



HHS Public Access

Author manuscript

Arch Phys Med Rehabil. Author manuscript; available in PMC 2019 August 01.

Published in final edited form as:

Arch Phys Med Rehabil. 2018 August ; 99(8): 1533–1539.e2. doi:10.1016/j.apmr.2018.02.016.

Predictive Validity of the STarT Back Tool for Risk of Persistent Disabling Back Pain in a United States Primary Care Setting

Pradeep Suri, MD, MS^{1,2,3}, Kristin Delaney, MPH⁴, Sean D. Rundell, DPT, PhD^{3,5}, and Daniel C. Cherkin, PhD⁴

¹Seattle Epidemiologic Research and Information Center (ERIC), Department of Veterans Affairs Office of Research and Development, Seattle, Washington, USA

²Division of Rehabilitation Care Services, VA Puget Sound Health Care System, Seattle, Washington, USA

³Department of Rehabilitation Medicine, University of Washington, Seattle, Washington, USA

⁴Kaiser Permanente Washington Health Research Institute, Kaiser Permanente, Seattle, Washington, USA

⁵Department of Health Services, University of Washington, Seattle, Washington, USA

Abstract

Objective—To examine the predictive validity of the STarT Back tool for classifying people with back pain into categories of low, medium, and high risk of persistent disabling back pain, in US primary care.

Design—Secondary analysis of data from participants receiving usual care in a randomized clinical trial.

Setting—Primary care clinics.

Participants—1109 adults with back pain 18 years of age. Those with specific causes of back pain (pregnancy, disc herniation, vertebral fracture, spinal stenosis) and work-related injuries were not included.

Interventions—N/A

Main Outcome Measures—The original 9-item version of the STarT Back tool, administered at baseline, stratified patients by their risk (low, medium, high) of persistent disabling back pain

Corresponding author: Pradeep Suri, MD, MS, Division of Rehabilitation Care Services/Seattle Epidemiologic Research and Information Center (ERIC), VA Puget Sound Healthcare System, 1660 S. Columbian Way, Seattle, WA 98108; tel: 206-764-2673; fax: 206-716-5977; pradeep.suri@va.gov.

Publisher's Disclaimer: This is a PDF file of an unedited manuscript that has been accepted for publication. As a service to our customers we are providing this early version of the manuscript. The manuscript will undergo copyediting, typesetting, and review of the resulting proof before it is published in its final citable form. Please note that during the production process errors may be discovered which could affect the content, and all legal disclaimers that apply to the journal pertain.

Reprints are not available

Trial Registration: National Clinical Trial Number NCT02286141.

Prior Presentations: None

The authors report no conflicts of interest.

(‘STarT Back risk group’). Persistent disabling back pain was defined as Roland-Morris Disability Questionnaire scores of ≥ 7 at 6-month follow-up.

Results—STarT Back risk group was a significant predictor of persistent disabling back pain ($p < 0.0001$) at 6-month follow-up. The proportion of individuals with persistent disabling back pain at follow-up was 22% (95% confidence interval [CI] 18–25%) in the low risk group, 62% (95% CI 57–67%) in the medium-risk group, and 80% (95% CI 75–85%) in the high-risk group. The relative risk of persistent disabling back pain was 2.9 (95% CI 2.4–3.5) in the medium-risk group as compared to the low-risk group, and 3.7 (95% CI 3.1–4.4) in the high-risk group.

Conclusions—We found that the STarT Back risk groups successfully separated people with back pain into distinct categories of risk for persistent disabling back pain at 6-month follow-up in US primary care. These results were very similar to those seen in the original STarT Back validation study. This validation study is a necessary first step towards identifying whether the entire STarT Back approach, including matched/targeted treatment, can be effectively used for primary care in the US.

Keywords

STarT Back; risk stratification; prognostic; prediction; musculoskeletal; outcomes; psychosocial; low back pain

Back pain causes more disability than any other health condition worldwide¹ and is responsible for estimated costs of \$88 billion annually in the United States (US).² A promising new approach to mitigating back-related disability and health care expenditures is the Keele University “Subgrouping for Targeted Treatment” (STarT Back) approach.³ Developed in primary care settings in the United Kingdom (UK), the STarT Back approach consists of 2 interrelated components. First, it uses a 9-item “STarT Back tool” to categorize patients into 3 subgroups according to their predicted risk of persistent back-related functional limitations (low risk vs. medium risk vs. high risk).⁴ Second, the approach identifies treatments thought to be most appropriate for patients in each subgroup. The matched treatments that comprise the STarT Back approach as originally developed in the UK involve a brief educational session without further treatment for those at low risk, physical therapy for those at medium risk, and psychologically informed physical therapy (incorporating principles of cognitive-behavioral therapy) for those at high risk.^{3, 5} The STarT Back approach was found to be effective for people with back pain with or without radiculopathy in primary care in the United Kingdom (UK), where it improved back-related functional limitations while reducing the costs of healthcare,³ and has subsequently been replicated in the UK and Ireland.^{5, 6} The STarT Back approach is considered broadly applicable to those with the symptom of back pain, except those with red flag conditions, irrespective of possible underlying subtypes or putative causes of back pain.³

