Can Using Patient Reports of Low Back Pain Help to Better Direct Patients to Treatments?

Dan Cherkin, PhD[1], Benjamin Balderson, PhD[1] ,Rob Wellman, MS[1] , Clarissa Hsu, PhD[1,2] , Karen J. Sherman, PhD[1] , Sarah C. Evers, MPH[1] , Rene Hawkes, BS[1] , Andrea Cook, PhD[1] , Martin D. Levine, MD[3] , Diane Piekara, PT[4] , Pam Rock, PT[4] , Katherine Talbert Estlin, MD[5] , Georgie Brewer[9], Mark Jensen[9], Anne-Marie LaPorte[9], , John Yeoman[9], , Gail Sowden, MSc[6] , Jonathan C. Hill, PhD[6] , Nadine E. Foster, DPhil[6]

[1] Group Health Research Institute, 1730 Minor Ave, Suite 1600, Seattle, WA 98101

[2] Iora Health, 15214 Aurora Avenue North, Shoreline, WA 98133

[3] Group Health Cooperative, 320 Westlake Ave N #100, Seattle, WA, 98109

[4] Arthritis Research UK Primary Care Centre, Research Institute for Primary Care and Health Sciences, Keele University, Keele, Staffordshire, ST5 5BG, UK

[5] IMPACT Service, Staffordshire and Stoke on Trent Partnership NHS Trust, Haywood Hospital, High Lane, Burslem, Stoke-on-Trent, ST6 7AG, UK

[6] Primary Care & Health Sciences, Keele University, Keele, Staffordshire, ST5 5BG, UK

[7] NIHR Professor of Musculoskeletal Health in Primary Care, Keele University, Staffordshire, UK, ST5 5BG

[8] Arthritis Research UK Primary Care Centre, Research Institute for Primary Care and Health Sciences, Keele University, Keele, Staffordshire, ST5 5BG, UK

[9] Patient Partner

Original Project Title: Evaluation of a Patient-Centered Risk Stratification Method for Improving Primary Care for Back Pain

PCORI ID: 398

HSRProj ID: 20143334

ClinicalTrials.gov ID: NCT02286141

To cite this document, please use: Cherkin D, Balderson B, Wellman R.et al. 2018. *Can Using Patient Reports of Low Back Pain Help to Better Direct Patients to Treatments?* Washington, DC: Patient-Centered Outcomes Research Institute (PCORI). https://doi.org/10.25302/11.2018.CER.398

Table of Contents

ABSTRACT **3**

ABREVIATIONS **5**

BACKGROUND **6**

STAKEHOLDER ENGAGEMENT 7

METHODS **8**

Design and Setting 8
Participating Clinics (Clusters) 9
Patient Inclusion and Exclusion Criteria 10
Randomization and Blinding 10
The Intervention 10
Outcomes 13
Data Collection 13
Sample Size 14
Statistical Methods 15

RESULTS 16

Patient Recruitment and Follow-up 16
Patient Characteristics 16
Effect of the Intervention 21

DISCUSION **29**

Context for Study Results 29
Generalizability of the Findings and Implementation of Study Results 30
Subpopulation Considerations 30
Study Limitations 30

CONCLUSION **31**

REFERENCES **32**

ACKNOWLEGEMENTS **34**

APPENDIX **35**

ABSTRACT

Background: The Subgroups for Targeted Treatment Risk Stratification (STarT) Back strategy for categorizing and treating patients with low back pain (LBP) improved patients' physical function while reducing costs in the United Kingdom. This trial evaluated the effect of implementing an adaptation of this approach in a health care setting in the United States.

Methods: The MATCH (Matching Appropriate Treatments to Consumers' Healthcare needs) trial was a pragmatic cluster randomized trial with a preintervention baseline period. T recruited patients for a baseline phase of observation and follow-up, then taught the providers how to conduct the interventions, and then recruited another study population and followed them for 6 months. We recruited patients from primary care clinics in an integrated health care system in western Washington State. Six primary care clinics were pair-randomized, 3 to training in the STarT Back strategy and 3 to serve as controls. We invited adult patients receiving primary care for nonspecific LBP to provide data 2 weeks after their primary care visit and follow-up data 2 and 6 months (primary endpoint) later. The STarT Back risk stratification strategy matches treatments for LBP to physical and psychosocial obstacles to recovery using patient-reported data (the STarT Back Tool) to categorize patients at low, medium or high risk of persistent disabling pain. Primary care providers in the intervention group attended 6 didactic sessions to improve their understanding of the management of LBP and received in-person training in the use of the risk stratification tool that had been incorporated into electronic health records. Physical therapists in the intervention clinics received 5 days of intensive training. Primary care providers and physical therapists in the control clinics received no training. We collected patient-reported data through telephone interviews.

Primary outcomes were back-related physical function and pain severity. We estimated intervention effects by comparing mean changes in patient outcomes at 2 and 6 months follow-up between intervention and control clinics. We estimated differences in change scores by trial arm and time period using linear mixed effect models. Secondary outcomes included patients'

health care utilization (e.g., physician and physical therapy visits, imaging studies, opioid prescriptions) using data from electronic health records**.**

Results: A total of 2138 patients with LBP visited the intervention clinics and 2571 the control clinics over the course of the study. Overall, 36% of patients provided baseline data on patient outcomes. Follow-up rates were 93% at 2 months and 91% at 6 months. Participation and follow-up rates were similar in the intervention and control groups. There was no significant difference between the intervention and control groups in the primary outcomes at 6 months. Specifically, mean improvement in function in the control group exceeded that in the intervention group by 0.50 on RMD scale (95% CI, –0.55-1.55; *P* = 0.349), and mean improvement in pain severity was 0.13 greater in the control group than in the intervention group (95% CI, –0.37-0.63; *p* = 0.61). The intervention had no significant effect on any patient outcomes at 2 or 6 months or on health care use.

Conclusions: A resource-intensive intervention to support stratified care for LBP in a US health care setting had no effect on patient outcomes or health care use.

Limitations: The main limitations of our evaluation were that less than half of patients visiting the clinics for LBP provided data for the outcomes analyses, and the trial was restricted to a single health care system.

Abbreviations

CAM: Complementary and alternative medicine

CBT: Cognitive behavioral therapy

EHR: Electronic health record

GAD: Generalized anxiety disorder

GH: Group Health

LBP: Low back pain

MATCH: Matching Appropriate Treatments to Consumers' Healthcare needs

NIHR:

OR: Odds ratio

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

RMDQ: Roland-Morris Disability Questionnaire

STarT: Subgroups for Targeted Treatment Risk Stratification

TSK: Tampa Scale of Kinesiophobia

WPAI: Work Productivity and Activity Impairment

BACKGROUND

Despite increasing US expenditures for low back pain (LBP), patient outcomes have deteriorated.[1] The current epidemic of opioid addiction and resultant deaths illustrates the urgency of finding safer and more effective approaches for treating chronic pain.[2,3] The traditional view of LBP as a largely biomedical problem[4] is being supplanted by the biopsychosocial model that acknowledges that, while pain usually has an underlying biological basis, psychosocial factors (e.g., pain beliefs/cognitions, distress, coping behaviors, social factors) also significantly influence the experience and impact of pain.[5,6] This broader conceptualization of chronic pain provides a clear rationale for incorporating cognitive behavioral principles into the management of distressed and disabled patients with LBP to minimize pain-related disability.

The Subgroups for Targeted Treatment Risk Stratification (STarT) approach, a promising strategy for categorizing and treating patients that considers both their physical and their psychosocial characteristics, was developed and evaluated in the United Kingdom.[7] This UK strategy focused on training physical therapists (PTs) and modifying their practice behavior. This approach improved patients' physical function and satisfaction with care while reducing costs.[8-11] The STarT Back strategy uses patient responses to a 9-item STarT Back tool questionnaire to allocate patients to a low-, a medium-, or a high-risk subgroup according to their risk of persistent disabling back pain. It takes less than 2 minutes to administer. Patients in each subgroup are then recommended evidence-based treatments matched to their prognostic profile.[7,12,13] Patients found to have at least 4 out of the 5 psychosocial risk factors (i.e., high pain bothersomeness, fear, worry, catastrophizing, depression) are considered high risk, and those with relatively few (i.e., 0 to 3) physical or psychosocial risk factors are considered low risk. The remaining patients, who have significant pain and/or activity limitations but fewer psychosocial risk factors, are considered medium risk. This tool, administered by general practitioners, has been validated in the United Kingdom for use with primary care adults with nonspecific LBP.[12] The success of this strategy in the United Kingdom

has generated high levels of interest in developed countries, providing new hope that meaningful improvements in primary care for LBP are within reach.[14-19]

To determine if the STarT Back risk stratification strategy would succeed in the United States, we conducted the MATCH (Matching Appropriate Treatments to Consumers' Healthcare needs) cluster randomized trial. This trial evaluates the effects of incorporating the STarT Back strategy into primary care practices within an integrated health care system. The goal was to give primary care providers (PCPs) and PTs the knowledge, tools, and confidence they needed to provide their patients with a broader understanding of their LBP, reassurance about their likely prognosis, and treatment options that matched their prognostic profile. We hypothesized that this intervention would improve patient outcomes by promoting the increased use of matched treatment options for patients in each subgroup, as determined by the STarT Back tool. To our knowledge, this is the first randomized and controlled evaluation of a risk stratification approach to improve care for back pain, based on the STarT Back strategy.

