Brazilian Journal of Physical Therapy 28 (2024) 101116

Brazilian Journal of Physical Therapy

https://www.journals.elsevier.com/brazilian-journal-of-physical-therapy

SYSTEMATIC REVIEW

Does risk stratification with a matched treatment pathway improve clinical outcomes for adults with acute back pain? A systematic review and meta-analysis

Anthony Frank Chiodo*, Melanie Haley

Physiotherapy Department, Eastern Health, Victoria, Australia

Received 13 October 2023; received in revised form 20 April 2024; accepted 21 August 2024
Available online 5 September 2024

KEYWORDS
Adult;
Back pain;
Chronic pain;
Disabled persons;
Humans;
Randomized controlled trials as topic

Abstract
Background: Risk stratification is an approach which has been recommended across a number of international guidelines for the management of back pain.
Objective: To assess whether the use of risk stratification with a matched treatment pathway improves clinical outcomes, when compared with usual care or other interventions, in adults with acute back pain.
Methods: A comprehensive search was conducted of the databases Medline, Embase, PEDro, CINAHL and Cochrane Library in November 2022. Studies of adults with back pain of less than 3 months' duration and who had been stratified according to their level of risk of a poor functional outcome and provided with a treatment matched to their level of risk were included. Participants with specific and/or serious spinal pathologies were excluded.
Results: Five trials involving 3519 participants were included. Meta-analysis found very-low certainty evidence that the use of a risk stratification approach with matched treatment may lead to a very small reduction in pain levels at 3–6 months compared with usual care (MD -0.62, 95% CI -0.88, -0.36). These results did not achieve clinical significance. No difference was found for the use of risk stratification compared to usual care for disability (MD -1.52, 95% CI -4.15, 1.11).
Conclusion: The use of risk stratification with matched treatment may be just as worthwhile as usual care for acute back pain, however the evidence is very uncertain. Further high quality research is required to confirm whether risk stratification is a useful approach for this population.
Systematic review registration number: CRD42022379987

* Corresponding author at: Eastern Health, Physiotherapy Department, 5 Arnold Street, Box Hill 3128, Victoria, Australia.
E-mail: anthony.chiodo@easternhealth.org.au (A.F. Chiodo).

https://doi.org/10.1016/j.bjpt.2024.101116

Introduction

It has been estimated that around 70−90 % of the population will experience an episode of back pain in their lifetime,[1] and while the prognosis for those with acute back pain is generally favourable within the first six weeks,[2] it is estimated that between 7 % to 42 % will develop chronic pain.[3,4] Chronic back pain is challenging to manage, with commonly used interventions, such as exercise and multidisciplinary rehabilitation, only yielding modest and mostly short-term improvements on pain and function.[5] With the prevalence of back pain disability and associated economic burden increasing worldwide,[6] cost-effective interventions, which can effectively prevent the transition from acute to chronic pain, are required.

Risk stratification is recommended in back pain guidelines across a number of jurisdictions, including Australia, the United Kingdom, and Belgium.[7-9] This approach involves classifying patients into low, medium, and high risk of long-term disability or poor outcomes, and can be used to tailor treatment strategies.[8] The most commonly used risk stratification tools are the STarT Back Screening Tool[10,11] and the Örebro Musculoskeletal Pain Screening Questionnaire.[12,13] These tools can accurately estimate the risk of transition from acute to chronic pain,[3] and have demonstrated validity in differentiating a person's risk of poor prognosis,[14] future disability,[15] and, for the Örebro Musculoskeletal Pain Screening Questionnaire, return-to-work outcomes.[15] Using risk stratification, with treatment matched to an individual's level of risk, has been demonstrated to be cost-effective; facilitating resource distribution by preventing over treatment of low-risk patients,[16] and reducing back-pain related absenteeism.[17] The authors of the STarT Back Screening Tool have suggested tailored treatments for each level of risk; education for patients at low-risk, with additional individual physical therapy for patients at medium risk, and psychologically-informed physical therapy for those at high risk.[10]

Despite the efficacy of risk stratification with matched treatment for patients with back pain of any duration,[17] to our knowledge, no systematic review has sought to evaluate the use of a risk stratification approach specifically for improving clinical outcomes for acute back pain. The aim of this systematic review is to assess whether the use of risk stratification with a matched treatment pathway improves clinical outcomes, including transition to chronic pain, when compared with usual care or other interventions, in adults with acute back pain.

Methods

This systematic review was prospectively registered on PROSPERO (CRD4202237998) and followed the statement for Preferred Reporting Items for Systematic Reviews and Meta−Analysis (PRISMA).[18] A study protocol was not published.

Participants

Studies were deemed eligible for inclusion if they included adult participants with acute pain, using the commonly accepted descriptor for acute pain as being less than 3 months.[19,20] Participants with pain greater than 3 months duration or with specific and/or serious spinal pathologies (fracture, spinal haematoma, infection, malignancy, acute cord compression) were excluded. Patients were required to have been stratified according to their level of risk for developing chronic pain or having a poor functional outcome, and then provided a treatment matched to this level of risk. This included all participants from those at low risk of a poor outcome (and hence more likely to have a positive outcome) through to high risk of a poor outcome, provided that they had been stratified. Only randomized-controlled trials (RCTs) were included as systematic reviews of RCTs are considered the highest level of evidence.[21,22] For a complete description of the inclusion and exclusion criteria, refer to supplementary material − Table S.1.