The effectiveness of the STarT Back approach relies first and foremost on the degree to which the STarT Back tool is valid for predicting persistent back-related functional limitations. Given that the tool was developed in UK primary care, its psychometric properties require evaluation to determine whether they are acceptable in other countries with different languages, health care systems, and/or sociocultural norms regarding back

pain. Various studies have examined the reliability, construct validity, and content validity of the STarT Back tool, however, few studies have examined the tool's external predictive validity, which is central to the question of whether the STarT Back approach can be exported to different clinical settings. Those studies that have examined the external predictive validity of the STarT Back tool in European primary care have generally found the tool to be successful in risk stratification. However, certain subgroups of patients have been identified in which the STarT Back tool may perform less optimally, including older adults,⁴ those with acute back pain of duration <1 month or 2 weeks^{7, 8}, and those with early trajectories of major improvement in back-related symptoms⁹. The STarT Back tool has also not performed as well in secondary settings such as physical therapy or chiropractic clinics.⁹⁻¹¹ To our knowledge, no prior study has examined the external predictive validity of the original 9-item STarT Back tool among people with back pain in US primary care.

The study aim was to examine the external predictive validity of the STarT Back tool when using the recommended cutoffs for classification into low-, medium-, and high-risk categories, among unselected people with back pain in US primary care.⁴ For the purposes of this article, these 3 categories are referred to as the 'STarT Back risk groups'. A secondary aim was to examine external predictive validity of the STarT Back risk groups among 3 subgroups where the tool's performance may be suboptimal:^{4, 9, 12, 13} older adults 65 years, those with acute back pain of duration <1 month, and those with a self-reported trajectory of recent major improvement in back pain-related symptoms.

METHODS

Study Design and Setting

We conducted a secondary analysis of data from the MATCH (Matching Appropriate Treatments to Consumers Healthcare needs) trial, a pragmatic cluster randomized trial involving people with back pain from primary care in the Group Health (now Kaiser Permanente Washington), an integrated healthcare system. The trial design and methods have been reported in detail elsewhere.¹⁴ Briefly, six primary care clinics were randomized in 1:1 ratio to an intervention or control. Intervention and control clinics were matched on geographic and socioeconomic characteristics. A pre-post design was used such that all clinics and participants in both intervention and control arms received usual care during the 1st phase of the study, the "pre-implementation period", which lasted 5 months (Figure 1). The 2nd phase of the study was a 7-month "implementation period" during which no participants were recruited and intervention clinics began to 1) incorporate the STarT Back tool into the electronic health record (EHR), 2) identify recommended treatment options for patients in each risk group, and 3) train providers in the STarT Back approach. The 3rd phase of the study was an 8-month 'post-implementation' period during which intervention clinics used the STarT Back tool in the EHR and identified STarT Back-recommended treatment option. Control clinics continued to deliver usual care during all 3 study phases. This secondary analysis was restricted to MATCH participants who received usual care after their index back pain visit, including all participants recruited during the pre-implementation period, and participants recruited from control clinics during the post-implementation period

(Figure 1). The MATCH trial was approved by the Group Health Institutional Review Board. Further details of study methods are provided in the online supplement.

Study Participants and Data Collection

We used the EHR to identify patients 18 years of age with an index clinical encounter for non-specific back pain identified using ICD-9 diagnosis codes. We excluded patients with specific causes of their pain (e.g., pregnancy, disc herniation, vertebral fracture, spinal stenosis). Those with back pain due to job-related injuries were not included since such patients were routinely evaluated in a separate Occupational Medicine clinic. We mailed letters to patients shortly after their index back pain visits and invited them to participate in the study. Research specialists called patients 1–3 weeks (mean: 12.7 days; SD=7.1) after their back pain visit to explain the study, confirm eligibility and obtain informed consent. Trained interviewers administered surveys by telephone at study baseline, 2-month follow-up, and 6-month follow-up.

Baseline Study Assessments—The original 9-item version of the STarT Back tool was administered by telephone at the baseline assessment, and was used to implement risk stratification per the usual recommended system of cutoffs (the ‘STarT Back risk groups’). This scoring system converts the 0 to 9 total score, and the 0 to 5 psychological score, into three categories: 1) low risk (a total score of 3 or less), 2) medium risk (a total score of 4 or more, and a psychological score of 3 or less), and 3) high risk (a total score of 4 or more, and a psychological score of 4 or more).

Participants reported their race and ethnicity; how long ago their current episode of back pain began, in weeks, months, and/or years; a global rating of improvement since their index back pain visit (“completely recovered”, “much better”, “better”, “not changed”, “worse”, vs. “much worse”); highest level of education attained; and back pain intensity measured on a 0 to 10 numerical pain rating scale (NRS).¹⁵

Outcomes—The primary study outcome was the modified Roland-Morris Disability Questionnaire (RMDQ) assessed at 6-month follow-up. The modified RMDQ is a measure of back-related functional limitations ranging from 0 to 23, with higher scores indicating greater functional limitations.¹⁶ This version of the RMDQ has been shown to have construct validity in comparison with other functional measures, internal consistency, reliability, and responsiveness.^{16, 17} To remain consistent with the RMDQ cut-offs used in the original STarT Back studies,^{3, 4} participants reporting RMDQ ≥ 7 at 6-month follow-up were classified as having persistent disabling back pain, and those with RMDQ scores <7 were classified as not having persistent disabling back pain.