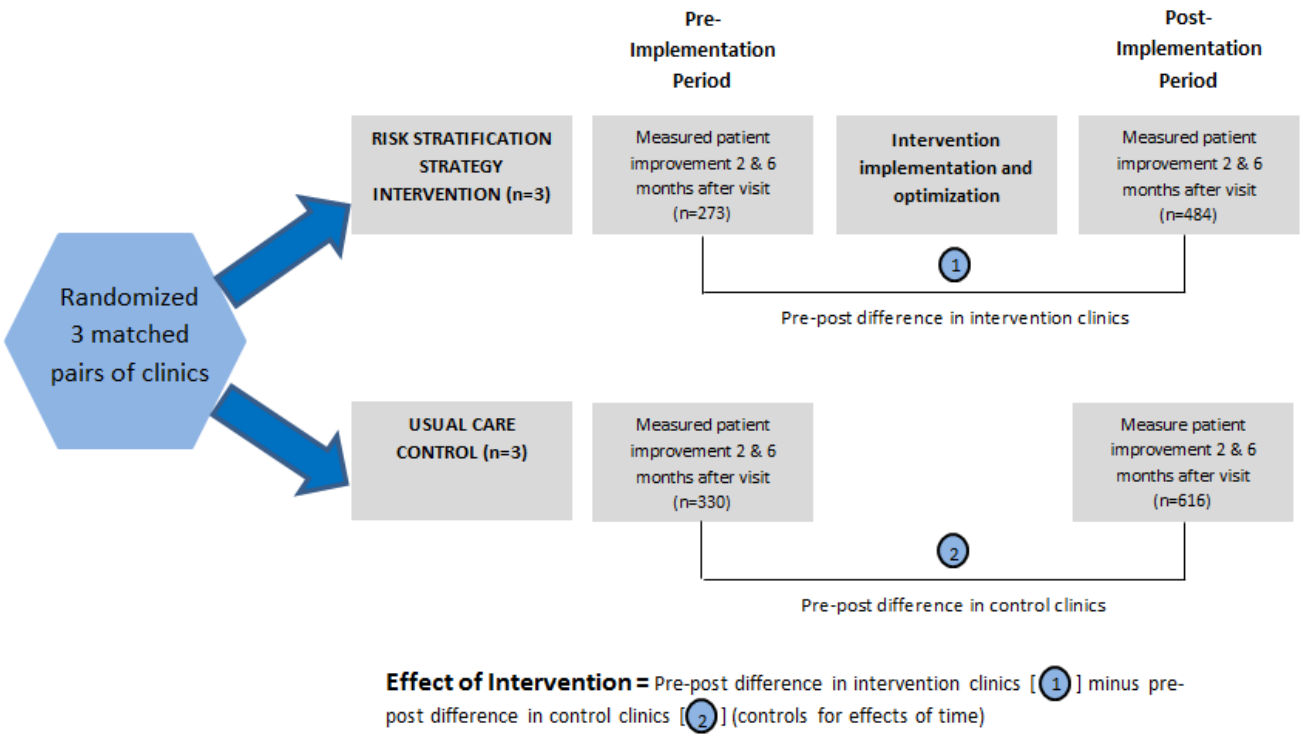
STAKEHOLDER ENGAGEMENT

We developed the intervention strategy between March 2013 and April 2014 with the support of key Group Health (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 3 key UK members (NF, JH, GS) on our team. We also invited 4 local individuals with extensive personal experience with chronic pain to serve on our project team (GB, MJ, AL, JY). Two were recruited through the governance office of our health care organization, and 2 were recommended by the executive director of the American Chronic Pain Association. These patient partners provided valuable perspectives on the implementation strategy; they identified ways that clinicians could more effectively communicate with patients and helped select outcome measures to use in our evaluation. The patient partners attended monthly team meetings, participating as full members of the team, and were included as coauthors on the protocol manuscript.[20]

To maximize the potential for efficiently implementing our intervention strategy in other US health care settings, we recruited 10 national advisers representing patients, large employers,

major governmental and independent health care systems, primary care practice networks, major government payers and insurers, and institutions that train complementary and alternative medical care (CAM) providers. Most advisers were in positions to influence policies within their organizations. The advisers participated in conference calls with the principal investigator (DC) and the project manager (RH) every 6 months, providing insights about the needs of the organizations they represented and advice on how to disseminate the findings should the results document the value of the risk stratification strategy under evaluation.

Figure 1: Study Design



METHODS

Design and Setting

The trial design has been reported in detail.[20] In brief, MATCH was a pragmatic, cluster (clinic) randomized trial with 2 parallel arms[21] each with a baseline data collection period.

Figure 1 shows that 3 clinics were randomized to receive the risk stratification intervention and 3 to continue to provide usual care without any intervention. During a pre-implementation phase, we collected data on patients' STarT Back risk levels and patient outcome measures at baseline, and after 2 and 6 months from those LBP patients in the intervention and control clinics who were willing to participate. This allowed estimation of the normal rate of recovery in the intervention and control clinics before the intervention period. This was followed by an intervention period wherein clinicians in the intervention clinics received training in the risk stratification approach. Once this training was completed, we again collected baseline and 2- and 6-month follow-up data from willing patients seen for LBP during this postintervention period. These data allowed estimation of the overall and risk level–specific effects of the intervention on patient outcomes—controlling for any differences that may have existed between the intervention and control groups during the preintervention period. We conducted the trial in the United States in an integrated health care delivery system serving more than 600 000 members: Group Health in Washington State. We collected data from patients and electronic health records (EHRs). GH partnered with the research team to evaluate the effect of stratified care in its primary care clinics. The intervention in the selected clinics was incorporated into a mandatory care improvement activity that was fully supported by clinical and administrative leadership at both the system and clinic levels.

Participating Clinics (Clusters)

We considered for inclusion GH's 10 largest primary care clinics in the Seattle metropolitan area, and we selected 3 pairs of geographically and socio-demographically similar clinics. All clinics agreed to participate. One clinic from each of the 3 pairs of similar clinics was randomized to the intervention, and the other clinics served as controls. All 6 clinics had an on-site PT department. The clinics had between 7 and 21 PCPs, most of whom worked part time. The 6 clinics served a total of 114 000 patients (range: 10 000-30 000).

Control clinics did not receive any intervention and continued to offer whatever care they normally provided. We measured preintervention levels of patient outcomes in all 6 clinics. We

then simultaneously implemented the intervention in the 3 intervention clinics over 6 months, after which we again assessed patient outcomes in both intervention and control clinics.

Patient Inclusion and Exclusion Criteria

We used the EHR to identify all patients 18+ years of age who received a primary diagnosis consistent with nonspecific LBP (e.g., lumbago, back pain not otherwise specified). To maintain broad applicability of the trial population, we excluded only patients with specific causes of their pain (e.g., pregnancy, disc herniation, vertebral fracture, spinal stenosis) or with job injuries, which were seen in the occupational medicine clinic. Otherwise, all patients visiting GH PCPs for nonspecific back pain were eligible for the study. Patients receiving care from the intervention and control clinics were not aware that a trial involving training clinicians in the intervention clinics was occurring.

Randomization and Blinding

Prior to the intervention, the trial biostatistician randomly assigned, by computer-generated random number, 1 clinic in each of the 3 geographically and socio-demographically matched pairs of clinics to receive the intervention and the other clinic to serve as the no-intervention control. We considered all eligible patients seen in the intervention clinics to have received the intervention.

The Intervention

We implemented the intervention in the 3 intervention clinics from April to September 2014. Key components of the intervention were incorporating the STarT Back tool[8] into the EHR, identifying recommended treatment options for patients in each risk subgroup, and training the primary care teams and PTs to use the tool and to use recommend treatments appropriate for patients in each risk subgroup.[20] Participation in the intervention was mandatory for the PCPs and PTs in the intervention sites, and time was made available for all of the intervention training. Clinicians in the control clinics did not receive any of the training received by clinicians in the intervention clinics. Although all Group Health Cooperative clinicians had access to the

STaRT Back tool in the EHR, only the clinicians in the intervention clinics were trained how to find it in EPIC, how to interpret the scores, and which treatments were most appropriate for patients in each risk stratum.

Incorporating Decision Support Tools Into Electronic Health Records: After clinicians or nursing staff entered patients' responses to the STarT Back questions, the EHR automatically calculated each patient's risk stratum and displayed the responses to each question, the patient's risk stratum, and the recommended treatment options for that stratum on a screen visible to clinicians and patients. The goal was to provide an opportunity for clinicians to discuss treatment options with their patients. To accommodate differences in practice style, individual PCPs were allowed to decide whether they or their medical assistants entered patients' responses to the STarT Back in the EHR. Because nursing staff sometimes collected and entered patient responses to the STarT Back tool, provider use of the tool varied greatly; therefore, it is unclear how many providers used the tools as intended with patients. In addition, shortcuts were incorporated into the EHR to help clinicians efficiently access the STarT Back tool, the Group Health back pain guidelines, existing GH educational resources (DVDs about acute and chronic back pain, and when surgery might be indicated), and GH's self-management groups for persons with chronic conditions (Living Well With Chronic Conditions). All such tools were available to all PCPs but were rarely used.

Identifying Recommended Treatment Options for Patients in Each Risk Stratum: We used the STarT Back tool without modification; however, we relied on GH's new low back pain guidelines to identify several evidence-based treatment options available at GH appropriate for each patient subgroup. This resulted in the following treatment recommendations for each risk stratum:

- <u>Low Risk (~ 40% of patients)</u>: Conduct a brief assessment to rule out potentially serious causes of back pain (i.e., "red flags"), do not refer to other health care professionals, elicit and listen to patients' concerns, and provide reassurance about the positive prognosis and self-care recommendations to relieve pain (e.g., physical activity, pain

medications, avoiding bed rest). Encourage patients to access online videos that reinforce information about acute or chronic back pain and the importance of self-care.

- Moderate Risk (~ 40% of patients): In addition to ruling out red flags and encouraging self-care, recommend activating treatments, such as PT-led exercise and yoga, that could reduce fear of movement. For patients not interested in activating treatments, consider more passive options (e.g., acupuncture, chiropractic, massage therapy) in the hope that these treatments will help decrease their pain and prepare them for more active approaches.
- High Risk (~ 20% of patients): In addition to ruling out red flags, recommend or refer patients to the GH PTs who are specially trained (described below) to offer patients in the intervention group a systematic approach to the integration of physical and psychological approaches to treatment of people with low back pain. Another evidence-based approach for chronic back pain, cognitive behavioral therapy (CBT) from a psychologist, could have been recommended, but access was very limited because of the lack of therapists trained in CBT for pain. PCPs were also encouraged to proactively follow-up with high-risk patients within 2 weeks.