Search strategy

A search was conducted across five databases (Medline, Embase, PEDro, CINAHL, and Cochrane Library) from database inception to November 21st, 2022. Studies were limited to humans and the English language. The reference lists of relevant articles were also scanned to identify any eligible studies which may have been missed by the database search. Refer to Table A.1 for an example search strategy. A complete search strategy for each database can be found in the supplementary material.

Selection process

Two reviewers independently screened titles and abstracts using the web-based Covidence software platform.[23] Relevant full texts were then independently reviewed by the two researchers to determine suitability for inclusion. Any conflicts in this process were resolved through consensus discussion between the two researchers with a third researcher available if consensus was unable to be achieved. One researcher independently collected and collated key data from the studies, which was then cross-referenced for accuracy with a second researcher. The extracted data were then converted to table form to allow for easy comparison between studies. Absent data were attempted to be obtained through review of published supplementary resources, search of ClinicalTrials.gov, and through contact with study authors. Authors were given 60 days to respond to written requests for missing information via email. Studies with absent data were assessed on a case-by-case basis. Studies without sufficient data for meta-analysis were analysed descriptively.

Outcome measures

Outcome data relating to participant outcomes such as pain and disability were extracted. Outcomes relating to cost and health-care utilization were not extracted as this was beyond the scope of the review. Outcomes were extracted for all available timeframes and then grouped into short-term (< 3 months), intermediate term (3−6 months), and long-term (>6 months). For studies that did not publish raw data relating to the standard deviations of outcomes, these were estimated using calculations outlined in the Cochrane Handbook for Systematic Reviews of Interventions.[24] Participant demographic information, settings, method of

stratification, and details of interventions, including frequency and duration, were also extracted.

Study risk of bias assessment

The quality of the included studies was assessed using the PEDro Scale; an 11-item tool that is validated to assess the methodological quality of studies.[25] Two reviewers independently appraised articles using this scale, with consensus discussion used for resolution of any conflicts.

Data analysis

Meta-analysis was conducted on trials with sufficient homogeneity, using a random effects model and presented as mean difference. Trials were considered sufficiently homogenous if there was a common outcome measure, comparator, and follow-up timeframe. Review Manager (RevMan) (Version 5.4)[26] was used to complete the meta-analysis. Where outcomes of interest (eg. pain, disability) were measured on the same numeric scale (eg. 0 − 100), these were synthesized using a mean difference approach. This was chosen over a standardized mean difference, to enhance comparability of findings. Outcomes for pain were presented using a 0 − 10 scale and outcomes for disability were presented using a 0 − 100 scale. The I^2 statistic[27] was used to quantify between-study heterogeneity. The results of meta-analysis were compared to pre-determined minimally clinically important differences (MCID), as proposed by Ostelo,[28] to establish clinical significance. The MCID was considered to be a change of 2 points of a 0 − 10 scale for pain outcomes and 10 points on a 0 − 100 scale for disability.[28]

Certainty assessment

To assess the certainty of evidence, the Grading of Recommendations, Assessment, Development and Evaluation (GRADE) approach[29] was utilised. This process involves evaluating each meta-analysis and downgrading the quality of evidence across five domains. Meta-analyses of RCTs are initially considered high quality and may be downgraded to moderate, low, and very-low. The downgrade criteria were informed by Guyatt [30-34] and are as follows: for risk of bias, analyses were downgraded by one if >25% of participants were from studies at a high risk of bias (PEDro <6) and by two if >50% of participants were from studies at high risk of bias. For inconsistency, analyses were downgraded by one if there was moderate heterogeneity (I^2 > 40%) and downgraded by two if there was very substantial heterogeneity (I^2 > 75%). For indirectness, analyses were downgraded by one if there was variation in one of the following, (or downgraded by two if there was variation in more than one): outcome measures used, timeframes, or interventions. For imprecision, analyses were downgraded if there were fewer than 2000 participants included or if the 95% CI width was greater than twice the minimally clinically important difference; and for publication bias, analyses were downgraded if <10 studies were included. Refer to supplementary material- Table S.2 for a complete summary of the GRADE downgrade criteria utilised.

Results

Study selection

Our initial search generated 3991 articles with removal of duplicates. From this, 3952 were excluded based on title and abstract screening, leaving 39 which was reduced to 5 following review of full text. Refer to Fig. 1 for the complete PRISMA flowchart. Five studies were included in the analysis, encompassing 3519 participants. Study size ranged from 47[35] to 2300 participants[36] and were completed in three high-income countries[37]: United States,[36,38,39] Hong Kong SAR (China),[35] and Germany.[40] Female participants made up, on average, 55% in each study and only one study[36] published demographic data relating to race and ethnicity. Hazard[39] did not publish baseline demographic data. Follow-up time ranged from 4 weeks to 12 months, with the majority of studies reporting outcomes between 3 and 6 months. The STarT Back Screening tool was the most common method of risk stratification, used in three studies[36,38,40] with the Örebro Musculoskeletal Pain Screening Questionnaire[35] and the Vermont Disability Prediction Questionnaire[39] used in the remaining two articles. These three tools are validated for predicting risk of poor outcomes; the STarT Back Screening Tool for future disability,[15] the Vermont Disability Prediction Questionnaire for return-to-work outcomes,[15] and the Örebro Musculoskeletal Pain Screening Questionnaire for both future disability and return-to-work outcomes.[15] Priebe et al[40] was the only study to also include participants classified as being low risk. Interventions varied across studies. Delitto et al[36] and Lee et al[35] incorporated treatment aligned with psychological principles, Magel et al[38] compared an early intervention with usual care, Priebe et al[40] involved consultation with a pain specialist and use of a back pain app and Hazard et al[39] involved letters sent to physicians with recommendations. Comparator interventions varied across studies. Hazard et al[39] compared risk stratification to no intervention, whereas the remaining studies compared risk stratification to usual care, which was considered management through a primary care clinician on an individual basis. Lee et al[35] was the only study to provide specific details on the nature of the usual care provided, incorporating electrophysical, manual, and exercise therapies delivered by a physical therapist. The most common outcomes of interest were pain and disability with only one study including transition to chronic back pain as an outcome measure.[36] The full details of the study characteristics are outlined in Table 1.