Statistical Analysis

We descriptively characterized the study sample of MATCH participants who did not receive implementation of the STarT Back approach (Figure 1). To examine calibration of the STarT Back Tool, we calculated the proportion of individuals with persistent disabling back pain at 6-month follow-up in each risk group, the relative risk (RR) of persistent disabling back pain at 6-month follow-up in each risk group, and 95% confidence intervals (CIs) for these

proportions and RRs. We descriptively compared these estimates with those from 3 prior studies that reported on the predictive validity of the STarT Back risk groups using the same definition of persistent disabling back pain at follow-up (RMDQ ≥ 7), including the original STarT Back validation study.^{4, 7, 20} Next, we examined discrimination by calculating the performance characteristics of the STarT Back categories, applying the same contrasts used in the original STarT Back validation study: contrasting high-risk with low/medium-risk, and contrasting low-risk with medium/high-risk. Last, we examined the external predictive validity of the 3 STarT Back risk groups within 3 subgroups: 1) older adults (age ≥ 65 years), 2) those with acute back pain (<4 weeks duration), and 3) those with self-reported recent major improvement reflecting recovery or near-recovery (those reporting being 'completely recovered' or 'much better' compared to their index back pain visit, about 2 weeks earlier).

We did not examine the performance characteristics of the continuous STarT Back score for two reasons: 1) the matched/targeted treatment component of the STarT Back approach is coupled to the specific cut points used to define low-, medium-, and high-risk groups, and 2) we expected that it would be infeasible for US primary care physicians in routine clinical practice to make predictions and tailor treatment decisions based on the continuous STarT Back score. Since the practical value of the STarT Back tool is as a stand-alone tool, we did not adjust for other covariates in our analysis.

RESULTS

Figure 1 illustrates the flow of participants in this study. The sample included 1218 participants who completed baseline assessments (Table 1). The STarT Back tool classified 510 participants (42%) as low-risk, 447 (37%) as medium-risk, and 261 (21%) as high-risk. Roughly one third of participants (36%) were age ≥ 65 years. The majority (80%) were white and had graduated from college (53%). Baseline mean (standard deviation [SD]) RMDQ score reflecting back-related functional limitations was 11.6 (6.2). 938 participants (77%) had RMDQ scores ≥ 7 : 261 participants (51%) in the low-risk group, 419 participants (94%) in the medium-risk group, and 258 participants (99%) in the high-risk group. Eighteen percent of participants reported acute back pain, and 24% reporting recent self-reported major improvement.

At 6-month follow-up 1109 participants (91%) completed the RMDQ (Figure 1/Table 2). The mean (SD) RMDQ score at follow-up was 7.8 (6.8). STarT Back risk group was significantly associated with persistent disabling back pain (RMDQ ≥ 7) at 6-month follow-up, and the proportion of individuals with persistent disabling back pain was 22% in the low risk group, 62% in the medium risk group, and 80% in the high risk group (Table 2). Figure 2 presents these data graphically as relative risks (RRs). Overall, the proportions and RRs for the STarT Back risk groups in MATCH were comparable to those reported in the original STarT Back validation study by Hill et al. at 6-month follow-up (Table 2/Figure 2). The proportions and RRs of those with persistent disabling back pain at 6-month follow-up according to STarT Back risk group in MATCH were also generally similar to those reported in Danish (Morso et al.⁷) and Dutch (Bier et al.²⁰) primary care at 3-month follow-up, when accounting for the width of confidence intervals (Table 2/Figure 2).

The sensitivities, specificities, positive likelihood ratios, and negative likelihood ratios of the STarT Back risk groups in MATCH for predicting persistent disabling back pain at 6-month follow-up were very similar to those reported in the original STarT Back validation study and suggest good discrimination (Table 3).

In secondary analyses of subgroups with age ≥ 65 years, acute back pain, or recent major improvement, STarT Back risk group was significantly associated with persistent disabling back pain at 6-month follow-up across all 3 subgroup analyses (Table 4). In older adults, the proportion of participants with persistent disabling back pain at 6 months was high in both the medium (71%) and high risk (79%) groups, with considerable overlap in the confidence intervals between these 2 risk groups and no significant difference between medium and high risk groups ($p=.17$). Similarly, in those with acute back pain, the proportion of participants with persistent disabling back pain at 6 months was comparable between the medium (51%) and high risk (55%) groups ($p=0.71$). In contrast, in subgroup analyses of those with recent major improvement, the proportions with persistent disabling back pain at 6 months were quite different between each of the 3 STarT Back risk groups (Table 4). Finally, the proportion of participants with persistent disabling back pain at 6 months within each STarT Back risk group were generally higher in older adults, and lower in those with acute back pain and recent major improvement, when compared to the corresponding proportions within each STarT Back risk group in the entire sample (from Table 2).

DISCUSSION

This study found that the STarT Back risk groups were significant predictors of persistent disabling back pain at 6-month follow-up in a US primary care sample, successfully separating patients into distinct categories of risk. The predictive ability of the STarT Back risk groups was remarkably similar to that seen in the original UK validation study in primary care and subsequent replication in Danish/Dutch primary care, despite likely differences in health care systems and characteristics of patient populations between these countries. Our findings support the external predictive validity of the STarT Back risk groups in a US primary care setting, a necessary prerequisite to future efforts to match treatments to predicted risk using the STarT Back tool.

