

## Supplementary Appendix S1. Search strategy

MEDLINE	EMBASE	CINAHL
1. RCT.mp.	1. RCT.mp.	1. (MH "Back Pain+")
2. randomised.mp.	2. randomised.mp.	2. MH low back pain/
3. Randomized Controlled Trial/	3. Randomized Controlled Trial/	3. mh radiculopathy/
4. randomised trial.mp.	4. randomised trial.mp.	4. mh sciatica
5. cluster.mp.	5. cluster.mp.	5. mh "musculoskeletal pain"
6. Prospective Studies/	6. Prospective Studies/	6. mh back pain or low back pain or radiculopathy or sciatica or back ache or lumbago
7. 1 or 2 or 3 or 4 or 5 or 6	7. 1 or 2 or 3 or 4 or 5 or 6	7. S1 OR S2 OR S3 OR S4 OR S5 OR S6
8. Back Pain/	8. Back Pain/	8. mh screen*
9. low back pain/	9. low back pain/	9. "risk screen"
10. radiculopathy/	10. radiculopathy/	10. "tool"
11. sciatica/	11. sciatica/	11. "instrument/"
12. (back pain or low back pain or radiculopathy or sciatica or back ache or lumbago).mp. [mp=title, book title, abstract, original title, name of substance word, subject heading word, floating sub-heading word, keyword heading word, organism supplementary concept word, protocol supplementary concept word, rare disease supplementary concept word, unique identifier, synonyms]	12. (back pain or low back pain or radiculopathy or sciatica or back ache or lumbago).mp. [mp=title, abstract, heading word, drug trade name, original title, device manufacturer, drug manufacturer, device trade name, keyword heading word, floating subheading word, candidate term word]	12. (mh questionnaires+ )
13. Musculoskeletal Pain/	13. Musculoskeletal Pain/	13. "start back/"
14. 8 or 9 or 10 or 11 or 12 or 13	14. 8 or 9 or 10 or 11 or 12 or 13	14. "stratified care"
15. screen\$.mp.	15. screen\$.mp.	15. TX stratif
16. risk screen\$.mp.	16. risk screen\$.mp.	16. mh prognosis
17. tool.mp.	17. tool.mp.	17. S8 OR S9 OR S10 OR S11 OR S12 OR S13 OR S14 OR S15 OR S16
18. instrument.mp.	18. instrument.mp.	18. TX RCT
19. questionnaire/	19. questionnaire/	19. "randomised/"
20. start back.mp.	20. start back.mp.	20. (MH "Randomized Controlled Trials+")
21. stratified care.mp.	21. stratified care.mp.	21. TX randomised trial
22. stratif\$.mp.	22. stratif\$.mp.	22. "cluster"
23. prognosis/	23. prognosis/	23. (MH "Prospective Studies+")
24. 15 or 16 or 17 or 18 or 19 or 20 or 21 or 22 or 23	24. 15 or 16 or 17 or 18 or 19 or 20 or 21 or 22 or 23	24. S18 OR S19 OR S20 OR S21 OR S22 OR S23
25. 7 and 14 and 24	25. 7 and 14 and 24	25. S7 AND S17 AND S24
26. limit 25 to humans	26. limit 25 to (human and "remove medline records")	

**Supplementary Appendix S2. List of excluded full-text articles and the primary reason for exclusion**

<b>Study</b>	<b>Title</b>	<b>Reason for exclusion</b>
Beneciuk et al 2017	Identifying Treatment Effect Modifiers in the STarT Back Trial: A Secondary Analysis	Wrong study design
Bier et al 2017	Can Primary Care for Back and/or Neck Pain in the Netherlands Benefit from Stratification for Risk Groups According to the STarT Back Tool Classification?	Wrong study design
Fritz et al 2011	Relationship Between Categorization with the STarT Back Screening Tool and Prognosis for People Receiving Physical Therapy for Low Back Pain	Treatment did not follow STarTBack algorithm
Hill et al 2022	Risk-based stratified primary care for common musculoskeletal pain presentations (STarT MSK): a cluster-randomised, controlled trial	STarTBack tool not used to stratify patients
Magel et al 2017	Outcomes of Patients with Acute Low Back Pain Stratified by the STarT Back Screening Tool: Secondary Analysis of a Randomized Trial	Treatment did not follow STarTBack algorithm
Mansell et al 2016	Exploring What Factors Mediate Treatment Effect: Example of the STarT Back Study High-Risk Intervention	Wrong study design
Middleton et al 2020	Implementing stratified care for acute low back pain in primary care using the STarT Back instrument: a process evaluation within the context of a large pragmatic cluster randomized trial	Wrong study design
Saunders et al 2020	Stratified primary care versus non-stratified care for musculoskeletal pain: qualitative findings from the STarT MSK feasibility and pilot cluster randomized controlled trial	Wrong study design
Treanor et al 2022	Prospective observational study investigating the predictive validity of the STarT Back tool and the clinical effectiveness of stratified care in an emergency department setting	Wrong study design
Whitehurst et al 2015	Implementing Stratified Primary Care Management for Low Back Pain: Cost-Utility Analysis Alongside a Prospective, Population-Based, Sequential Comparison Study	Wrong study design