Risk of bias of individual studies

The study quality ranged from a PEDro score of 3 to 7, with a mean score of 5.4, indicating overall fair quality.[41] The risk of bias for each article is outlined in Table 2.

Summary of syntheses

Meta-analysis comparing risk stratification to usual care was completed incorporating 5 cohorts from 4 studies (Fig. 2). Hazard et al[39] was the only study to compare risk stratification to no intervention and did not contain sufficient data for quantitative analysis. For these reasons, this study was not included in the meta-analysis.

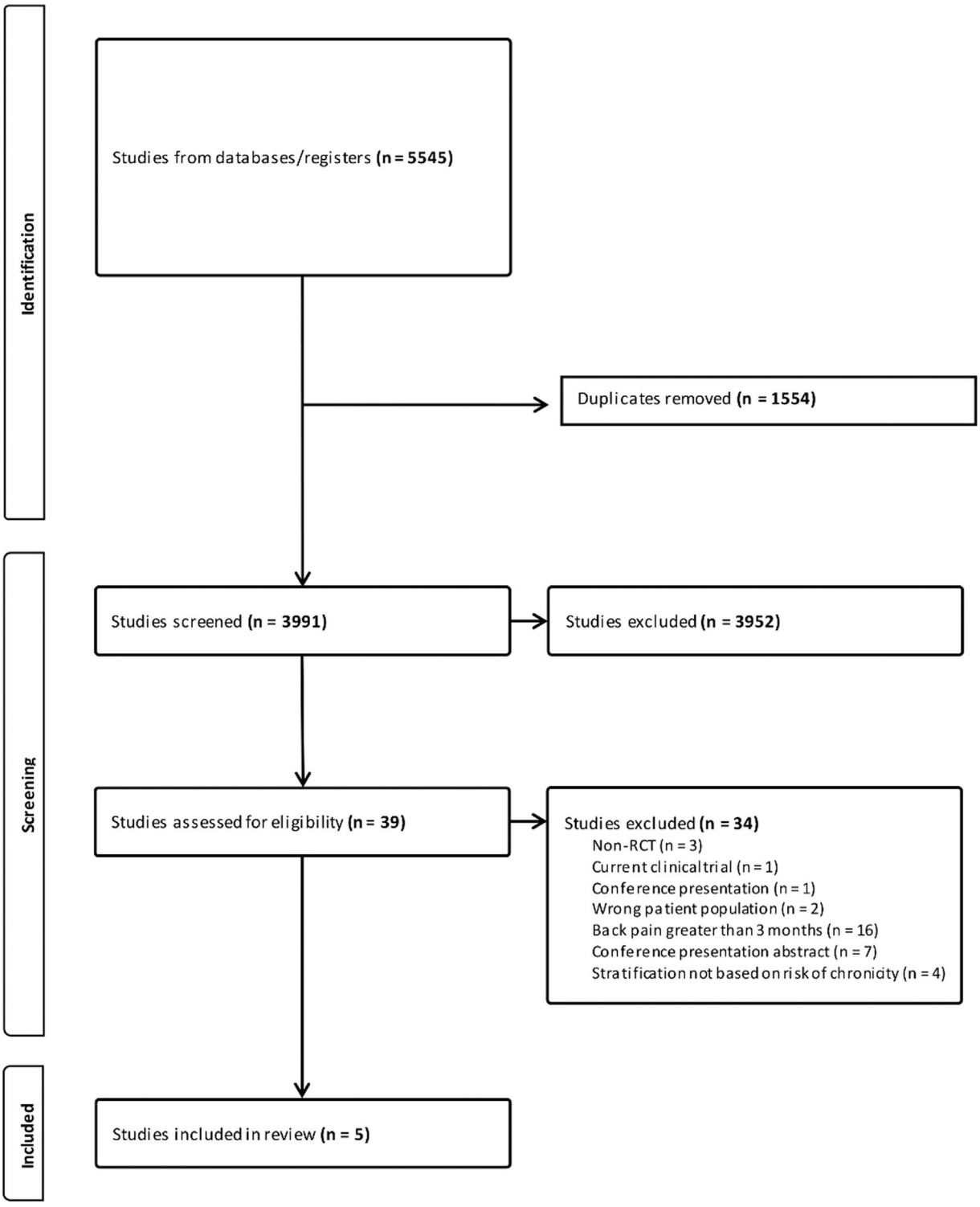

Fig. 1 PRISMA flowchart.

Magel et al,[38] Priebe et al,[40] and Delitto et al[36] shared common outcome measures, taken between 3 and 6 months (intermediate time-frame), and were therefore included in the meta-analysis. For Lee et al,[35] the timeframe for assessment of outcomes was not specified, however, participants were assessed at the conclusion of the intervention, which was up to 3 months in duration. Given the proximity to the timeframe of other included studies (3–6 months), this study was therefore included in the meta-analysis. As Magel et al[38] included separate outcome data for its medium- and high-risk groups, these were included separately in this meta-analysis despite being part of the same study. Magel et al[38] was also the only study to assess outcomes at a short-term (4 weeks) and long-term (12 months) period. As this was limited to one study, no synthesis was possible for these timeframes.