In contrast to the original STarT Back development and validation studies, the MATCH trial allowed the inclusion of participants ≥ 65 years of age, a subgroup of patients with relatively poor back-related functional outcomes,²¹ so that the study findings would have broad generalizability to all primary care patients with nonspecific back pain. The inclusion of older participants, and the higher levels of baseline disability in the current study (mean baseline RMDQ=11.6) as compared to the original STarT Back validation study (mean baseline RMDQ=9.1), may in part explain why the proportions of patients with persistent disabling back pain at follow-up within each STarT Back risk group were higher in the current study as compared to the original. Despite these possible differences, the sensitivities, specificities, and likelihood ratios for the two major distinctions permitted by the STarT Back risk groups (low vs. medium/high, low/medium vs. high) in MATCH were similar to those from the original STarT Back validation study. As used in the UK,^{3,5} each of these two distinctions guide a decision pertinent to matching treatments to predicted risk:

1) whether or not to refer to physical therapy ('no' in low risk vs. 'yes' in medium/high risk), and 2) whether or not to refer for a more rigorous, psychologically-informed physical therapy ('no' in low/medium risk vs. 'yes' in high risk). Given that the STarT Back risk groups appear similarly accurate in predicting persistent disabling back pain in the US as compared to the UK, if the treatments matched to predicted risk are also successfully implemented and similarly effective in the US, the STarT Back approach taken as a whole (stratification + targeted treatment) holds promise as an effective strategy in US primary care.

Contrary to a prior study,⁹ we found that the STarT Back risk groups defined distinct and increasing risk of persistent disabling back pain in those with early trajectories of major improvement, despite better overall outcomes in this subgroup as compared to all individuals with back pain. In contrast, in older adults, the STarT Back medium, and high-risk groups both had a high proportion (>70%) of participants with persistent disabling pain at 6-month follow-up. Similarly, in those with acute back pain <1 month duration, the STarT Back medium, and high-risk groups both had a moderate proportion (>50%) of participants with persistent disabling pain at 6-month follow-up. Thus, there may not be a meaningful difference in persistent disabling pain outcomes between medium, and high-risk groups among older adults and those with acute back pain. Our study examined only certain subgroups where we suspected the STarT Back tool might not perform well; future US studies may wish to attempt replication of these findings, and such studies may also examine the performance of the STarT Back Tool in other relevant subgroups, where applicable.

Limitations

Strengths of this study include its large sample size and low proportion with (91%) loss to follow-up. Primary care within the integrated Group Health system (now Kaiser Permanente Washington) likely reflects typical care in the Pacific Northwest region of the US, and risk prediction tools developed here (such as the Chronic Pain Grade)^{22, 23} have previously been exported to other clinical contexts. However, these results still may not be generalizable to other clinical settings in the US. Another potential limitation of the study is that we compared our results with that of prior studies in a descriptive manner, rather than in hypothesis-driven comparisons using individual-level data. Some features of our study also warrant attention to understand the limited conditions to which our study findings might be generalized. We examined the predictive ability of the STarT Back risk groups. Studies examining the predictive ability of the continuous STarT Back score might find different results, however, since it is as yet unclear how alternative cutoffs using the continuous STarT Back score might inform matched/targeted treatment options in the future, we did not choose such a path. In addition, we examined persistent disabling pain as a dichotomous outcome, defined using the same cutoff (RMDQ = 7) as the original STarT Back study. Other studies using outcomes such as the continuous RMDQ score, 30% improvement from baseline, or other outcomes (pain improvement, global perceived recovery, health care utilization, etc) may have results that are quite different from ours. On the other hand, such studies are applying conditions different than that used in the original STarT Back study, and it is reasonable to expect that doing so would decrease the tool's predictive capability. We note that several studies finding less impressive outcomes when replicating the STarT Back

method of risk stratification have modified the tool or risk categories in some way^{13, 24}, used an alternative outcome,^{9, 10, 12} and/or been conducted in settings other than primary care physicians' clinics.^{9,11} It may not be fair to expect the STarT Back tool to perform for risk stratification in such a scenario, given the deliberate development of the tool as a simple-to-complete yet relatively crude predictor of outcomes, rather than an elaborate measure meant to capture all subtleties of risk prediction even while using a range of continuous outcome measures.

CONCLUSIONS

We found that the STarT Back risk groups successfully separated people with back pain into distinct categories of risk for persistent disabling back pain at 6-month follow-up in US primary care. This is a necessary first step towards identifying whether the entire STarT Back approach, including matched/targeted treatment, can be effectively used for primary care in the US. Future research studies should prioritize understanding how the STarT Back risk groups may best inform choices among the various treatment options commonly available in the US.

Supplementary Material

Refer to Web version on PubMed Central for supplementary material.

Acknowledgments

We would like to thank Dr. Jonathan Hill for providing comments during the preparation of this manuscript.

We would like to thank the MATCH study participants.

Funding sources:

Funding for this research 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). Dr. Suri is employed by VA Puget Sound Health Care System. Dr. Suri is funded by Career Development Award #1IK2RX001515 from VA Rehabilitation Research and Development. The views and opinions expressed in this paper are those of the authors and do not necessarily reflect those of the VA, Patient-Centered Outcomes Research Institute, the National Institutes of Health, or an US government agency. None of the agencies that funded the study had any input into the design, conduct, data collection, analysis or interpretation, reporting, or writing of this article.