Training for Primary Care Teams: Six 1-hour training sessions on separate topics were given in each of the 3 primary care intervention clinics. Each topic was presented on several occasions to ensure that all PCPs participated. Sessions were presented roughly monthly during the 6-month intervention 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 complementary and alternative medial (CAM) therapies. PTs and members of the nursing staff were invited to attend several of the sessions. In addition, PCPs and staff received coaching on how to locate and correctly use the STarT Back and other related tools in the EHR. Most PCPs participated in at least 1 such coaching session. Finally, to reduce knowledge barriers to recommending

matched treatment options, we compiled a list of the names and contact information of recommended local CAM providers and made them available to clinic staff.

Training for PTs: The PTs in the intervention clinics received 5 days of training from a UK instructor (GS) who had trained the PTs in the original studies of the STarT Back strategy. This training aimed to provide PTs with a better understanding of how psychosocial factors contribute to the pain experience and helped them apply pain-relevant psychosocial theories and practice to maximize their effectiveness in reducing their patients' pain-related disability.

Clinicians in the control clinics did not receive any special training and were completely unaffected by the intervention. They continued to provide whatever care they normally provided; this typically included ordering imaging studies, prescribing medications, and referrals to PT.

Outcomes

Telephone interviewers collected patient outcome data during the preintervention (November 2013-April 2014) and postintervention (December 2014-August 2016) periods. Interviews occurred 2 weeks (range: 1-3 weeks) after the LBP visit (baseline) and again 2 and 6 months later. Our primary outcomes were LBP-related physical function in the previous week (measured with the modified Roland-Morris Disability Questionnaire [RMDQ])[22] and LBP severity during the previous week (measured on a 0-to-10 scale in which 0 represents "no pain" and 10 "pain as bad as it could be").[23] Secondary outcomes included patient outcomes (depression, anxiety, fear of movement, global improvement, self-efficacy, satisfaction, and work productivity and activity impairment[20] and health care utilization from the EHR [e.g., lumbar imaging, PT, complementary and alternative medical therapies, cognitive behavioral therapy, opioid medications, epidural steroid injections, spine surgeon consultations]). Because the intervention targeted PCPs, it was not possible to identify adverse effects.

Data Collection

We collected patient data for measuring changes in patient outcomes 2 and 6 months after LBP visits in the intervention and control clinics during the preintervention and postintervention periods (Figure 1). Because collecting baseline data during the visit was not feasible, shortly after the visits we mailed patients letters explaining that GH was conducting a study to improve LBP care and that we would call to invite their participation. No mention was made that an intervention trial involving training clinicians in some Group Health Cooperative clinics was occurring. Patients not wishing to be contacted were provided a phone number to opt out. Research specialists called patients between 1 and 3 weeks after their visits to explain the study, answer questions, confirm eligibility, and obtain verbal informed consent to complete a baseline and 2 follow-up interviews. Patients were offered $20 for completing each questionnaire. Trained interviewers used computer-assisted telephone interviewing 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. Interviewers were blinded to patients' clinics. We used similar methods to collect postintervention data. Because we did not meet our recruitment goal during the preintervention period, we increased postintervention recruitment to maintain overall statistical power (see Sample Size section). We also improved the recruitment letter, increased staffing, and lengthened the recruitment period. The mean interval between visit date and baseline data collection was 12.7 (SD = 7.1) days and was similar in the intervention and control clinics.

Sample Size

We performed a priori sample size calculations targeting 80% power to detect a 1.5 point difference in 6-month LBP-related change in patient function (RMDQ) preintervention and postintervention between control and intervention clinics (0-point difference in the low-risk subgroups and 2.5-point difference in medium- and high-risk subgroups) and a 0.9-point difference in LBP pain severity score (0-point difference in low-risk subgroups and 1.5-point difference in medium- and high-risk subgroups).[24] We planned for a sample size of 1760 participants balanced equally between the preintervention and postintervention periods and

the control and intervention clinics, allowing for a loss to follow-up rate of 20%.[20] Because we recruited only 603 participants (goal was 880) during the preintervention period, we determined that we would need a sample size of 1334 during the postintervention period to maintain 80% power. The final numbers recruited were 603 participants in the preintervention period (546 with complete follow-up) and 1098 in the postintervention period (1008 with complete follow-up). Our post hoc calculation of power based on the observed data (accounting for imbalance between intervention arms) found we had 80% power to detect a difference between trial arms of 1.5 points on the change in RMDQ score before versus after the intervention. For simplicity, we determined the sample size assuming no correlation of outcomes within provider or clinic, yielding conservative estimates of sample size.

Statistical Methods

We first estimated the change in mean score by clinic assignment between the preintervention and postintervention periods (Figure 1). We then compared these differences to estimate the change attributable to the intervention (i.e., we made inferences on the interaction between clinic assignment and intervention period). We used a linear mixed effects model with random effects[25] for patient participants (repeated outcome measurements on participants at 2 and 6 months post–LBP visit) and clinic (randomization at clinic level) to account for correlation within individuals and clinics. This model accounted for clustering of outcomes by study. The primary analysis time point was 6 months following the LBP visit. To account for potential confounding variables, we adjusted for participant-level baseline covariates shown to be associated with LBP physical function and pain intensity, as well as variables that were imbalanced at baseline at the patient level between intervention and control arms: sex, age, education, race, employment, function (RMDQ), and pain intensity. We calculated risk subgroup–specific estimates and secondary outcomes using an identical framework to that described above, with 1 exception: for binary secondary outcomes we used generalized linear mixed models[26] with logit and/or log link functions to estimate odds ratios and/or relative risks instead of mean change scores. We assumed the standard alpha level of 0.05 for a 2-sided test.

We used the same analytic approach with EHR data to evaluate the effect of the intervention on health care utilization for LBP. We examined if the use of STarT Back recommended treatments for patients at medium- and high-risk of persistent disabling pain increased and the use of treatments not recommended for nonspecific LBP decreased. The primary analyses included all eligible patients (not just those providing patient data). We also analyzed data for the subset of patients who participated in the telephone questionnaires. Comparison of the data from these 2 populations allowed us to determine the representativeness of participants. We also examined the frequency with which STarT Back risk scores were recorded in the EHR by PCPs in the intervention clinics. More detail can be found elsewhere.[20]

RESULTS

Patient Recruitment and Follow-up

Figures 2A and 2B present flow diagrams showing the 6 clinics in this cluster randomized controlled trial and the flow of trial participants separately for the preintervention and postintervention periods. Because we included a preintervention "baseline" period (Figure 1), we present flow data separately for the preintervention and postintervention periods as well. A total of 2,138 LBP patients visited the intervention clinics and 2571 the control clinics. The characteristics of the intervention and control patients were very similar both preintervention and postintervention. Overall, 36% of patients provided baseline data on the telephone. Participating patients were slightly older than nonparticipants (mean ages of 57.1 and 54.8, respectively) and more likely to be white (83.0% and 77.0%, respectively). There were no differences by gender or Hispanic ethnicity. Follow-up rates among patients who agreed to participate were 93% at 2 months and 91% at 6 months. Participation and follow-up rates were similar in the intervention and control clinics (Figures 2A and 2B).

Patient Characteristics

Table 1 presents participants' characteristics based on baseline data from both the preintervention and postintervention periods. Reflecting the GH membership, participants had

relatively high levels of education and income and were primarily white and non-Hispanic. About half the participants were older than age 60; 56% reported that their current pain episode had lasted less than 3 months. Participants had moderately high levels of functional disability and pain severity, 48% reported leg pain, and about 30% were using opioids for their pain. Data from the STarT Back tool showed that 41% were categorized at low risk, 37% at medium risk, and 22% at high risk of persistent disabling pain. The STarT Back tool successfully distinguished among the prognoses of the 3 risk groups (i.e., the high-risk group had the worst outcomes and the low-risk group had the best outcomes).[27] Participants' characteristics were similar in the intervention and control arms.