Table 1 Details of included studies.

Authors & Year	Population	Inclusion Criteria	Exclusion Criteria	Participants	Stratification Method	Group Details	Intervention Details	Outcomes	Study Results
Delitto et al., 2021	Primary care clinics (USA) Female: 59 %	-Primary complaint of LBP -Adults -Score of ≥4 on distress subscale STarT Back Screening Tool	-Patients with chronic symptoms, -Patients with non-musculoskeletal causes of LBP	*n* = 2300 Intervention (*n* = 1207) Control (*n* = 1093)	STarT Back Screening Tool High risk patients included	Both Groups: Education to participating primary care clinics including up-to date evidence-based management of back pain Usual Care + PIPT Group -PIPT approach (education, reducing fear of movement, improving coping skills, addressing physical impairments) Usual Care Group -At discretion of primary care provider	*Providers:* PTs who had completed PIPT training *Duration/Timing/Frequency:* not set (at discretion of individual clinician)	Primary *(6 months)* Transition to chronic LBP, ODI Secondary *(12 months)* LBP-related processes of care Medical Utilization	No difference between the groups in rates of transition to chronic pain (OR 0.83 95 % CI 0.64, 1.09) or disability (MD −2.1 95 % CI −4.9, 0.6)
Magel et al., 2017	Primary care clinics (Utah, USA) Female: 49.2 %	-ODI score of ≥20 % -Aged 18−60 years -LBP or numbness -Symptom duration < 16 days	-Spinal surgery, -Current pregnancy, -Currently receiving treatment for LBP, -Neurological symptoms, -Presence of "red flags", -LBP extending below the knee	*n* = 181 Medium Risk (*n* = 120; Intervention *n* = 59, Control *n* = 36) High Risk (*n* = 61; Intervention *n* = 25, Control *n* = 36)	STarT Back Screening Tool High and medium risk patients included	Both Groups: 20-minute education session based on the *Back Book* Early Intervention Group, -Management by primary care provider on as-needed basis, −4 PT sessions over first 4 weeks Usual Care Group -Follow-up with primary care provider on as needed basis	*Duration:* 4 weeks *Timing:* PT session within 3 days of baseline assessment, 2nd session 2−3 days after 1st and 3rd/4th sessions at 1 week intervals after 2nd *Content:* spinal manipulation, range of motion, and trunk strengthening exercises	Primary *(4 weeks, 3 months, 12 months)* ODI, Secondary *(4 weeks, 3 months, 12 months)* NPRS	Significant differences at 3 months in favour of the intervention group for disability (MD −3.31 95 % CI −6.43, −0.18) and pain (MD −0.58 95 % CI −0.06, 0.09) across all participants. No differences were identified for the medium risk subgroup alone. For the high-risk subgroup, there was a significant difference between the early intervention and usual care groups for disability (MD −5.87 95 % CI −11.24, −0.50) and pain (MD −0.98 95 % CI −1.81, −0.14) at 3-months.
Hazard et al., 1997	Workers with reported workplace back injury (Vermont, USA)	-Aged 18−60 years, -Notification of workplace injury within 12 days	Not specified	*n* = 50 Intervention (*n* = 28) Control (*n* = 25)	Vermont Disability Prediction Questionnaire High risk workers included	Intervention, -Letters sent to physician with recommendations for assessment and treatment based on AHCPR algorithms* (initially and at 1 month) Control No intervention	*No detail provided*	Primary *(3 months)* Work Absence Rate Secondary *(3 months)*, Pain, healthcare satisfaction, return to work, work loss, and days until first return to work	There were no significant differences between the intervention and control groups for return to work, self-assessed pain, satisfaction with health care and days until first return to work ($p < 0.05$)

Table 1 (*Continued*)