Abbreviations

US	United States
UK	United Kingdom
MATCH	Matching Appropriate Treatments to Consumers Healthcare
EHR	Electronic health Record
NRS	Numerical Pain Rating Scale
RMDQ	Roland-Morris Disability Questionnaire

SD	Standard deviation
CI	Confidence interval
RR	Relative risk

References

1. Hoy D, March L, Brooks P, Blyth F, Woolf A, Bain C, et al. The global burden of low back pain: estimates from the Global Burden of Disease 2010 study. *Annals of the rheumatic diseases*. 2014; 73(6):968–74. [PubMed: 24665116]
2. Dieleman JL, Baral R, Birger M, Bui AL, Bulchis A, Chapin A, et al. US Spending on Personal Health Care and Public Health, 1996–2013. *JAMA*. 2016; 316(24):2627–46. [PubMed: 28027366]
3. Hill JC, Whitehurst DG, Lewis M, Bryan S, Dunn KM, Foster NE, 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(9802):1560–71. [PubMed: 21963002]
4. Hill JC, Dunn KM, Lewis M, Mullis R, Main CJ, Foster NE, et al. A primary care back pain screening tool: identifying patient subgroups for initial treatment. *Arthritis Rheum*. 2008; 59(5): 632–41. [PubMed: 18438893]
5. Foster NE, Mullis R, Hill JC, Lewis M, Whitehurst DG, Doyle C, 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(2):102–11. [PubMed: 24615305]
6. 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 Nonrandomized Controlled Trial. *Spine (Phila Pa 1976)*. 2016; 41(8):645–52. [PubMed: 26630423]
7. Morso L, Kent P, Albert HB, Hill JC, Kongsted A, Manniche C. The predictive and external validity of the STarT Back Tool in Danish primary care. *Eur Spine J*. 2013; 22(8):1859–67. [PubMed: 23397189]
8. Mehling WE, Avins AL, Acree MC, Carey TS, Hecht FM. Can a back pain screening tool help classify patients with acute pain into risk levels for chronic pain? *Eur J Pain*. 2015; 19(3):439–46. [PubMed: 25381748]
9. Newell D, Field J, Pollard D. Using the STarT Back Tool: Does timing of stratification matter? *Manual therapy*. 2015; 20(4):533–9. [PubMed: 25175750]
10. Fritz JM, Beneciuk JM, George SZ. Relationship between categorization with the STarT Back Screening Tool and prognosis for people receiving physical therapy for low back pain. *Phys Ther*. 2011; 91(5):722–32. [PubMed: 21451094]
11. Morso L, Kent P, Manniche C, Albert HB. The predictive ability of the STarT Back Screening Tool in a Danish secondary care setting. *Eur Spine J*. 2014; 23(1):120–8. [PubMed: 23793455]
12. Morso L, Kongsted A, Hestbaek L, Kent P. The prognostic ability of the STarT Back Tool was affected by episode duration. *Eur Spine J*. 2016; 25(3):936–44. [PubMed: 25835771]
13. Mehling WE, Avins AL, Acree MC, Carey TS, Hecht FM. Can a back pain screening tool help classify patients with acute pain into risk levels for chronic pain? *European journal of pain*. 2015; 19(3):439–46. [PubMed: 25381748]
14. Cherkin D, Balderson B, Brewer G, Cook A, Estlin KT, Evers SC, et al. Evaluation of a risk-stratification strategy to improve primary care for low back pain: the MATCH cluster randomized trial protocol. *BMC Musculoskelet Disord*. 2016; 17(1):361. [PubMed: 27553626]
15. Jensen MP, Turner JA, Romano JM, Fisher LD. Comparative reliability and validity of chronic pain intensity measures. *Pain*. 1999; 83(2):157–62. [PubMed: 10534586]
16. 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(17):1899–908. discussion 909. [PubMed: 8560339]
17. Lauridsen HH, Hartvigsen J, Manniche C, Korsholm L, Grunnet-Nilsson N. Responsiveness and minimal clinically important difference for pain and disability instruments in low back pain patients. *BMC Musculoskelet Disord*. 2006; 7:82. [PubMed: 17064410]

18. Jensen MP, Strom SE, Turner JA, Romano JM. Validity of the Sickness Impact Profile Roland scale as a measure of dysfunction in chronic pain patients. *Pain*. 1992; 50(2):157–62. [PubMed: 1408311]
19. Roland M, Fairbank J. The Roland-Morris Disability Questionnaire and the Oswestry Disability Questionnaire. *Spine (Phila Pa 1976)*. 2000; 25(24):3115–24. [PubMed: 11124727]
20. Bier JD, Ostelo R, van Hooff ML, Koes BW, Verhagen AP. Validity and Reproducibility of the STarT Back Tool (Dutch Version) in Patients With Low Back Pain in Primary Care Settings. *Phys Ther*. 2017; 97(5):561–70. [PubMed: 28340202]
21. Rundell SD, Sherman KJ, Heagerty PJ, Mock CN, Jarvik JG. The clinical course of pain and function in older adults with a new primary care visit for back pain. *J Am Geriatr Soc*. 2015; 63(3):524–30. [PubMed: 25754841]
22. Von Korff M, Dunn KM. Chronic pain reconsidered. *Pain*. 2008; 138(2):267–76. [PubMed: 18226858]
23. Von Korff M, Ormel J, Keefe FJ, Dworkin SF. Grading the severity of chronic pain. *Pain*. 1992; 50(2):133–49. [PubMed: 1408309]
24. Von Korff M, Shortreed SM, Saunders KW, LeResche L, Berlin JA, Stang P, et al. Comparison of back pain prognostic risk stratification item sets. *J Pain*. 2014; 15(1):81–9. [PubMed: 24295877]