Figure 2A. Flow of Patients Through Trial: Pre-Intervention Period

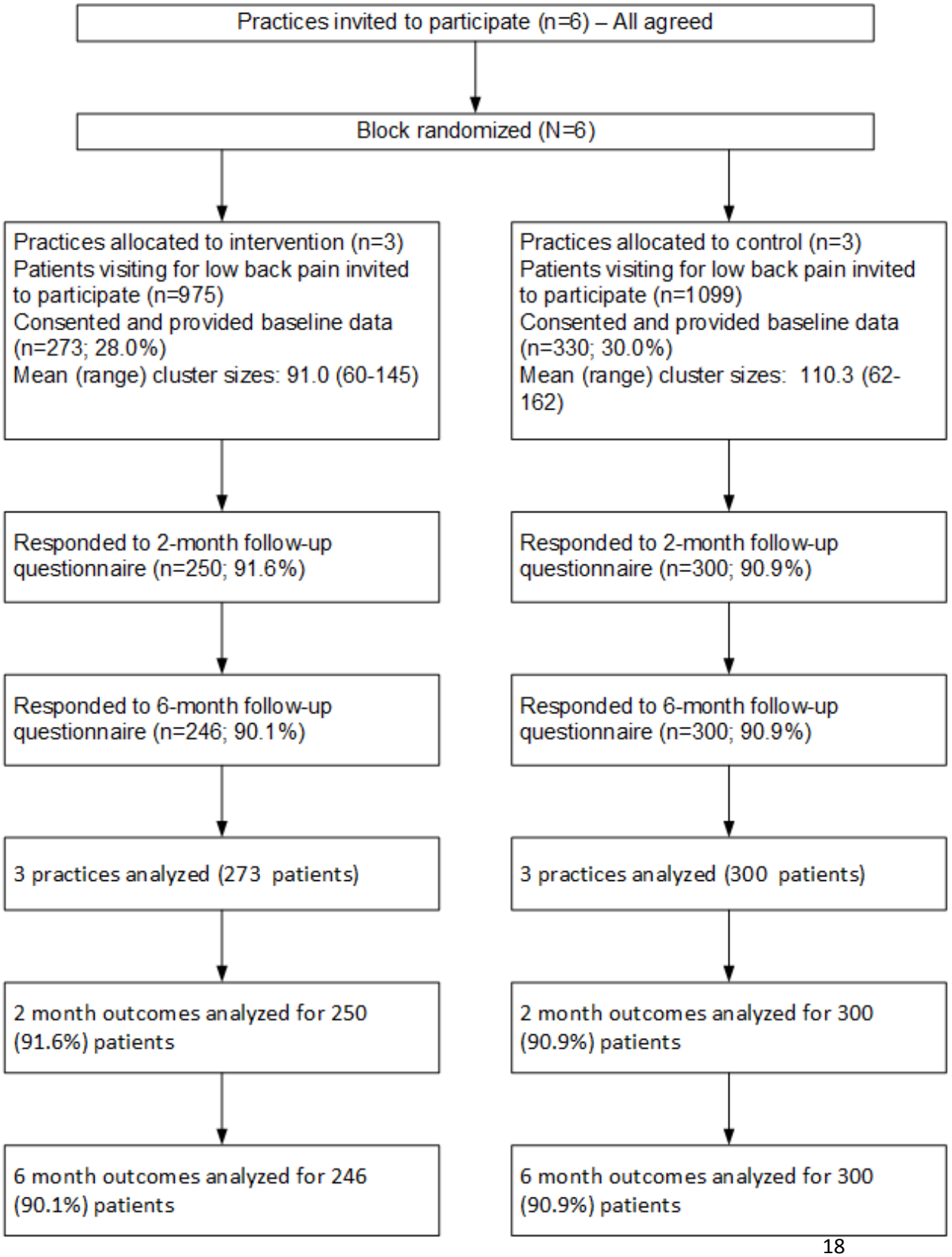


Figure 2B. Flow of Patients Through Trial: Post-Intervention Period

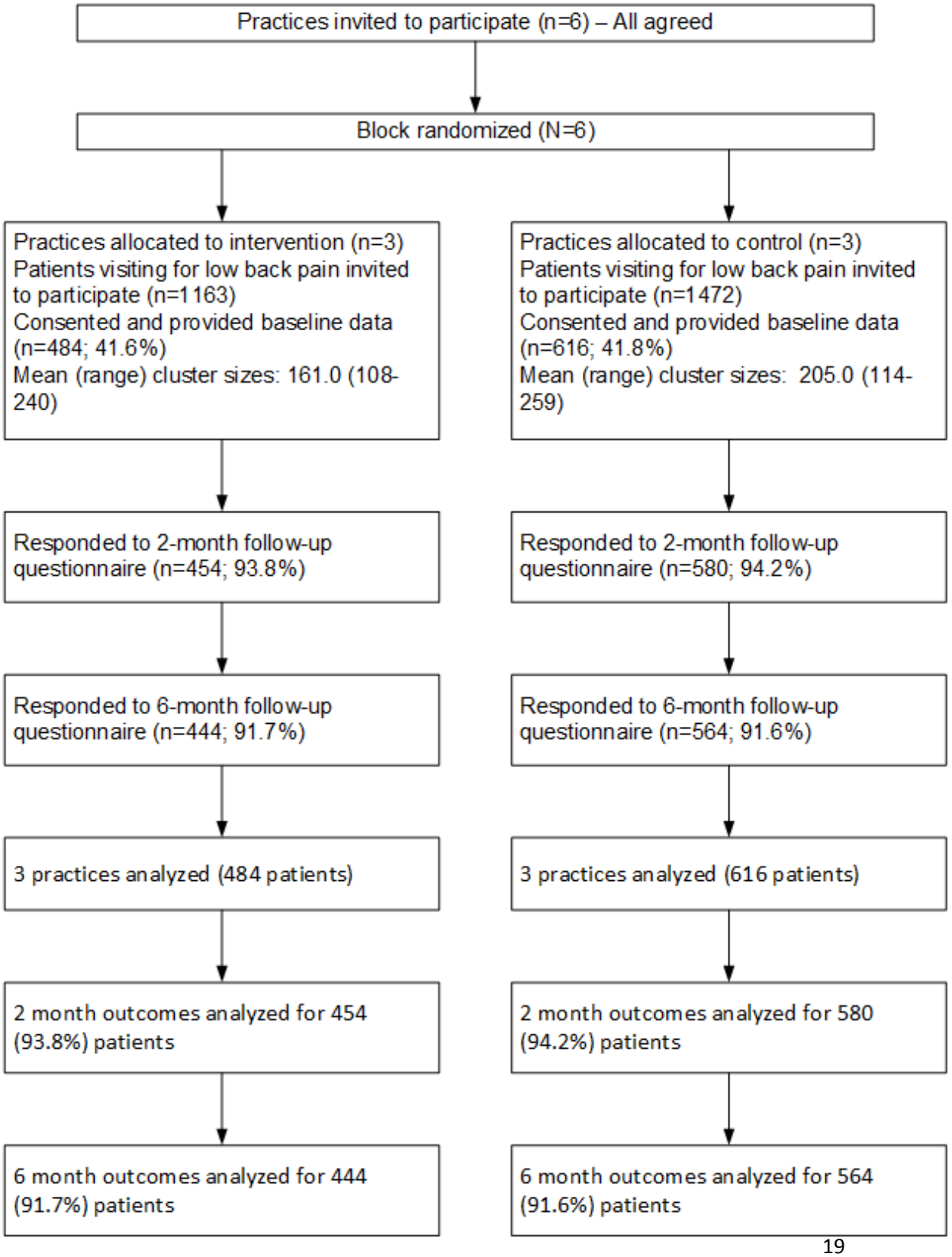


Table 1. Baseline Characteristics of Study Participants in the Control (n = 3) and Intervention (n = 3) Clinics^

Characteristic	Control (N = 945)		Intervention (N = 756)	
Sex, female, no. (%)	512	(54.2)	441	(58.3)
Age, mean (SD), y	55	(17.3)	58	(18.4)
18-39, no. (%)	215	(22.7)	160	(21.1)
40-59, no. (%)	310	(32.8)	204	(26.9)
60+, no. (%)	420	(44.4)	392	(51.8)
Education				
High school or less, no. (%)	143	(15.1)	105	(13.9)
Some college, no. (%)	303	(32.1)	231	(30.6)
College/postgraduate, no. (%)	498	(52.8)	418	(55.4)
Income*				
< $35K, no. (%)	161	(18.3)	141	(20.2)
$45-55K, no. (%)	203	(23.1)	186	(26.6)
$55-85K, no. (%)	216	(24.6)	149	(21.3)
$85K+, no. (%)	298	(33.9)	223	(31.9)
Employed, no. (%)	557	(58.9)	417	(55.2)
White, no. (%)*	743	(79.9)	582	(78.0)
Hispanic, no. (%)*	53	(5.8)	40	(5.5)
Back-related function (RMDQ), (0-23 scale) mean (SD)#	11.8	(6.3)	11.8	(6.1)
Back pain severity, (0-10 scale), mean (SD)##	5.4	(2.5)	5.5	(2.5)
StartBack risk group				
Low, no. (%)	392	(41.5)	305	(40.3)
Medium, no. (%)	348	(36.8)	286	(37.8)
High, no. (%)	205	(21.7)	165	(21.8)
Duration of current episode of LBP@				
< 3 months (%)	186	(56.4)	153	(56.0)
3-12 months (%)	58	(17.6)	47	(17.2)
> 12 months (%)	86	(26.0)	73	(26.7)
Hrs. of work missed past week due to LBP, mean (SD)	4.4	(10.4)	3.6	(8.9)
Effect of LBP on work in past week (0-10), mean (SD)	3.0	(2.9)	3.0	(2.7)
Leg pain in leg, no. (%)	457	(48.4)	360	(47.7)
Anxiety (GAD-7), mean (SD)	4.2	(4.6)	4.4	(4.6)
Depression (PHQ-8), mean (SD)	6.1	(5.4)	6.2	(5.3)
Self-efficacy (PSEQ), mean (SD)	44.4	(13.3)	45.0	(12.7)
Fear of movement (TSK), mean (SD)	39.7	(10.1)	39.4	(10.3)
Used medications for back pain in past week, no. (%)	735	(77.8)	570	(75.5)
Used narcotics for back pain in past week, no. (%)	104	(31.5)	76	(27.8)

* Missing data: Income and pain duration (7%), race and Hispanic (3%), hours of work missed (1%).

All other variables had < 1% missing.

Higher scores indicate greater dysfunction

Higher scores indicate greater pain severity

@ Pain duration was measured only during the preimplementation phase, resulting in smaller sample sizes than for the other measures.

^ Includes baseline data from the preintervention and postintervention periods combined.

Abbreviations: GAD = generalized anxiety disorder; LBP = low back pain; PHQ = patient health questionnaire PSEQ = patient self-efficacy questionnaire; TSK = Tampa Scale for Kinesiophobia; RMDQ = Roland-Morris Disability Questionnaire.

Effect of the Intervention

Patient Outcomes: At 6 months, no statistically significant differences existed between participants in the intervention and control arms for either the primary patient outcome overall or within risk subgroups (Table 2) or for secondary patient outcomes (Table 3). For example, Table 2 shows that improvement in mean functional disability after 6 months in the control group was 0.55 points greater on the disability scale during the postintervention period than the preintervention period (–3.89 versus –3.33 points), while the intervention group improved by only 0.05 points (–4.03 versus –3.98 points). The net effect was 0.50 greater improvement in the control group on the RMD scale (95% CI, –0.55-1.55; *P* = 0.349). This indicates that postimplementation improvements in the intervention clinics were slightly worse than those in the clinics not receiving the intervention. The absolute magnitudes of the between-group differences were generally small and for the primary outcome measures slightly favored the control group. We found similar results at 2-months follow-up (Appendix).

Health Care Utilization: The STarT Back tool data were entered into the EHR for about 50% of LBP visits during the 6-month intervention period, decreasing to about 40% over the ensuing 20 months. Among the 32 PCPs in the intervention clinics who saw at least 10 patients with LBP during both the intervention period and the postintervention period, the median percentage of visits with a STarT Back tool score in the EHR was 47% (range: 23%-71%) during the intervention period and 42% (range: 8%-71%) during the postintervention period. Thus, the tool continued to be used for patients of all intervention PCPs, at least occasionally, long after the intervention ended.