Authors & Year	Population	Inclusion Criteria	Exclusion Criteria	Participants	Stratification Method	Group Details	Intervention Details	Outcomes	Study Results
Lee et al., 2013	Physiotherapy department of Alice Ho Mui Ling Nethersole Hospital (Hong Kong SAR, China) Female: 49 %	-Back pain <12 weeks -Injured on duty or on sick leave due to back pain -Aged 18−55 years	-Sick leave >7 days (past 12 months), -Medical consultation for musculoskeletal problem (past 12 months) -Back surgery in past 12 months -Radiculopathy -Specific diagnoses (MRI-verified disc herniation, spondylolisthesis, spinal stenosis, and inflammatory diseases) -Spinal instability -Serious spinal abnormality -Pregnancy -Contraindications to exercise -Psychiatric diagnosis -Drug abuse -Illiteracy -BMI $\geq$ 30 kg/m^2	n = 47 Intervention (n = 24) Control (n = 23)	Örebro Musculoskeletal Pain Questionnaire (OMPQ) Moderate and high risk patients (OMPQ 106−145)	Work Rehabilitation Group, -Individualized cognitive behavioural approach with graded activity, pacing, work conditioning, return-to-work goal setting, self-management strategies, job analysis, and ergonomic advice Conventional Treatment Group -Combination of interferential therapy, transcutaneous electrical nerve stimulation, lumbar traction, manual therapy, and exercise therapy	*Duration:* max of 3 months (both groups), patients were discharged program when able to return to work, had a subjective improvement of $\geq$ 70 %, or their condition plateaued *Provider:* PTs with postgraduate qualifications and training on cognitive behavioural approach *Sessions:* average of 30 min for work rehabilitation group and 15 min for conventional treatment group	Primary *(discharge)* ERQ Secondary *(discharge)* NGRCS, Pain Level (0 − 10), RMDQ, Chinese ODI, Patient's Satisfaction (-10 to +10), PSEQ. TSK-11	Significant differences favouring the intervention group for work-related recovery expectations (p = 0.002), pain self-efficacy (p = 0.035), numeric global rate of change score (p = 0.044) and patient satisfaction (p = 0.001). No differences were identified for pain level and disability.
Priebe et al., 2020	Participating GPs and via Facebook advertising (Bavaria, Germany) Female: 65 %	-Non-specific acute low back pain < 12 weeks, -<6 recurrent episodes of LBP, of < 12 weeks duration -Email/smartphone access, -German language, -Member of health insurances	-Other kinds of back pain, -Aged below 18 or above 65 years -History of back or vertebral surgery -LBP of any specific cause requiring treatment (eg fractures or tumors), -Serious medical conditions, -Psychiatric disorders	n = 941 Intervention (n = 680) Control Group (n = 261)	STarT Back Screening Tool (SBST), High risk participants received a teleconsultation from a pain specialist	Rise-uP Group, -Classified based on StarT Back scores, -*High risk patients' GPs* received a pain specialist teleconsultation, -Patients used Kaia back pain app (involves education, exercises and mindfulness/relaxation) Usual Care -Standard of care treatment based on German national guideline	Not specified; participant dependent on usage of app	Primary *(3 months)* Pain Intensity (NPRS 0−10 for current pain, maximum pain over last 4 weeks and average pain over last 4 weeks), Secondary *(3 months)*, DASS, HFAQ, VR-12	Significant differences favouring the intervention group for pain (p < 0.001), anxiety, depression, stress and disability (p < 0.001) at 3 months.

LBP, low back pain; PIPT, Psychologically Informed Physical Therapy; PT, physical therapist/physiotherapy, ODI, Oswestry Disability Index; NPRS, Numeric Pain Rating Scale; AHCPR, Agency for Health Care Policy and Research; OMPQ, Örebro Musculoskeletal Pain Questionnaire; ERQ, Work Related Recovery Expectations Questionnaire; NGRCS, Numeric Global Rate of Change Score; RMDQ, Roland Morris Disability Questionnaire; PSEQ, Pain Self-Efficacy Questionnaire; TSK-11, Tampa Scale for Kinesiophobia; DASS, Depression-Anxiety-Stress Scale; HFAQ, Hannover Functional Ability Questionnaire; VR-12, Veterans RAND 12 item Health Survey.

[*] AHCPR algorithms includes screening for red flags, observation, neurological testing, neurological tension tests, education and assurance, patient comfort, oral pharmaceuticals, manipulation, traction, massage, diathermy, ultrasound, cutaneous laser treatment, biofeedback and TENS, needle acupuncture and injections ERQ.

Table 2 Risk of bias using PEDro Scale.

Study	Score/10[1]	1. Eligibility criteria	2. Random allocation	3. Concealed allocation	4. Baseline comparability	5. Blinding of subjects	6. Blinding of therapists	7. Blinding of assessors	8. Adequate follow-up	9. Intention to treat	10. Bewteen-group comparisons	11. Point estimates and variability
Delitto et al., 2021	6	Yes	Yes	Yes	Yes	No	No	No	No	Yes	Yes	Yes
Hazard et al., 1997	4	Yes	Yes	No[2]	No	No	No	No	Yes	No[2]	Yes	Yes
Priebe et al., 2020	3	Yes	Yes	No	No	No	No	No	No	No	Yes	Yes
Magel et al., 2017	7	Yes	Yes	Yes	Yes	No	No	No[3]	Yes	Yes	Yes	Yes
Lee et al., 2013	7	Yes	Yes	Yes	Yes	No	No	No	Yes[4]	Yes	Yes	Yes

[1] eligibility criteria not included in total score.
[2] unclear from article.
[3] score differs from that of PEDro database; both outcome measures were considered to rely on self-reported information. Therefore, the participants were also considered to be the assessors and as the participants were not blinded, by extension the assessors were also considered to not be blinded.
[4] score differs from that of PEDro database; Table 3 appears to indicate that outcomes were obtained from all participants ($n = 47$) reported to be initially allocated to each group.

Table 3 Certainty of evidence assessment (GRADE analysis).

Outcome	No of participants (cohorts)	Design	Effect	Study limitations	Inconsistency	Indirectness	Imprecision	Publication bias	Overall GRADE
Disability	3469 (5)	RCT	−1.52 (−4.15, 1.11)	Serious [a]	Serious [b]	Serious [c]	Not serious	Suspected [d]	Very low ⊕○○○
Pain	1169 (4)	RCT	−0.62 (−0.88, −0.36)	Very serious [e]	Not serious	Serious [c]	Serious [f]	Suspected [d]	Very low ⊕○○○

Effect expressed as mean difference (95 % confidence interval).
[a] >25 % of participants were from studies at high risk of bias (27 %)
[b] statistical heterogeneity (I^2) >40 % ($I^2 = 55$ %)
[c] variations in nature of included interventions
[d] <10 studies included in analysis
[e] >50 % for participants from studies at high risk of bias (80 %)
[f] fewer than 2000 participants