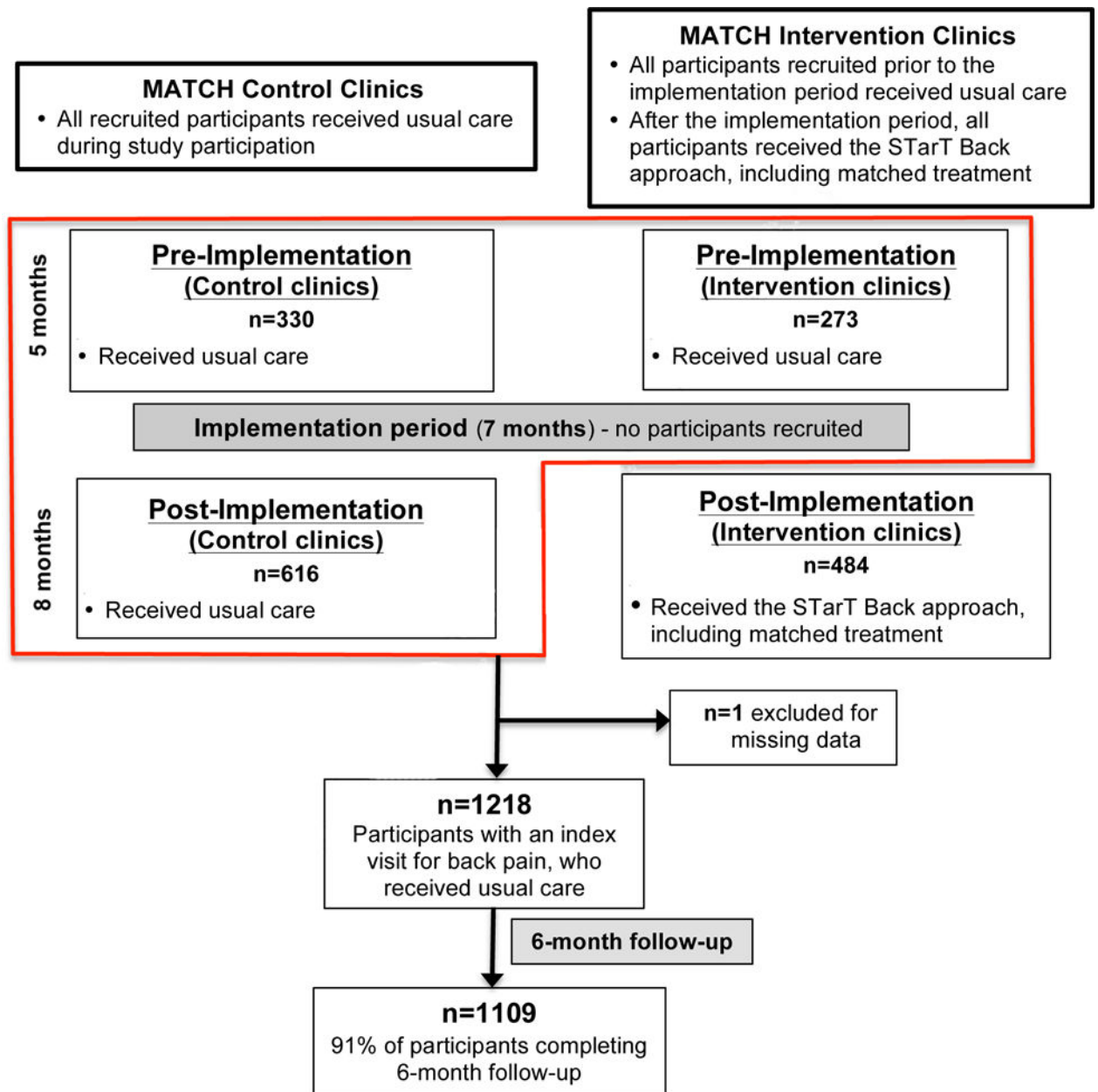


Figure 1. Flowchart of Study Participants
MATCH= Matching Appropriate Treatments to Consumers Healthcare study