Despite PCP or nursing staff entry of the STarT Back tool data for almost half of the visits for LBP, knowledge of the patients' risk subgroup did not affect the type or frequency of health care provided (Table 4). Specifically, there was no evidence that the intervention increased the

use of treatments recommended for medium- and high-risk patients (e.g., PT, complementary and alternative medicine, cognitive behavioral therapy), or decreased the use of non-recommended tests or treatments (i.e., imaging, opioid medications, spine injections, surgical referrals) for LBP patients at any risk level. There was also no evidence of any changes in the tests or treatments recommended by clinicians in the intervention clinics for any of the patient risk subgroups (Table 5).

A prespecified secondary analysis restricted to patients providing telephone outcome data showed similar results.

Table 2. Primary Patient Outcomes Overall and by Risk Subgroup at 6-month Follow-up

	Preintervention Period			Postintervention Period				
	N	Change	95% CI	N	Change	95% CI	Difference	95% CI
Function (RMDQ)*								
Overall								
Control clinics	297	−3.33	(−4.18, −2.49)	546	−3.89	(−4.64, −3.14)	−0.55	(−1.26, 0.15)
Intervention clinics	245	−3.98	(−4.86, −3.09)	428	−4.03	(−4.81, −3.25)	−0.05	(−0.83, 0.73)
Difference		−0.64	(−1.86, 0.58)		−0.14	(−1.22, 0.94)	0.50	(−0.55, 1.55)
P value								0.349
Subgroup, Low Risk								
Control clinics	122	−5.18	(−5.98, −4.39)	234	−5.75	(−6.37, −5.12)	−0.56	(−1.42, 0.29)
Intervention clinics	109	−5.60	(−6.41, −4.78)	170	−5.60	(−6.29, −4.91)	−0.01	(−0.94, 0.93)
Difference		−0.41	(−1.48, 0.65)		0.14	(−0.70, 0.98)	0.56	(−0.71, 1.82)
P value								0.389
Subgroup, Moderate Risk								
Control clinics	108	−3.54	(−5.11, −1.97)	203	−3.90	(−5.30, −2.50)	−0.36	(−1.61, 0.89)
Intervention clinics	88	−3.34	(−4.98, −1.70)	165	−3.77	(−5.21, −2.32)	−0.43	(−1.82, 0.96)
Difference		0.20	(−2.06, 2.46)		0.13	(−1.86, 2.13)	−0.06	(−1.94, 1.81)
P value								0.946
Subgroup, High Risk								
Control clinics	67	−1.42	(−3.33, 0.49)	109	−2.49	(−4.15, −0.82)	−1.07	(−2.92, 0.79)
Intervention clinics	48	−3.78	(−5.82, −1.74)	93	−3.09	(−4.77, −1.41)	0.69	(−1.46, 2.84)
Difference		−2.36	(−4.77, 0.05)		−0.60	(−2.48, 1.27)	1.76	(−1.10, 4.62)
P value								0.229
Pain Intensity*								
Overall								
Control clinics	297	−1.64	(−1.94, −1.34)	534	−1.96	(−2.20, −1.72)	−0.32	(−0.66, 0.01)
Intervention clinics	245	−1.81	(−2.13, −1.48)	415	−2.00	(−2.26, −1.74)	−0.19	(−0.56, 0.18)
Difference		−0.17	(−0.61, 0.27)		−0.039	(−0.39, 0.31)	0.13	(−0.37, 0.63)
P value								0.61
Subgroup, Low Risk								
Control clinics	122	−2.13	(−2.54, −1.72)	228	−2.25	(−2.56, −1.94)	−0.12	(−0.59, 0.35)

Table 2. Primary Patient Outcomes Overall and by Risk Subgroup at 6-month Follow-up

	Preintervention Period			Postintervention Period				
	N	Change	95% CI	N	Change	95% CI	Difference	95% CI
Intervention clinics	109	−2.27	(−2.69, −1.85)	168	−2.42	(−2.78, −2.07)	−0.15	(−0.66, 0.36)
Difference		−0.14	(−0.69, 0.41)		−0.173	(−0.59, 0.25)	−0.03	(−0.72, 0.66)
P value								0.926
Subgroup, Moderate Risk								
Control clinics	108	−1.59	(−2.23, −0.95)	200	−1.93	(−2.48, −1.38)	−0.34	(−0.91, 0.23)
Intervention clinics	88	−1.47	(−2.14, −0.79)	157	−1.79	(−2.37, −1.21)	−0.33	(−0.97, 0.32)
Difference		0.13	(−0.80, 1.05)		0.14	(−0.65, 0.93)	0.02	(−0.85, 0.88)
P value								0.973
Subgroup, High Risk								
Control clinics	67	−0.74	(−1.55, 0.07)	106	−1.54	(−2.24, −0.84)	−0.8	(−1.62, 0.03)
Intervention clinics	48	−1.6	(−2.47, −0.73)	90	−1.54	(−2.25, −0.82)	0.06	(−0.89, 1.01)
Difference		−0.86	(−1.88, 0.17)		0.002	(−0.79, 0.79)	0.86	(−0.41, 2.12)
P value								0.183

***** Negative values indicate decreased dysfunction and pain severity (i.e., improvement).

Table 3. Secondary Patient Outcomes at 6-month Follow-up

	Preintervention			Postintervention				
	N	Change	95% CI	N	Change	95% CI	**Difference**	**95% CI**
Continuous Outcomes								
Depression (PHQ-8)								
Control clinics	296	−1.90	(−2.39, −1.41)	544	−1.41	(−1.80, −1.03)	0.49	(−0.06, 1.04)
Intervention clinics	245	−1.94	(−2.47, −1.41)	427	−1.57	(−1.99, −1.15)	0.36	(−0.24, 0.97)
Difference		−0.04	(−0.76, 0.68)		−0.16	(−0.73, 0.41)	−0.12	(−0.94, 0.70)
P value								0.770
Anxiety (GAD-7)								
Control clinics	297	1.13	(0.74, 1.51)	541	0.92	(0.64, 1.21)	−0.21	(−0.68, 0.27)
Intervention clinics	245	1.02	(0.60, 1.45)	425	0.83	(0.50, 1.15)	−0.20	(−0.73, 0.33)
Difference		−0.10	(−0.67, 0.47)		−0.10	(−0.53, 0.33)	0.01	(−0.71, 0.72)
P value								0.988
Fear of Movement (TSK)								
Control clinics	297	−4.19	(−5.19, −3.18)	537	−4.31	(−5.10, −3.52)	−0.12	(−1.27, 1.02)
Intervention Clinics	244	−4.53	(−5.63, −3.44)	422	−4.08	(−4.95, −3.21)	0.46	(−0.82, 1.73)
Difference		−0.35	(−1.83, 1.14)		0.23	(−0.95, 1.41)	0.58	(−1.13, 2.29)
P value								0.506
Self-efficacy (PSEQ)								
Control clinics	297	3.17	(1.94, 4.40)	543	3.96	(2.97, 4.94)	0.79	(−0.56, 2.13)
Intervention clinics	244	3.25	(1.92, 4.58)	423	4.31	(3.23, 5.39)	1.06	(−0.44, 2.56)
Difference		0.08	(−1.73, 1.89)		0.36	(−1.11, 1.81)	0.28	(−1.74, 2.29)
P value								0.789

Table 3. Secondary Patient Outcomes at 6-month Follow-up

	Preintervention			Postintervention				
	N	Change	95% CI	N	Change	95% CI	**Difference**	**95% CI**
Effect on Work Productivity								
Control clinics	155	–1.37	(–1.69, –1.05)	284	–1.44	(–1.69, –1.20)	–0.08	(–0.45, 0.30)
Intervention clinics	119	–1.68	(–2.04, –1.32)	209	–1.45	(–1.73, –1.17)	0.23	(–0.20, 0.67)
Difference		–0.32	(–0.80, 0.16)		–0.01	(–0.38, 0.36)	0.31	(–0.26, 0.88)
P value								0.288
Hours of Work Lost due to LBP								
Control clinics	155	–2.41	(–3.20, –1.61)	283	–2.94	(–3.53, –2.35)	–0.53	(–1.52, 0.46)
Intervention clinics	119	–3.33	(–4.24, –2.42)	209	–2.49	(–3.17, –1.81)	0.84	(–0.30, 1.98)
Difference		–0.92	(–2.13, 0.28)		0.45	(–0.45, 1.35)	1.37	(–0.13, 2.88)
P value								0.074

Binary Outcomes	N	Prop.	95% CI	N	Prop.	95% CI	OR	95% CI
Very Satisfied With Care								
Control clinics	284	0.50	(0.33, 0.67)	519	0.56	(0.42, 0.70)	1.27	(0.83, 1.94)
Intervention clinics	235	0.41	(0.25, 0.56)	420	0.41	(0.29, 0.52)	1.01	(0.63, 1.61)
OR		0.69	(0.42, 1.15)		0.55	(0.37, 0.80)	0.79	(0.42, 1.49)
P value								0.471
Very Satisfied With Treatment								
Control clinics	256	0.41	(0.26, 0.55)	415	0.38	(0.27, 0.49)	0.91	(0.57, 1.44)
Intervention clinics	209	0.26	(0.15, 0.37)	333	0.29	(0.19, 0.38)	1.15	(0.68, 1.95)
OR		0.51	(0.29, 0.89)		0.65	(0.42, 1.00)	1.26	(0.63, 2.54)
P value								0.516
Very Satisfied With Information About Cause of Pain								
Control clinics	289	0.31	(0.19, 0.43)	517	0.41	(0.29, 0.52)	1.52	(0.94, 2.46)
Intervention clinics	243	0.30	(0.17, 0.43)	416	0.36	(0.24, 0.48)	1.29	(0.76, 2.20)
OR		0.967	0.54 1.72		0.82	0.53 1.26	0.85	0.41 1.740
P value								0.65
Completely Recovered or Much Better (PGIC)								
Control clinics	297	0.26	(0.15, 0.36)	537	0.34	(0.24, 0.44)	1.51	(0.93, 2.46)
Intervention clinics	245	0.31	(0.17, 0.44)	423	0.38	(0.26, 0.51)	1.38	(0.81, 2.37)
OR		1.31	(0.73, 2.33)		1.19	(0.77, 1.84)	0.91	(0.44, 1.89)
P value								0.810