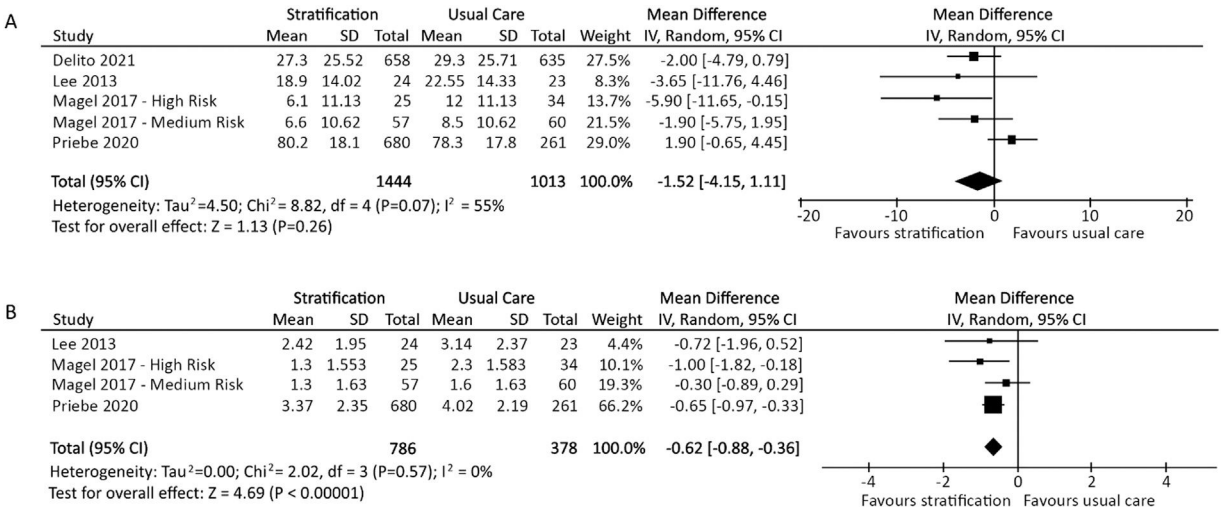

Fig. 2 Forest Plot A) Mean Difference: Disability at Intermediate Term (3–6 months) B) Mean Difference: Pain at Intermediate Term (3–6 months).

Results of syntheses

Risk stratification compared with usual care

The results of the meta-analysis demonstrate a small but significant effect favouring risk stratification over usual care for pain at 3–6 month follow-up (MD= −0.62, 95% CI: −0.88, −0.36). The upper limits of the confidence intervals were well within the pre-determined MCID of 2 out of 10 points, suggesting that while the effect was statistically significant, it was not clinically significant. There was no difference between risk stratification and usual care for disability (MD= −1.52, 95% CI: −4.15, 1.11).

Risk stratification compared with no intervention

Hazard et al[39] was the only study to compare risk stratification to no intervention. This study found no significant differences between control and intervention groups. However, it is worth noting that it did not publish, nor provide, on request, raw data relating to participant outcome measures.

Certainty of evidence

Using the GRADE approach, each meta-analysis was assessed as having very-low certainty. Refer to Table 3 for complete details of the GRADE outcome for each analysis.

Sensitivity analysis was completed by conducting repeat meta-analysis with small methodological alterations. These alterations included removing Priebe et al[40] due to its low methodological quality, and repeating meta-analysis using both a random effects model and standardized mean difference. With the removal of Priebe et al[40] from the meta-analysis relating to disability, the results demonstrated a statistically significant difference favouring risk stratification (MD= −0.11, 95% CI: −0.22, −0.00). This contrasts with the original findings. No other sensitivity analysis generated contrasting findings.

Discussion

Summary

This review found very-low certainty evidence that the use of a risk stratification approach with matched treatment for patients with acute back pain may lead to a very small reduction in pain levels at 3–6 months compared with usual care. This result, however, did not achieve the pre-determined clinically significant threshold.

This review also found very-low certainty evidence that risk stratification has no significant difference on disability outcomes. While the removal of Priebe et al[40] during sensitivity analysis resulted in contrasting findings to the original analysis, these results would fall well short of the proposed MCID.[28] Hence, overall, the use of risk stratification for adults with acute back pain may not be superior when compared with usual care, but the evidence is very uncertain. Due to the very-low certainty of evidence, future, high-quality, research may significantly change the conclusions of this review and is required to confirm the utility of risk stratification in the management of acute back pain.

Limitations

There was a focus on participants at high risk of developing chronic pain in relation to their inclusion in each trial and the provision of interventions. Priebe[40] was the only study to provide a different intervention for different levels of risk, including consultation with a pain specialist for those at high-risk. All other studies matched interventions for participants stratified as high-risk only (and medium-risk for Magel et al[38]), rather than providing different interventions to those with different levels of risk. It is possible that different conclusions could have been drawn if interventions were also developed and prescribed for those in the low and medium risk groups.

Furthermore, no study fully complied with the suggested management for each level of risk, as recommended by the

authors of the STarT Back Screening Tool. The only studies which followed the recommendations for high-risk patients were Delitto et al,[36] which applied a psychologically-informed intervention, and Lee et al,[35] which provided an intervention based on a cognitive-behavioral approach. This lack of concordance between the care provided and suggested management for each level of risk may have impacted the results of the studies as the care provided in each study may have been less effective. Studies did not publish specific detail regarding the nature of the usual care intervention provided, therefore it is unknown whether the usual care interventions were truly comparable. It is reasonable to assume that, given the regional variation of where the studies were conducted, usual care interventions would have differed between each study.