## Supplementary Appendix S3. All study level variables explored

Study	Individual administering tool	Clinician providing care	Did the tool lead to changes in the recommended treatment?	Proportion receiving matched treatment?	Risk level group	Allocated treatment	Was the tool collected in control group	Trained to use tool
Beneciuk 2015	Clinician (Physiotherapist)	All risk groups: Community physiotherapist	Not reported	Unclear	IG: Low (23%), Medium (44%), High (33%) CG: Low (36%), Medium (41%), High (23%)	IG: Treatment based on STarTBack score CG: Physiotherapy as 'usual'	Yes	Yes
Cherkin 2018	Researcher	All risk groups: A PCP	Not reported	Unclear	IG: Low (40%), Medium (38%), High (22%) CG: Low (42%), Medium (37%), High (22%)	IG: Treatment based on STarTBack score CG: "Usual care"	Yes	Yes
Choudhry 2022	Unclear	All risk groups: Physiotherapist, "spine coach" and an "ICE MD"	Not reported	Not reported	Unclear	IG: Treatment based on STarTBack score CG1: IPT CG2: No intervention	Yes, but has combined medium and high risk	Not reported
Delitto 2021	Unclear	High risk: Community physiotherapist	Higher referral for Physiotherapy in IG (high risk group) (57% vs 30%)	Not reported	IG: Low (36%), Medium (39%), High (24%) CG: Low (36%), Medium (40%), High (24%)	IG: Treatment based on STarTBack score CG: 'Usual care'	Yes, but only high-risk group	Yes
Foster 2014	Clinician (GP)	Low risk: Community GP Medium and high risk: Referral to community physiotherapist	Higher referral for physiotherapy in medium and high-risk group (72% vs 40%) No difference for low-risk group (65% vs 68% not referred)	Clinician followed recommended matched treatment in 71% of patients <sup>a</sup>	IG: Low (37%), Medium (41%), High (22%) CG: Low (39%), Medium (42%), High (20%)	IG: Treatment based on STarTBack score CG: GP +/- physiotherapy referral if appropriate	Yes	Yes
Hill 2011	Researcher	Low risk: Research physiotherapist Medium and high risk: Research physiotherapist	Higher referral for Physiotherapy in IG (75% vs 58%)	Clinician followed recommended matched treatment in 93% of patients <sup>a</sup>	IG: Low (26%), Medium (46%), High (28%) CG: Low (26%), Medium (46%), High (28%)	IG: Treatment based on STarTBack score CG: Physiotherapy referral if appropriate	Yes	Yes
Konstantinou 2020	Researcher	Low and medium risk: Research physiotherapist High risk: Referred for MRI and Specialist	Not reported	Clinician followed recommended matched treatment in: Low risk: 98.8% Medium risk: 82.5% High risk: 81.3%	IG: Low (22%), Medium (44%), High (34%) CG: Low (23%), Medium (45%), High (33%)	IG: Treatment based on STarTBack score ( <i>high risk referred for MRI and to specialist</i> ) CG: GP and Physiotherapy consult	Yes	Not reported
Koppenaar 2022	Unclear	All risk groups: Community physiotherapist	Not reported	Not reported	IG: Low (55%), Medium (36%), High (9%) CG: Low (61%), Medium (36%), High (3%)	IG: Treatment based on STarTBack score CG: Guideline based care as per RDAP by Physiotherapist	Yes	Unclear
Morsø 2021	Researcher	All risk groups: Community physiotherapist	Not reported	Not reported	IG: Low (31%), Medium (31%), High (38%) CG: Low (31%), Medium (31%), High (38%)	IG: Treatment based on STarTBack score CG: Guideline based care by GP and Physiotherapist	Yes	Yes
Murphy 2016	Clinician (Physiotherapist)	All risk groups: Community physiotherapist	Not reported	Unclear	IG: Low (17%), Medium (54%), High (29%)	IG: Treatment based on STarTBack score CG: Generic physiotherapy	Yes	Yes

Rhon 2023	Self-administered	All risk groups: Community physiotherapist	Not reported	Not reported	CG: Low (21%), Medium (50%), High (29%) IG: Low (57%), Medium (36%), High (7%)  CG: Low (53%), Medium (39%), High (8%)	IG: Treatment based on STarTBack score CG: "Usual care"	Yes	Yes
-----------	-------------------	--	--------------	--------------	--	---	-----	-----

Use of the tool=who administered the tool/who stratified patients; EMR=electronic medical records; EHR=electronic health record; GP=general practitioner; PCP=primary care practitioner; PIPT=psychologically informed physiotherapy; ICE MD= identify, coordinate, enhance (ICE) model of care; MD=medical doctor; IG=intervention group; CG=control group; RDAP=Royal Dutch Association for Physiotherapy  
Unclear=reporting of the presence of a variable did occur, however not enough information was provided for a clear definitive answer  
Not reported=no reporting of the variable occurred