Table 4. Preintervention Versus Postintervention Proportion and Odds Ratio (OR) for Selected Health Services for Low Back Pain Between the Control and Intervention Arms in the 6 Months After Visit

	Preintervention Period			Postintervention Period				
	N	Prop	95% CI	N	Prop	95% CI	**OR**	**95% CI**
Lumbar Spine Imaging*								
Control clinics	1061	0.14	(0.07, 0.22)	1473	0.18	(0.09, 0.27)	1.34	(1.08, 1.67)
Intervention clinics	943	0.16	(0.08, 0.25)	1163	0.22	(0.11, 0.34)	1.46	(1.17, 1.84)
OR		1.20	(0.57, 2.54)		1.31	(0.63, 2.73)	1.09	(0.80, 1.50)
P value								0.578
Additional Primary Care Visits								
Control	1061	0.15	(0.12, 0.17)	1473	0.24	(0.21, 0.27)	1.86	(1.51, 2.29)
Intervention	943	0.12	(0.09, 0.14)	1163	0.24	(0.21, 0.28)	2.43	(1.91, 3.10)
OR		0.77	(0.59, 1.01)		1.01	(0.82, 1.23)	1.31	(0.95, 1.79)
P value								0.095
Emergency Department Visits								
Control clinics	1061	0.05	(0.04, 0.06)	1473	0.04	(0.03, 0.05)	0.77	(0.53, 1.11)
Intervention clinics	943	0.04	(0.03, 0.06)	1163	0.03	(0.02, 0.04)	0.75	(0.48, 1.18)
OR		0.81	(0.54, 1.23)		0.80	(0.53, 1.21)	0.98	(0.55, 1.76)
P value								0.959
Narcotic Analgesics								
Control Clinics	1061	0.39	(0.30, 0.48)	1473	0.45	(0.35, 0.55)	1.28	(1.07, 1.53)
Intervention Clinics	943	0.37	(0.28, 0.45)	1163	0.41	(0.32, 0.51)	1.23	(1.01, 1.49)
OR		0.91	(0.65, 1.27)		0.87	(0.64, 1.20)	0.96	(0.74, 1.25)
P value								0.757
Physical Therapy Visits								
Control clinics	1061	0.21	(0.15, 0.28)	1473	0.22	(0.15, 0.29)	1.04	(0.86, 1.26)
Intervention clinics	943	0.24	(0.16, 0.31)	1163	0.26	(0.18, 0.34)	1.13	(0.93, 1.38)
OR		1.13	(0.73, 1.76)		1.23	(0.81, 1.89)	1.09	(0.83, 1.43)
P value								0.546
CAM Visits								
Control clinics	1061	0.14	(0.12, 0.17)	1473	0.14	(0.12, 0.16)	0.94	(0.75, 1.18)
Intervention clinics	943	0.11	(0.09, 0.13)	1163	0.12	(0.10, 0.15)	1.12	(0.85, 1.46)
OR		0.74	(0.57, 0.97)		0.88	(0.70, 1.12)	1.19	(0.83, 1.70)
P value								0.338
Behavioral Health Visits								
Control clinics	1061	0.00	(0.00, 0.01)	1473	0.00	(0.00, 0.00)	0.76	(0.18, 3.15)
Intervention clinics	943	0.00	(0.00, 0.01)	1163	0.00	(0.00, 0.00)	0.45	(0.07, 2.77)
OR		1.13	(0.21, 6.08)		0.67	(0.10, 4.25)	0.59	(0.06, 5.96)

Table 4. Preintervention Versus Postintervention Proportion and Odds Ratio (OR) for Selected Health Services for Low Back Pain Between the Control and Intervention Arms in the 6 Months After Visit

	Preintervention Period			**Postintervention Period**				
	N	Prop	95% CI	N	Prop	95% CI	**OR**	**95% CI**
P value								0.655
Spine Surgeon Visits								
Control clinics	1061	0.02	(0.01, 0.03)	1473	0.03	(0.02, 0.03)	1.14	(0.69, 1.90)
Intervention clinics	943	0.02	(0.01, 0.03)	1163	0.03	(0.02, 0.04)	1.27	(0.75, 2.17)
OR		1.02	(0.58, 1.81)		1.14	(0.72, 1.80)	1.11	(0.53, 2.32)
P value								0.777
Injections of Lumbar Spine								
Control clinics	1061	0.01	(0.00, 0.02)	1473	0.01	(0.00, 0.01)	0.68	(0.32, 1.47)
Intervention clinics	943	0.01	(0.00, 0.01)	1163	0.01	(0.00, 0.01)	0.75	(0.28, 2.00)
OR		0.67	(0.26, 1.68)		0.73	(0.28, 1.92)	1.10	(0.32, 3.82)
P value								0.878
Back-related Hospitalizations								
Control clinics	1061	0.01	(0.00, 0.02)	1473	0.02	(0.01, 0.02)	1.60	(0.78, 3.29)
Intervention clinics	943	0.02	(0.01, 0.03)	1163	0.01	(0.01, 0.02)	0.71	(0.35, 1.43)
OR		1.80	(0.83, 3.89)		0.80	(0.41, 1.54)	0.44	(0.16, 1.21)
P value								0.112

** Includes plain films, CT scans, and MRIs.*

Table 5. Preintervention Versus Postintervention Odds Ratios for Selected Health Services for LBP Between the Control and Intervention Groups in the 6 Months After an Index Visit Overall and by Risk Subgroup, Entire Study Cohort (n = 1699)

	All Study Enrollees											
	Preintervention Period				Postintervention Period							
	N	Prop	95% CI		N	Prop	95% CI		OR	95% CI		
Any PT Visits												
Low Risk												
Control	134	0.28	(0.15,	0.40)	258	0.26	(0.16,	0.36)	0.93	(0.58,	1.48)	
Intervention	118	0.27	(0.14,	0.40)	187	0.31	(0.19,	0.44)	1.25	(0.75,	2.08)	
OR		0.97	(0.50,	1.86)		1.30	(0.76,	2.22)	1.34	(0.67,	2.68)	
P value												0.4
Medium Risk												
Control	120	0.25	(0.09,	0.41)	227	0.32	(0.15,	0.50)	1.45	(0.88,	2.39)	
Intervention	99	0.32	(0.11,	0.53)	187	0.34	(0.15,	0.53)	1.08	(0.64,	1.83)	
OR		1.42	(0.58,	3.49)		1.06	(0.48,	2.31)	0.74	(0.36,	1.53)	
P value												0.42
High Risk												
Control	76	0.27	(0.13,	0.41)	129	0.26	(0.16,	0.36)	0.95	(0.50,	1.82)	
Intervention	56	0.32	(0.14,	0.49)	108	0.36	(0.22,	0.50)	1.20	(0.61,	2.40)	
OR		1.24	(0.58,	2.67)		1.57	(0.90,	2.73)	1.26	(0.49,	3.25)	
P value												0.63
Any CAM Visits												
Low Risk												
Control	134	0.17	(0.08,	0.26)	258	0.13	(0.08,	0.18)	0.71	(0.39,	1.31)	
Intervention	118	0.11	(0.04,	0.17)	187	0.14	(0.08,	0.20)	1.38	(0.67,	2.84)	
OR		0.57	(0.27,	1.23)		1.11	(0.62,	2.01)	1.94	(0.75,	5.01)	
P value												0.17
Medium Risk												
Control	120	0.19	(0.10,	0.28)	227	0.17	(0.11,	0.23)	0.85	(0.47,	1.55)	
Intervention	99	0.13	(0.05,	0.21)	187	0.11	(0.06,	0.16)	0.82	(0.39,	1.76)	
OR		0.64	(0.30,	1.37)		0.61	(0.34,	1.09)	0.96	(0.37,	2.52)	
P value												0.94
High Risk												
Control	76	0.27	(0.13,	0.41)	129	0.26	(0.16,	0.36)	0.95	(0.50,	1.82)	
Intervention	56	0.32	(0.14,	0.49)	108	0.36	(0.22,	0.50)	1.20	(0.61,	2.40)	
OR		1.24	(0.58,	2.67)		1.57	(0.90,	2.73)	1.26	(0.49,	3.25)	
P value												0.63

Table 5. Preintervention Versus Postintervention Odds Ratios for Selected Health Services for LBP Between the Control and Intervention Groups in the 6 Months After an Index Visit Overall and by Risk Subgroup, Entire Study Cohort (n = 1699)