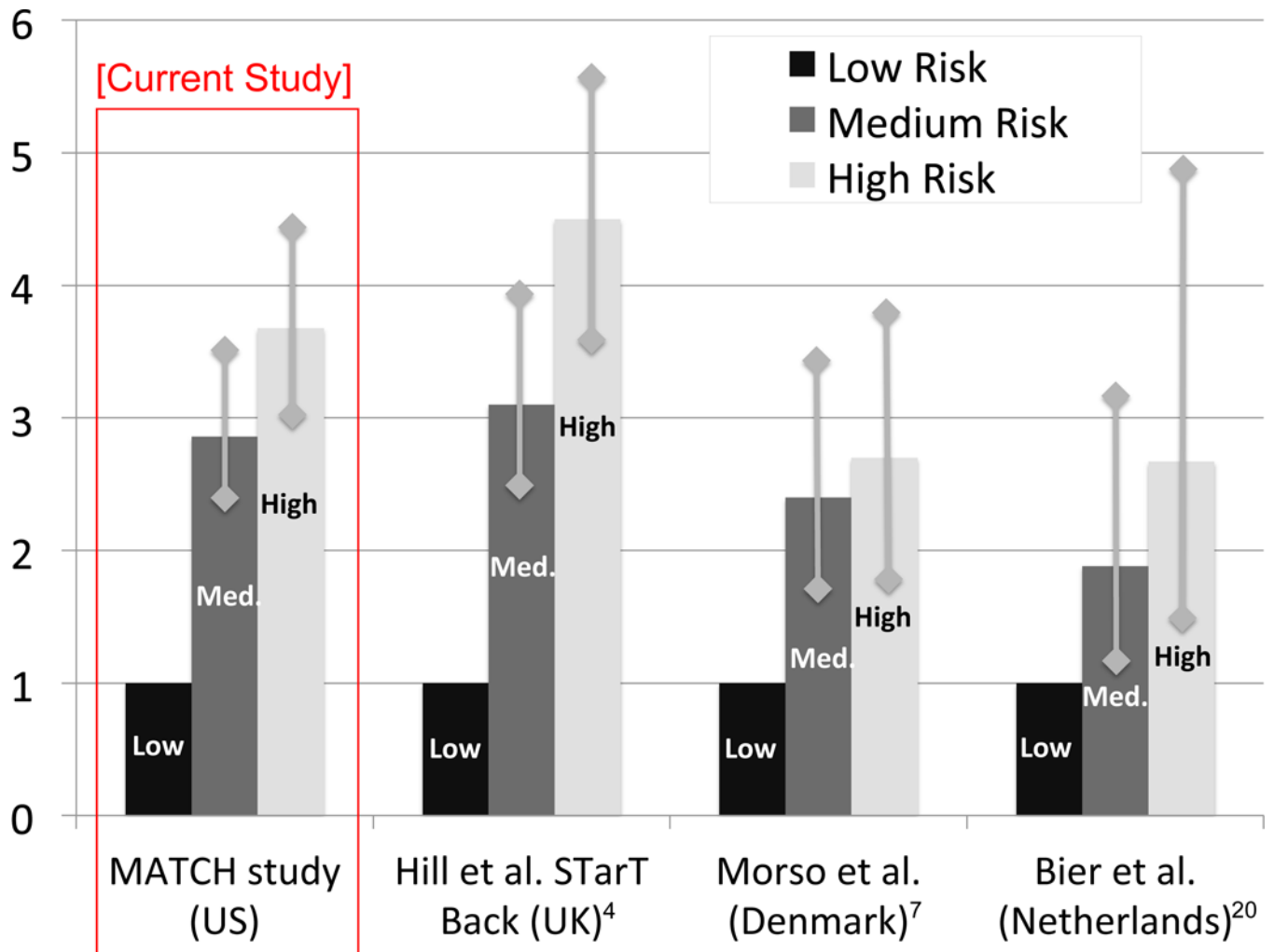


Figure 2. Relative Risks of Persistent Disability Pain in MATCH and prior studies
 MATCH= Matching Appropriate Treatments to Consumers Healthcare study; US=United States, UK=United Kingdom Vertical lines with diamond heads represent 95% confidence intervals.

Table 1

Characteristics of the Study Sample at the Baseline Assessment (n=1218)

	n (%) or mean \pm SD
Age	
Age (years)	58.0 \pm 17.6
Age groups	
18–39 years	268 (22.0%)
40–54 years	269 (22.1%)
55–64 years	245 (20.1%)
65+ years	436 (35.8%)
Female Sex	662 (54.4%)
Race	
American Indian/Alaskan Native	3 (0.3%)
Asian	45 (3.8%)
Black/African-American	85 (7.1%)
Native Hawaiian/Pacific Islander	8 (0.7%)
White	991 (82.9%)
Other	20 (1.7%)
Mixed	44 (3.7%)
Hispanic	60 (5.0%)
Education	
High school or less	187 (15.4%)
Technical/trade school	62 (5.1%)
Some college	326 (26.8%)
College graduate	325 (26.7%)
Graduate school	317 (26.1%)
Back-related functional limitations (RMDQ) (range 0–24)	11.6 \pm 6.2
Back Pain Intensity (NRS) (range 0–10)	5.4 \pm 2.5
S^TarT Back risk group	
Low	510 (41.9%)
Medium	447 (36.7%)
High	261 (21.4%)
Anxiety (GAD-7) (range 0–21)	4.2 \pm 4.6
Depression (PHQ-8) (range 0–28)	6.2 \pm 5.4
Duration of back pain	

	n (%) or mean \pm SD
<2 weeks	19 (1.7%)
2–4 weeks	188 (16.5%)
1–3 months	172 (15.1%)
4–6 months	128 (11.2%)
7 mos.–3 yrs.	333 (29.2%)
>3 years	299 (26.3%)
Global rating of improvement since the index visit	
Completely recovered	41 (3.4%)
Much better	246 (20.3%)
Better	384 (31.7%)
Not changed	455 (37.6%)
Worse/much worse	85 (7.0%)