The inclusion criteria of the individual studies may have also limited the generalizability and practicality of conclusions. All included studies took place within high-income countries, limiting generalizability to low- and middle-income countries. Lee et al[35] and Magel et al,[38] excluded participants with pain extending below the knees and with neurological symptoms, and the former only included participants aged 18 to 55 who had been injured at work. Magel et al[38] was a secondary analysis of an RCT, and as such, the original study was not designed to evaluate the use of a risk stratification tool.

In reference to the design of this review, the exclusion of non-RCTs from the search strategy likely limited the yield of evidence. Although, to our knowledge, no other non-RCT has evaluated risk stratification specifically for acute back pain. Furthermore, the limitation of studies to only the English language may have resulted in non-English studies being missed; limiting the diversity and generalizability of findings. Additionally, there was likely a publication bias in both meta-analyses due to the small number of cohorts included.

Comparison with other research

The results of this review contrast those of a recent systematic review by Ogbeivor and Elsabbagh[17] which found evidence to support the use of stratified management with matched treatment for back pain of any duration. It is worth noting that Ogbeivor and Elsabbagh[17] differed from this review as it included participants with chronic pain and non-RCTs, and excluded patients stratified using tools other than the STarT Back Screening Tool. In reference to the management of acute back pain more broadly, there is high-quality evidence to support the use of exercise, heat, manual therapy, and education.[42,43]

Implications

While risk stratification is proposed as a method of managing acute back pain, this review found only very-low certainty evidence of improvement in pain at 3–6 months post intervention compared with usual care, and this finding did not reach clinical significance. The very-low certainty results were due to a general paucity of high-quality studies specific to only acute back pain. Future high-quality RCTs are needed to conclusively determine the effectiveness of risk stratification for patients with acute back pain, and outcomes should include the transition from acute to chronic pain.

Conclusions

With chronic pain becoming increasingly prevalent worldwide, preventative interventions are required. Findings from this systematic review indicate that the use of risk stratification with matched treatment may be just as worthwhile as usual care for acute back pain, however the evidence is very uncertain. Further research is required to confirm whether risk stratification with matched treatment is a worthwhile approach for managing acute back pain.

Conflicts of interest

The authors declare no conflicts of interest.

Acknowledgements

The authors would like to thank Professor Nicholas Taylor, Annie Lewis, Dr Katherine Harding, Sue Parslow, Leanne Roddy and the Eastern Health Library Service.

This research did not receive any specific grant from funding agencies in the public, commercial, or not-for-profit sectors

Supplementary materials

Supplementary material associated with this article can be found in the online version at doi:10.1016/j.bjpt.2024.101116.

References

1. Australian Institute of Health and Welfare. *Back problems*. 2019. https://www.aihw.gov.au/reports/chronic-musculoskeletal-conditions/back-problems.
2. LdCM Costa, Maher CG, Hancock MJ, McAuley JH, Herbert RD, Costa LOP. The prognosis of acute and persistent low-back pain: a meta-analysis. *Can Med Assoc J*. 2012;184(11):E613–E624.
3. Stevans JM, Delitto A, Khoja SS, et al. Risk factors associated with transition from acute to chronic low back pain in US patients seeking primary care. *JAMA Network Open*. 2021;4(2):e2037371.
4. Chou R, Shekelle P. Will this patient develop persistent disabling low back pain? *JAMA*. 2010;303(13):1295–1302.
5. Chou R, Deyo R, Friedly J, et al. Nonpharmacologic therapies for low back pain: a systematic review for an American College of Physicians Clinical Practice Guideline. *Ann Intern Med*. 2017;166(7):493–505.
6. GBD 2019 Risk Factor Collaborators. Global burden of 87 risk factors in 204 countries and territories, 1990–2019: a systematic analysis for the Global Burden of Disease Study 2019. *Lancet*. 2020;396(10258):1223–1249.
7. Back Pain Clinical Care Standard. 2022. https://www.safetyandquality.gov.au/sites/default/files/2022-08/low_back_pain_clinical_care_standard.pdf.
8. National Institute for Health and Care Excellence. Low back pain and sciatica in over 16s: assessment and management [NICE Guideline No. 59]. https://www.nice.org.uk/guidance/ng59.
9. Van Wambeke P., Desomer A., Ailliet I., et al. Low Back Pain and Radicular Pain: Assessment and Management. 2017.