	All Study Enrollees											
	Preintervention Period				Postintervention Period							
	N	Prop	95% CI		N	Prop	95% CI		OR	95% CI		
Any LBP Imaging												
Low Risk												
Control	134	0.16	(0.00,	0.32)	258	0.20	(0.02,	0.39)	1.30	(0.70,	2.44)	
Intervention	118	0.25	(0.01,	0.49)	187	0.18	(0.01,	0.35)	0.65	(0.37,	1.15)	
OR		1.74	(0.44,	6.91)		0.86	(0.23,	3.20)	0.50	(0.21,	1.16)	
P value												0.11
Medium Risk												
Control	120	0.25	(0.07,	0.44)	227	0.16	(0.05,	0.28)	0.57	(0.33,	1.00)	
Intervention	99	0.25	(0.06,	0.44)	187	0.28	(0.09,	0.47)	1.19	(0.68,	2.08)	
OR		0.98	(0.34,	2.83)		2.03	(0.76,	5.40)	2.07	(0.94,	4.57)	
P value												0.07
High Risk												
Control	76	0.34	(0.17,	0.52)	129	0.34	(0.20,	0.47)	0.96	(0.53,	1.77)	
Intervention	56	0.29	(0.11,	0.47)	108	0.30	(0.17,	0.44)	1.06	(0.52,	2.17)	
OR		0.79	(0.36,	1.76)		0.87	(0.47,	1.59)	1.10	(0.43,	2.80)	
P value												0.85
Any Narcotic Rx												
Low Risk												
Control	134	0.28	(0.16,	0.41)	258	0.23	(0.16,	0.31)	0.76	(0.46,	1.27)	
Intervention	118	0.34	(0.19,	0.48)	187	0.24	(0.15,	0.33)	0.62	(0.36,	1.08)	
OR		1.29	(0.71,	2.33)		1.06	(0.65,	1.73)	0.82	(0.39,	1.73)	
P value												0.61
Medium Risk												
Control	120	0.56	(0.34,	0.78)	227	0.47	(0.34,	0.61)	0.70	(0.43,	1.13)	
Intervention	99	0.48	(0.27,	0.68)	187	0.48	(0.33,	0.63)	1.00	(0.59,	1.70)	
OR		0.71	(0.40,	1.27)		1.02	(0.67,	1.56)	1.43	(0.70,	2.92)	
P value												0.33
High Risk												
Control	76	0.72	(0.33,	1.11)	129	0.67	(0.39,	0.95)	0.81	(0.42,	1.55)	
Intervention	56	0.66	(0.26,	1.05)	108	0.61	(0.33,	0.88)	0.81	(0.40,	1.64)	
OR		0.74	(0.33,	1.68)		0.74	(0.40,	1.39)	1.00	(0.38,	2.61)	
P value												1.00

DISCUSSION

Context for Study Results

The MATCH trial is the first major evaluation of the implementation of an adaption of the STarT Back risk stratification strategy in the United States. Although the intervention resulted in use of the STarT Back tool for approximately half of patient visits for LBP, it did not change PCP practice decisions. Another recent cluster randomized controlled trial evaluated use of a multifaceted strategy (including embedding the STarT Back Tool in the EHR) to implement LBP guidelines into Danish general practices.[28] That trial found lower secondary care referral rates in the intervention clinics (5.0%) than in the control clinics (10.5%), but no improvement in patient outcomes.

Complex interventions such as the one evaluated in this trial could fail to improve patient outcomes for many reasons, including unacceptability to clinicians, inadequate leadership and system support, ineffective implementation, inadequate potency, and lack of room for improvement in already high-functioning organizations. Although a comprehensive evaluation of the implementation process found high levels of clinician engagement and system support (Hsu C, submitted for publication), there were limitations in our intervention that could explain why PCP behavior did not change, most notably (1) we did not conduct feedback audits to encourage clinician adherence to matching treatments to patient subgroups, and (2) compared with the study in the United Kingdom, general practices that focused on PT interventions,[9] our matched treatment options were more numerous, less familiar, and more difficult to access, thereby placing a greater burden on our PCPs. We also used a different recruitment strategy than that of the UK study. Differences between the UK study population[9] and our study population could also explain outcome differences. For example, although the patient populations were similar in age, gender, employment, risk subgroup distribution, and pain severity, US patients had substantially higher baseline levels of LBP-related physical disability (RMDQ scores of 11.8 versus 8.4, respectively).

Generalizability of the Findings and Implementation of Study Results

We designed our intervention[20] to be as potent as possible without making it impossible to implement in primary care clinics; however, even if our intervention had improved outcomes, we now recognize that it may not have been feasible to implement in most US primary care settings. The high levels of burnout among PCPs and the continued turmoil in US health care[29] make complex changes in clinical practice difficult.

Subpopulation Considerations

Although our overall analyses of the effect of the intervention on patient outcomes clearly showed no benefit (and in fact slightly favored the control group), we conducted exploratory analyses of the effects on the 3 STarT Back risk groups (ie, low, medium, and high risk of a poor outcome). These analyses failed to provide evidence that the intervention differentially affected patients in any of the risk subgroups.

Study Limitations

Major strengths of the MATCH trial include randomization of matched pairs of clinics to serve as intervention or control clinics; adequate sample sizes and power to detect meaningful differences; system allotment of time for clinician training; high follow-up rates; and an adaptive and pragmatic intervention design including substantial PT training, training modules based on requests of primary care teams, and inclusion of the whole primary care team.[20] Limitations include low participation rates of LBP patients, lack of clarity about who collected the STaRT Back data and entered it into the EHR (i.e., PCP versus nursing staff), collection of data for risk stratification for only half of LBP patients, and the need to defer baseline data collection until 2 weeks after the PCP visit.

CONCLUSIONS

In contrast to the positive results of implementing a risk stratification strategy to improve primary care for LBP in the United Kingdom,[15,16] our adaptation of that strategy to the different circumstances in our setting did not improve health care utilization or patient outcomes. This

illustrates the risk of failure when complex interventions developed and found effective in one setting are implemented in a different setting—even when the intervention was adapted to local needs and circumstances, clinicians were given time away from clinical responsibilities to participate, and most clinicians viewed the experience as worthwhile. Future initiatives to implement complex interventions in primary care that include simple and easily implemented and supported treatment recommendations, automatic alerts in the EHR to make it easy for clinicians to remember to collect risk stratification information, and the provision of regular feedback on their performance adhering to the matched treatment recommendations for patients at each risk stratum are likely to improve the chances of success.

References

1. Martin BI, Deyo RA, Mirza SK, et al. Expenditures and health status among adults with back and neck problems. *JAMA*. 2008;(299):656-664.
2. Von Korff M, Kolodny A, Deyo RA, Chou R. Long-term opioid therapy reconsidered. *Ann Intern Med*. 2011;(155):325-328.
3. Von Korff M, Franklin G. *Responding to America's Iatrogenic Epidemic of Prescription Opioid Addiction and Overdose. Med Care.* 2016;54(5):426-9.
4. Moore JE. Chronic low back pain and psychosocial issues. *Phys Med Rehabil Clin N Am*. 2010;21(4):801-815.
5. Boothby JL, Thorn BE, Stroud MW, Jensen MP. Coping with pain. In: Gatchel RJ, Turk DC, eds. *Psychosocial Factors in Pain: Critical Perspectives*. New York, NY: Guilford Press; 1999:343-359.
6. Jensen MP, Karoly P. Control beliefs, coping efforts, and adjustment to chronic pain. *J Consult Clin Psychol*. 1991;(59):431-438.
7. Hill JC, Dunn KM, Lewis M, et al. A primary care back pain screening tool: identifying patient subgroups for initial treatment. *Arthritis Rheum*. 2008;(59):632-641.
8. Hill JC, Whitehurst DG, Lewis M, et al. Comparison of stratified primary care management for low back pain with current best practice (STarT Back): a randomised controlled trial. *Lancet*. 2011;(378):1560-1571.
9. Foster NE, Mullis R, Hill JC, et al. Effect of stratified care for low back pain in family practice (IMPaCT Back): a prospective population-based sequential comparison. *Ann Fam Med*. 2014;(12):102-111.
10. 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(4):405-414.
11. Hay EM, Dunn KM, Hill JC, et al. 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.
12. 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):783-787.
13. Sowden G, Hill J, Konstantinou K, et al. 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.
14. 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):110-117.
15. 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)*. 2016;41(8):645-52.

16. 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):1712-1721.
17. Karstens S, Krug K, Hill JC, 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.
18. Karstens S, Joos S, Hill JC, Krug K, Szecsenyi J, Steinhauser 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. DOI: 10.1371/journal.pone.0136119
19. 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. DOI: 10.1371/journal.pone.0152019
20. Cherkin D, Balderson B, Brewer G, et al. Evaluation of a risk-stratification strategy to improve primary care for back pain: the MATCH cluster randomized trial protocol. *BMC Musculoskelet Disord*. 2016;(17):361.
21. Hemming K, Haines TP, Chilton AJ, et al. The stepped wedge cluster randomized trial: rationale, design, analysis, and reporting. *BMJ.* 2015;(351):h391.
22. Bombardier C, Hayden J, Beaton DE. Minimal clinically important difference. low back pain: outcome measures. *J Rheumatol*. 2001;28(2):431-438.
23. Chiarotto A, Deyo RA, Terwee CB, et al. Core outcome domains for clinical trials in nonspecific low back pain. *Eur Spine J*. 2015;24(6):1127-1142.
24. Patrick D, Deyo RA, Atlas SJ, Singer DE, Chapin A, Keller RB. Assessing health-related quality of life in ptients with sciatica. *Spine (Phila Pa 1976).* 1995;(20):1899-1908.
25. Laird NM, Ware JH. Random-effects models for longitudinal data. *Biometrics*. 1982;(38):963-974.
26. Pinheiro J, Bates D,. Approximations to the log-likelihood function in the nonlinear mixed-effects model. *J Comp Graph Stat*. 1995;(4):12-35.
27. Suri P, Delaney K, Rundell SD, Cherkin D. Predictive validity of the STarT Back tool for risk stratification in a United States primary care setting. *Arch Phys Med Rehabil*. In press.
28. Riis A, Jensen CE, Bro F, et al. A multifaceted implementation strategy versus passive implementation of low back pain guidelines in general practice: a cluster randomised controlled trial. *Implement Sci*. 2016;11(1):143.
29. Shanafelt TD, Dyrbye LN, West CP. Addressing physician burnout: the way forward. *JAMA*. 2017;317(9):901-902.

Key Words: low back pain, risk stratification, STarT Back, primary care, patient outcomes

Funding Support: Funding for this trial was provided by the Patient-Centered Outcomes 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/National Institutes of Health ("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 E Foster, an NIHR senior investigator, and Jonathan C. Hill were supported through an NIHR Research Professorship (NIHR-RP-011-015) awarded to Nadine Foster.

Disclaimer: The views and opinions expressed in this paper are those of the authors and do not necessarily reflect those of the Patient-Centered Outcomes Research Institute (PCORI), the National Institutes of Health, NIHR, NHS, or the Department of Health. PCORI reviewed the draft final report and requested changes in the reporting.