RMDQ=Roland Morris Disability Questionnaire, NRS=numerical rating scale, GAD-7= Generalized Anxiety Disorder-7, PHQ-8= Patient Health Questionnaire-8

Author Manuscript

Author Manuscript

Author Manuscript

Author Manuscript

Table 2

Proportions of primary care patients with persistent, disabling back pain at follow-up, by STarT Back risk group*

	RMDQ 7 at 6 months MATCH (US) (n=1109)		p-value ^b	RMDQ 7 at 6 months STarT Back (UK) ^d (n=500)		RMDQ 7 at 3 months Denmark ^e (n=344)		Netherlands ²⁰ (n=150)
	n/n _d	proportion (95% CI) ^a		n/n _d	proportion (95% CI) ^{a,c}	proportion ^c	n/n _d	
STarT Back risk group (at baseline)								
Low	102/471	21.7% (18.0%–25.7%)		39/234	16.7% (12.1%–22.1%)	24%	16/76	21.1% (12.5%–31.9%)
Medium	252/407	61.9% (57.0%–66.7%)	<0.0001	99/186	53.2% (45.8%–60.6%)	57%	22/58	37.9% (25.5%–51.6%)
High	184/231	79.7% (73.9%–84.7%)		58/74	78.3% (67.3%–87.1%)	64%	9/16	56.3% (29.8%–80.2%)

RMDQ= Roland-Morris Disability Questionnaire, MATCH= Matching Appropriate Treatments to Consumers Healthcare needs trial, US=United States, UK=United Kingdom, n=number with the outcome, n_d=number in the denominator, CI=confidence intervals

* Persistent, disabling back pain is defined as RMDQ 7 at 6 months

^a 95% CIs for binomial proportions using Clopper-Pearson method

^b p-for trend

^c Data extracted from publications or provided by study authors. Where numbers/CIs are not presented, they were not available. Precision is as stated in publications or as provided by authors.

Table 3

Performance characteristics for STarT Back risk group cutoffs to predict persistent, disabling back pain at 6 months*

	Sensitivity	Specificity	Negative LR (95% CI)	Positive LR (95% CI)
STarT Back risk group cutoffs in MATCH (United States)				
Low vs. Medium/High	81.0%	64.6%	0.29 (0.24–0.35)	2.29 (2.04–2.58)
Low/Medium vs. High	34.2%	91.8%	0.72 (0.67–0.77)	4.16 (3.08–5.60)
STarT Back risk group cutoffs in original validation study^a (United Kingdom⁴)				
Low vs. Medium/High	80.1%	65.4%	0.30 (0.23–0.40)	2.32 (1.96–2.76)
Low/Medium vs. High	39.6%	94.6%	0.74 (0.67–0.81)	5.51 (3.30–9.28)

MATCH= Matching Appropriate Treatments to Consumers Healthcare needs trial, LR=likelihood ratio, CI=confidence interval

* Persistent, disabling back pain is defined as RMDQ ≥ 7 at 6 months

^a data extracted from publications or provided by study authors

Table 4
Proportions with persistent, disabling back pain at follow-up, by STarT Back risk group, in specific subgroups*

	RMDQ 7 at 6 months in older adults (age 65 years) n=414		RMDQ 7 at 6 months in acute back pain (< 1 month) n=183		RMDQ 7 at 6 months in those with recent improvement n=268		p-value ^b
	n/n (proportion)	95% CI for proportion ^a	n/n (proportion)	95% CI for proportion ^a	n/n (proportion)	95% CI for proportion ^a	
STarT Back risk group (at baseline)							
Low	44/165 (26.7%)	20.1%–34.1%	12/87 (13.8%)	7.3%–22.9%	27/168 (16.1%)	10.9%–22.5%	
Medium	116/164 70.7%	63.1%–77.6%	33/65 (50.8%)	38.1%–63.4%	35/83 (42.2%)	31.4%–53.5%	<0.0001
High	67/85 78.8%	68.6%–86.9%	17/31 (54.8%)	36.0%–72.7%	12/18 (66.7%)	41.0%–86.7%	
n/n _d (proportion)		proportion (95% CI) ^a	n/n _d	proportion (95% CI) ^a	n/n _d	proportion (95% CI) ^a	p-value ^b
STarT Back risk group (at baseline)							
Low	44/165	26.7% (20.1%–34.1%)	12/87	13.8% (7.3%–22.9%)	27/168	16.1% (10.9%–22.5%)	
Medium	116/164	70.7% (63.1%–77.6%)	33/65	50.8% (38.1%–63.4%)	35/83	42.2% (31.4%–53.5%)	<0.0001
High	67/85	78.8% (68.6%–86.9%)	17/31	54.8% (36.0%–72.7%)	12/18	66.7% (41.0%–86.7%)	

RMDQ= Roland-Morris Disability Questionnaire, MATCH= Matching Appropriate Treatments to Consumers Healthcare needs trial, US=United States, UK=United Kingdom

* Persistent, disabling back pain is defined as RMDQ 7 at 6 months

^a 95% CIs for binomial proportions using Clopper-Pearson method

^b p-for trend