Running is associated with intervertebral disc adaptations: a pre-planned secondary analysis of the ASTEROID randomised controlled trial

European Spine Journal (2026) 35:3822–3838
https://doi.org/10.1007/s00586-026-09759-7

RESEARCH

Running is associated with intervertebral disc adaptations: a pre-planned secondary analysis of the ASTEROID randomised controlled trial

Claire L Samanna[1,2] · Christopher Neason[1,2] · Scott D Tagliaferri[3,4] · Daniel L Belavý[5] · Ulrike H Mitchell[5] · Hanna Rae Nez[6] · Paul Buntine[7,8] · Clint T Miller[9] · David Scott[8,9] · Niamh L Mundell[9] · Matthew J Clarkson[10] · Patrick J Owen[1,7]

Received: 13 October 2025 / Revised: 14 December 2025 / Accepted: 21 January 2026 / Published online: 5 March 2026

Abstract

Purpose To identify patient and intervention factors that moderate the effects of running on intervertebral disc health (IVD) in adults with non-specific chronic low back pain.

Methods Pre-planned secondary analysis of a 12-week parallel-group (1:1) randomised controlled trial of 40 adults with non-specific CLBP (mean [SD] age: 33 [6] years, female: 50%). Participants were allocated to a digitally-delivered progressive run-walk interval exercise training program (3 days/week, 30 min/session) or waitlist control. Magnetic resonance imaging at baseline, six, and 12 weeks quantified primary outcome IVD composition as whole-disc T2 (ms) from T11/T12 to L5/S1. Moderators included baseline degeneration (Pfirrmann grade; aggregated Pfirrmann score), sex, body mass index, and intervention factors (cumulative running volume, mean speed, and dominant surface). Linear mixed models were fitted at the IVD level to estimate group-by-time effects within each moderator-defined subgroup ($\alpha = 0.05$).

Results The highest aggregated Pfirrmann scores reflecting greater multi-level degeneration (estimated marginal mean net difference [95%CI]: 3.42 [0.22, 6.62] ms, $P=0.036$), cumulative running volumes between 28.6 and 46.1 kms (4.80 [2.51, 7.10] ms, $P<0.001$), mean running speeds between 10.5 and 11.7 km/h (2.46 [0.14, 4.79] ms, $P=0.038$), and running on grass (2.93 [0.10, 5.76] ms, $P=0.043$) positively moderated between-group IVD T2 changes.

Conclusion In this pre-planned secondary analysis of a randomised control trial of a run-walk program, our data suggest running speed, volume and surface, as well as more multi-level IVD degeneration, may be associated with more favourable 12-week IVD T2 changes. Findings are hypothesis-generating and may inform future trials designed to optimise IVD health.

Keywords Spine · Moderator · Magnetic resonance imaging · jogging · disc degeneration

Introduction

Individuals with chronic low back pain (CLBP) exhibit a higher prevalence of intervertebral disc (IVD) degeneration on magnetic resonance images than asymptomatic peers [1, 2]. Interventions targeting IVD degeneration pose a risk of adverse events [3] and repeat procedures [4]; however, improving IVD health may alleviate symptoms in some individuals [5]. Here, we define IVD health as the structural and functional integrity of the IVD, characterised by preserved hydration, disc height, and collagen composition, enabling normal load distribution [6]. Cell [7] and animal [8, 9] models suggest loading via mechanotransduction (thereafter referred to as physical loading) can impact IVD properties. A recent systematic review and meta-analysis revealed running was associated with better IVD health in cross-sectional studies. To date, two randomised controlled trials have investigated the effects of physical loading via exercise training on IVD health. Both were measured via magnetic resonance imaging and conducted in individuals with CLBP [10, 11]. The first trial was a 6-month strength and conditioning intervention compared to control [11].

Extended author information available on the last page of the article

However, there was suboptimal intervention adherence (60%), and the multimodal exercise intervention may have been too variable to support sufficient IVD loading [11]. More recently, authors conducted a 12-week run-walk intervention (ASTEROID) that led to between-group improvements in low back pain (LBP) and disability [12]. Exploring potential moderating factors influencing IVD adaptations observed in this trial may inform future studies designed to promote IVD health.