Prior Presentation: Plenary presentation ("Evaluation of Implementation of a Risk Stratification Strategy to Improve Primary Care for Low Back Pain") at the International Back and Neck Pain Research Forum in Oslo, Norway, on September 14, 2017

Acknowledgments: We are very grateful to the primary care clinicians, physical therapists, and staff at the Group Health Lynnwood, Northgate, and Renton clinics for their support for and engagement in the intervention. In addition, we would like to thank the following members of the research team: Zoe Bermet; Kristin Delaney, MPH; John Ewing; Kevin Filocamo; Vina Graham; Margie Wilcox; survey interviewers; Jennifer Brandlin; Donna Luce; and Patricia Scott.

Appendix

Appendix. Patient Outcomes at 2-month Follow-up

	Preintervention Period			Post\intervention Period				
	N	Change	95% CI	N	Change	95% CI	Difference	95% CI
PRIMARY OUTCOMES								
Function (RMDQ)								
Main Analysis								
Control clinics	297	–2.67	(–3.51, –1.82)	563	–3.08	(–3.82, –2.33)	–0.41	(–1.11, 0.29)
Intervention clinics	249	–2.41	(–3.29, –1.52)	437	–3.15	(–3.93, –2.38)	–0.75	(–1.52, 0.03)
Difference		0.26	(–0.96, 1.48)		–0.08	(–1.15, 1.00)	–0.34	(–1.38, 0.71)
P value								0.530
Subgroup, Low Risk								
Control clinics	119	–5.19	(–5.99, –4.39)	237	–4.96	(–5.58, –4.33)	0.24	(–0.62, 1.10)
Intervention clinics	112	–4.17	(–4.98, –3.37)	174	–5.12	(–5.81, –4.43)	–0.95	(–1.88, –0.02)
Difference		1.02	(–0.04, 2.08)		–0.16	(–1.00, 0.67)	–1.18	(–2.44, 0.08)
P value								0.066
Subgroup, Moderate Risk								
Control clinics	110	–1.85	(–3.41, –0.28)	210	–2.97	(–4.37, –1.58)	–1.12	(–2.37, 0.12)
Intervention clinics	89	–1.92	(–3.56, –0.27)	168	–2.87	(–4.31, –1.43)	–0.95	(–2.34, 0.44)
Difference		–0.07	(–2.32, 2.19)		0.10	(–1.88, 2.09)	0.17	(–1.70, 2.04)
P value								0.858
Subgroup, High Risk								
Control clinics	68	–1.21	(–3.11, 0.69)	116	–1.85	(–3.49, –0.20)	–0.64	(–2.47, 1.20)
Intervention clinics	48	–1.56	(–3.60, 0.49)	95	–1.62	(–3.31, 0.06)	–0.07	(–2.21, 2.08)
Difference		–0.35	(–2.75, 2.05)		0.23	(–1.63, 2.08)	0.57	(–2.27, 3.42)
P value								0.693
Pain Intensity								
Main Analysis								
Control clinics	297	–1.48	(–1.78, –1.18)	558	–1.56	(–1.80, –1.33)	–0.08	(–0.41, 0.25)
Intervention clinics	249	–1.31	(–1.63, –0.99)	430	–1.82	(–2.08, –1.56)	–0.51	(–0.88, –0.14)
Difference		0.17	(–0.27, 0.61)		–0.26	(–0.61, 0.09)	–0.43	(–0.93, 0.07)
P value								0.09
Subgroup, Low Risk								
Control clinics	119	–2.22	(–2.64, –1.81)	234	–2.04	(–2.35, –1.72)	0.19	(–0.28, 0.65)

Intervention clinics	112	–1.78 (–2.20, –1.37)	172	–2.24 (–2.59, –1.89)	–0.45 (–0.96, 0.05)	
Difference		0.44 (–0.11, 0.98)		–0.20 (–0.62, 0.22)	–0.64 (–1.32, 0.05)	
P value					0.069	
Subgroup, Moderate Risk						
Control clinics	110	–1.12 (–1.75, –0.48)	209	–1.49 (–2.03, –0.94)	–0.37 (–0.94, 0.20)	
Intervention clinics	89	–1.1 (–1.77, –0.43)	164	–1.79 (–2.37, –1.22)	–0.7 (–1.33, –0.06)	
Difference		0.02 (–0.90, 0.94)		–0.31 (–1.09, 0.48)	–0.33 (–1.18, 0.53)	
P value					0.452	
Control clinics	68	–0.66 (–1.47, 0.14)	115	–0.83 (–1.52, –0.14)	–0.17 (–0.98, 0.64)	
Intervention clinics	48	–0.82 (–1.70, 0.05)	94	–1.08 (–1.79, –0.37)	–0.25 (–1.20, 0.69)	
Difference		–0.16 (–1.19, 0.86)		–0.25 (–1.02, 0.53)	–0.08 (–1.34, 1.17)	
P value					0.895	
SECONDARY OUTCOMES						
Continuous Outcomes						
Depression (PHQ-8)						
Control clinics	296	–1.42 (–1.91, –0.93)	557	–1.34 (–1.72, –0.96)	0.08 (–0.47, 0.63)	
Intervention clinics	249	–1.28 (–1.80, –0.75)	437	–1.42 (–1.84, –1.00)	–0.14 (–0.75, 0.46)	
Difference		0.14 (–0.58, 0.86)		–0.08 (–0.65, 0.49)	–0.22 (–1.04, 0.59)	
P value					0.592	
Anxiety (GAD-7)						
Control clinics	297	0.76 (0.38, 1.14)	557	0.88 (0.60, 1.16)	0.12 (–0.36, 0.60)	
Intervention clinics	249	0.87 (0.45, 1.29)	436	0.71 (0.39, 1.03)	–0.16 (–0.68, 0.37)	
Difference		0.11 (–0.46, 0.68)		–0.17 (–0.59, 0.26)	–0.28 (–0.99, 0.44)	
P value					0.448	
Fear of Movement (TSK)						
Control clinics	297	–2.92 (–3.92, –1.91)	555	–2.8 (–3.59, –2.02)	0.11 (–1.02, 1.25)	
Intervention clinics	248	–2.51 (–3.59, –1.42)	434	–3.33 (–4.19, –2.47)	–0.82 (–2.09, 0.44)	
Difference		0.41 (–1.07, 1.89)		–0.53 (–1.70, 0.64)	–0.94 (–2.64, 0.76)	
P value					0.28	
Self-efficacy (PSEQ)						
Control clinics	296	2.53 (1.30, 3.76)	561	2.87 (1.89, 3.84)	0.34 (–1.00, 1.68)	
Intervention clinics	249	2.55 (1.23, 3.87)	434	2.87 (1.80, 3.94)	0.32 (–1.17, 1.80)	
Difference		0.02 (–1.78, 1.83)		0 (–1.45, 1.45)	–0.02 (–2.03, 1.98)	
P value					0.982	
Effect on Work Productivity (0-10)						
Control clinics	161	–1.06 (–1.37, –0.74)	303	–1.19 (–1.42, –0.95)	–0.13 (–0.50, 0.24)	

Intervention clinics	123	–1.23	(–1.58,	–0.87)	226	–1.16	(–1.43,	–0.89)	0.07	(–0.36,	0.49)
Difference		–0.17	(–0.64,	0.31)		0.03	(–0.33,	0.39)	0.19	(–0.37,	0.76)
P value											0.499
Hours of Work Lost due to Back Pain											
Control clinics	162	–2.15	(–2.93,	–1.37)	299	–2.65	(–3.23,	–2.08)	–0.5	(–1.47,	0.46)
Intervention clinics	123	–3.4	(–4.29,	–2.50)	227	–2.69	(–3.35,	–2.03)	0.71	(–0.40,	1.83)
Difference		–1.25	(–2.43,	–0.06)		–0.03	(–0.91,	0.85)	1.22	(–0.26,	2.69)
P value											0.107
Binary Outcomes											
Completely Recovered or Much Better (PGIC)											
Control clinics	296	0.27	(0.16,	0.38)	554	0.3	(0.21,	0.38)	1.14	(0.70,	1.84)
Intervention	249	0.21	(0.12,	0.31)	432	0.33	(0.22,	0.44)	1.85	(1.07,	3.19)
OR		0.72	(0.40,	1.30)		1.18	(0.76,	1.81)	1.63	(0.79,	3.37)
P value											0.189
Very Satisfied With Care											
Control clinics	285	0.47	(0.31,	0.63)	531	0.54	(0.41,	0.68)	1.35	(0.88,	2.05)
Intervention	240	0.36	(0.22,	0.50)	425	0.48	(0.34,	0.61)	1.61	(1.00,	2.58)
OR		0.65	(0.39,	1.07)		0.77	(0.53,	1.12)	1.19	(0.63,	2.24)
P value											0.583
Very Satisfied With Treatment											
Control clinics	243	0.35	(0.22,	0.48)	434	0.3	(0.21,	0.39)	0.81	(0.50,	1.29)
Intervention	205	0.28	(0.16,	0.40)	344	0.31	(0.21,	0.41)	1.13	(0.67,	1.91)
OR		0.73	(0.42,	1.28)		1.03	(0.67,	1.58)	1.4	(0.69,	2.84)
P value											0.344
Very Satisfied With Information About Cause											
Control clinics	288	0.31	(0.19,	0.43)	530	0.38	(0.27,	0.49)	1.38	(0.85,	2.23)
Intervention	247	0.29	(0.16,	0.41)	427	0.34	(0.23,	0.45)	1.3	(0.76,	2.20)
OR		0.894	0.5	1.59		0.84	0.55	1.29	0.94	0.46	1.92
P value											0.864

Disclaimer:

The [views, statements, opinions] presented in this report are solely the responsibility of the author(s) and do not necessarily represent the views of the Patient-Centered Outcomes Research Institute® (PCORI®), its Board of Governors or Methodology Committee.

Acknowledgement:

Research reported in this report was [partially] funded through a Patient-Centered Outcomes Research Institute® (PCORI®) Award (398). Further information available at: https://www.pcori.org/research-results/2012/can-using-patient-reports-low-back-pain-help-better-direct-patients-treatments