;147:478–91.
27. Practitioners RCoG. Low Back Pain: Early Management of Persistent Non-specific Low Back Pain National Institute for Health and Clinical Excellence: Guidance, 2009
28. Chou R, Deyo R, Friedly J, et al. Noninvasive treatments for Low back pain. Comparative effectiveness review No. 169. (prepared by the pacific northwest evidence-based practice center under contract No. 290-2012-00014-1) AHRQ publication No. 16-EH. Rockville: Agency for Healthcare Research and Quality; 2016. [www.effectivehealthcare.ahrq.gov/reports/final.cfm](http://www.effectivehealthcare.ahrq.gov/reports/final.cfm).
29. O'Keefe M, Purtill H, Kennedy N, et al. Comparative Effectiveness of Conservative Interventions for Nonspecific Chronic Spinal Pain: Physical, Behavioral/Psychologically Informed, or Combine? A Systematic Review and Meta-Analysis. *J Pain*. 2016 Feb 1
30. Sowden G, Hill JC, Konstantinou K, Khanna M, Main CJ, Salmon P, Somerville S, Wathall S, Foster NE. Targeted treatment in primary care for low back pain: the treatment system and clinical training programmes used in the IMPACT Back study (ISRCTN 55174281). *Fam Pract*. 2012;29:50–62. PMID: 3261797.6 month.
31. Patrick DL, Deyo RA, Atlas SJ, Singer DE, Chapin A, Keller RB. Assessing health-related quality of life in patients with sciatica. *Spine*. 1995;20:1899–908.
32. Underwood MR, Barnett AG, Vickers MR. Evaluation of two time-specific back pain outcome measures. *Spine*. 1999;24:1104–12.
33. Roland M, Fairbank J. The Roland-Morris Disability Questionnaire and the Oswestry Disability Questionnaire. *Spine*. 2000;25:3115–24.
34. Hjermstad MJ, Fayers PM, Haugen DF, Caraceni A, Hanks GW, Loge JH, Fainsinger R, Aass N, Kaasa S, European Palliative Care Research C. Studies comparing numerical rating scales, verbal rating scales, and visual analogue scales for assessment of pain intensity in adults: a systematic literature review. *J Pain Symptom Manage*. 2011;41:1073–93.
35. Kroenke K, Strine TW, Spitzer RL, Williams JB, Berry JT, Mokdad AH. The PHQ-8 as a measure of current depression in the general population. *J Affect Disord*. 2009;114:163–73.
36. Lowe B, Unutzer J, Callahan CM, Perkins AJ, Kroenke K. Monitoring depression treatment outcomes with the patient health questionnaire-9. *Med Care*. 2004;42:1194–201.
37. Kroenke K, Spitzer RL, Williams JB. The PHQ-9: validity of a brief depression severity measure. *J Gen Intern Med*. 2001;16:606–13. PMID: 1495268.
38. Spitzer RL, Kroenke K, Williams JB, Lowe B. A brief measure for assessing generalized anxiety disorder: the GAD-7. *Arch Intern Med*. 2006;166:1092–7.
39. Lowe B, Decker O, Muller S, Brahler E, Schellberg D, Herzog W, Herzberg PY. Validation and standardization of the Generalized Anxiety Disorder Screener (GAD-7) in the general population. *Med Care*. 2008;46:266–74.
40. Crombez G, Vlaeyen JW, Heuts PH, Lysens R. Pain-related fear is more disabling than pain itself: evidence on the role of pain-related fear in chronic back pain disability. *Pain*. 1999;80:329–39. PMID: DC - Pain.
41. Von Korff M, Balderson BHK, Saunders K, Miglioretti DL, Lin EHB, Berry S, Moore JE, Turner JA. *Pain*. 2005;113:323–30.
42. Guy W. *ECDEU Assessment Manual for Psychopharmacology*. Washington, D.C.; 1976
43. Nicholas MK. The pain self-efficacy questionnaire: taking pain into account. *Eur J Pain*. 2007;11:153–63.
44. Cherkin D, Deyo RA, Berg AO, Bergman JJ, Lishner DM. Evaluation of a physician education intervention to improve primary care for low-back pain. I Impact on physicians. *Spine*. 1991;16:1168–72.
45. Reilly MC, Zbrozek AS, Dukes EM. The validity and reproducibility of a work productivity and activity impairment instrument. *Pharmacoeconomics*. 1993; 4:353–65.
46. Laird NM, Ware JH. Random-effects models for longitudinal data. *Biometrics*. 1982;38:963–74.
47. Dunn KM, Jordan KP, Croft PR. Contributions of prognostic factors for poor outcome in primary care low back pain patients. *Eur J Pain*. 2011;15:313–9. PMID: 3062783.
48. Hayden JA, Chou R, Hogg-Johnson S, Bombardier C. Systematic reviews of low back pain prognosis had variable methods and results: guidance for future prognosis reviews. *J Clin Epidemiol*. 2009;62:781–96. e1.
49. Grotle M, Foster NE, Dunn KM, Croft P. Are prognostic indicators for poor outcome different for acute and chronic low back pain consulters in primary care? *Pain*. 2010;151:790–7.
50. Pinheiro JC, Bates DM. Approximations to the Log-likelihood function in the nonlinear mixed-effects model. *J Comp Graph Stat*. 1995;4:12–35.
51. Foster NE, Mullis R, Young J, Doyle C, Lewis M, Whitehurst D, Hay EM. IMPACT Back study protocol. Implementation of subgrouping for targeted treatment systems for low back pain patients in primary care: a prospective population-based sequential comparison. *BMC Musculoskelet Disord*. 2010;11:186. PMID: 2940850.
52. Cherkin DC, Sherman KJ, Avins AL, Erro JH, Ichikawa L, Barlow W, Delaney K, Hawkes R, Hamilton L, Pressman A, Khalsa PS, Deyo RA. A randomized controlled trial comparing acupuncture, simulated acupuncture, and usual care for chronic low back pain. *Arch Intern Med*. 2009;169:858–66. PMID: PMC2832641.
53. Cherkin D, Sherman K, Kahn J, Wellman R, Cook AJ, Johnson E, Erro J, Delaney K, Deyo R. A comparison of the effects of 2 types of massage and usual care on chronic low back pain. *Ann Intern Med*. 2011;155:1–9.
54. Sherman KJ, Cherkin DC, Wellman RD, Cook AJ, Hawkes RJ, Delaney K, Deyo RA. A randomized trial comparing yoga, stretching, and a self-care book for chronic low back pain. *Arch Intern Med*. 2011;171:2019–26. PMID: PMC3279296.