10. 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(9802):1560—1571.
11. 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(5):632—641.
12. Linton SJ, Nicholas M, MacDonald S. Development of a short form of the Örebro Musculoskeletal Pain Screening Questionnaire. *Spine*. 2011;36(22):1891—1895.
13. Linton SJ, Halldén K. Can we screen for problematic back pain? A screening questionnaire for predicting outcome in acute and subacute back pain. *Clin J Pain*. Sep 1998;14(3):209—215.
14. Hill JC, Dunn KM, Main CJ, Hay EM. Subgrouping low back pain: a comparison of the STarT Back Tool with the Orebro Musculoskeletal Pain Screening Questionnaire. *Eur J Pain*. 2010;14(1):83—89.
15. Karran EL, McAuley JH, Traegar AC, et al. Can screening instruments accurately determine poor outcome risk in adults with recent onset low back pain? A systematic review and meta-analysis. *BMC Med*. 2017;15(1):13.
16. Meyer C, Denis CM, Berquin AD. Secondary prevention of chronic musculoskeletal pain: a systematic review of clinical trials. *Ann Phys Rehabil Med*. 2018;61(5):323—338.
17. Ogbeivor C, Elsabbagh L. Management approach combining prognostic screening and targeted treatment for patients with low back pain compared with standard physiotherapy: a systematic review & meta-analysis. *Musculoskeletal Care*. 2021;19 (4):436—456.
18. Page MJ, McKenzie JE, Bossuyt PM, et al. The PRISMA 2020 statement: an updated guideline for reporting systematic reviews. *BMJ*. 2021;372:n71.
19. Merskey H. Pain terms: a list with definitions and notes on usage recommended by the IASP Subcommittee on Taxonomy. *Pain*. 1979;6:249—252.
20. Merskey H., Bogduk N. *Classification of Chronic pain. Descriptions of Chronic Pain Syndromes and Definitions of Pain Terms* (2nd ed). IASP Press; 1994.
21. National Health and Medical Research Council. NHMRC levels of evidence and grades for recommendations for developers of guidelines. 2009.
22. Charrois TL. Systematic reviews: what do you need to know to get started? *Can J Hosp Pharm*. 2015;68(2):144—148.
23. Covidence Veritas Health Innovation www.covidence.org.
24. Higgins J.P., Thomas J., Chandler J., et al. Higgins JPTTJ, Chandler J, Cumpston M, Li T, Page MJ, Welch VA (editors), ed. *Chapter 6: Choosing effect measures and computing estimates of effect*. Cochrane; 2022. www.training.cochrane.org/handbook.
25. de Morton NA. The PEDro scale is a valid measure of the methodological quality of clinical trials: a demographic study. *Aust J Physiother*. 2009;55(2):129—133.
26. *Review Manager (RevMan)*. 2020.
27. Higgins JP, Thompson SG. Quantifying heterogeneity in a meta-analysis. *Stat Med*. 2002;21(11):1539—1558.
28. Ostelo RW, Deyo RA, Stratford P, et al. Interpreting change scores for pain and functional status in low back pain: towards international consensus regarding minimal important change. *Spine*. 2008;33(1):90—94.
29. Guyatt GH, Oxman AD, Vist GE, et al. GRADE: an emerging consensus on rating quality of evidence and strength of recommendations. *BMJ*. 2008;336(7650):924—926.
30. Guyatt GH, Oxman AD, Kunz R, et al. GRADE guidelines 6. Rating the quality of evidence—imprecision. *J Clin Epidemiol*. 2011;64 (12):1283—1293.
31. Guyatt GH, Oxman AD, Kunz R, et al. GRADE guidelines: 8. Rating the quality of evidence—indirectness. *J Clin Epidemiol*. 2011;64 (12):1303—1310.
32. Guyatt GH, Oxman AD, Kunz R, et al. GRADE guidelines: 7. Rating the quality of evidence—inconsistency. *J Clin Epidemiol*. 2011;64(12):1294—1302.
33. Guyatt GH, Oxman AD, Montori V, et al. GRADE guidelines: 5. Rating the quality of evidence—publication bias. *J Clin Epidemiol*. 2011;64(12):1277—1282.
34. Guyatt GH, Oxman AD, Vist G, et al. GRADE guidelines: 4. Rating the quality of evidence—study limitations (risk of bias). *J Clin Epidemiol*. 2011;64(4):407—415.
35. Lee WYA, Lee WCE, Law SW, et al. Managing psychosocial contributors in low back pain patients-a randomised controlled trial. *J Orthopaed, Trauma Rehabil*. 2013;17 (1):46—51.
36. Delitto A, Patterson CG, Stevans JM, et al. Stratified care to prevent chronic low back pain in high-risk patients: the TARGET trial. A multi-site pragmatic cluster randomized trial. *eClinicalMedicine*. 2021.. 34((Delitto) School of Health and Rehabilitation Sciences, University of Pittsburgh, 4029 Forbes Tower, Pittsburgh 15260, United States(Patterson, Stevans, Freburger, Khoja, Schneider) School of Health and Rehabilitation Sciences, University of Pittsburgh, 1):100795; doi:10.1016/j.eclinm.2021.100795.
37. The World Bank Group. High income. https://data.worldbank.org/country/XD.
38. Magel J, Fritz JM, Greene T, Kjaer P, Marcus RL, Brennan GP. Outcomes of patients with acute low back pain stratified by the STarT Back Screening Tool: secondary analysis of a randomized trial. *Phys Ther*. 2017;97(3):330—337.
39. Hazard RG, Haugh LD, Reid S, McFarlane G, MacDonald L. Early physician notification of patient disability risk and clinical guidelines after low back injury: a randomized, controlled trial. *Spine*. 1997;22(24):2951—2958.
40. Priebe JA, Haas KK, Sanchez LFM, et al. Digital treatment of back pain versus standard of care: the cluster-randomized controlled trial, rise-up. *J Pain Res*. 2020;13:1823—1838.
41. Cashin AG, McAuley JH. Clinimetrics: Physiotherapy Evidence Database (PEDro) scale. *J Physiother*. 2020;66(1):59.
42. Gianola S, Bargeri S, Castillo GD, et al. Effectiveness of treatments for acute and subacute mechanical non-specific low back pain: a systematic review with network meta-analysis. *Br J Sports Med*. 2022;56(1):41—50.
43. George SZ, Fritz JM, Silfies SP, et al. Interventions for the management of acute and chronic low back pain: revision 2021. *J Orthopaed Sports Phys Ther*. 2021;51(11): CPG1—CPG60.