Patient factors may moderate the IVD response to physical loading. For example, the higher prevalence of degeneration in individuals with CLBP may, in part, explain the lack of changes in the run-walk intervention. Whether degenerated IVDs respond to dynamic loading similar to healthy IVDs reveals mixed evidence according to cell-based studies [13, 14]. However, it is unclear if the same relationship applies in vivo to human degenerated IVDs. Further, the influence of body mass index (BMI) warrants consideration. Higher BMI was associated with greater magnetic resonance imaging-derived reductions in IVD height after 30 min of treadmill walking in eight asymptomatic individuals [15]. Nonetheless, the influence of BMI on IVD physical loading across multiple exercise sessions remains unexplored. Finally, sex differences warrant consideration. A validated sex-specific biomechanical spinal model demonstrated that, under identical loading conditions, female IVDs experienced greater stress and intersegmental rotation than male IVDs [16]. Further, at matched running speeds, females experience greater vertical ground reaction forces than males, which may also influence IVD loading [17]. Therefore, the moderating influence of IVD degeneration status, BMI, and sex warrants investigation.

Adaptation of the IVD occurs through physical loading, which originates in part from ground reaction forces transmitted through the lower limbs. These forces lead to cellular responses that regulate IVD matrix synthesis and hydration [7]. Exercise training factors moderate the way physical loads are absorbed and distributed at the lumbar spine [18, 19]. For example, higher running cadence attenuates spinal loading compared to a lower cadence [18] and running on a concrete surface is associated with greater ground reaction forces [20]. The ASTEROID trial measured running speed, distance, and surface; therefore, another key question remaining is the potential moderating effect of these exercise training factors on IVD outcomes. Given the novelty of running among participants with CLBP, the run-walk intervention was designed to be conservative, with the aim of determining trial feasibility, which was one of the primary outcomes [12]. This resulted in wide variation in mean running speeds from 4 to 11 km/h and cumulative running volumes from 2 to 110 km over 12 weeks, providing an opportunity for subgroup analysis. Exploring the moderating effect of exercise intervention factors has marked potential to inform future trials.

Magnetic resonance imaging-derived IVD T2 is a highly precise measure of hydration and IVD composition [21]. Lower T2 values indicate greater degeneration [22] due to reduced hydration and increased fibrous collagen content [23]. Therefore, understanding the moderating effects of patient and exercise training intervention factors by exploring patient subgroups will inform future exercise interventions designed to improve IVD T2. To this end, in participants with CLBP enrolled in a 12-week run-walk interval training program, we examined whether patient factors moderated primary outcome IVD T2 changes across spinal segments T11/T12 to L5/S1. These included baseline degeneration (Pfirrmann grade and aggregated Pfirrmann grade), BMI and sex. Additionally, we examined the moderating effect of intervention factors, including cumulative running volume, mean running speed, and running surface, on IVD T2 changes in the intervention group compared to the complete control sample. These findings will inform the efficacy of physical loading for already degenerated IVDs and provide insight into appropriate exercise training factors for promoting IVD health.

Methods

Trial design

This pre-planned secondary analysis of a prospectively registered single-blinded parallel 12-week randomised controlled trial (Australian New Zealand Clinical Trials Registry ACTRN12622001276741, date registered: 29/09/2022) examined the efficacy of an exercise intervention for IVD health compared to a waitlist control in participants with CLBP. The intervention effect on IVD health was hypothesised to vary by baseline degeneration and running volume. The trial was conducted from December 2022 to May 2023 with a total sample size of 40 patients (1:1 allocation ratio). Magnetic resonance imaging data were collected at a radiological imaging facility in Melbourne (Imaging @ Olympic Park), and the intervention was digitally delivered. The study was approved by the Deakin University Human Research Ethics Committee review board (ID: 2022-162, date registered: 29/09/2022) and reported in accordance with the CONSORT 2010 checklist (Supplement A).

Participants

Forty participants aged 18–45 years with non-specific CLBP (no definitive underlying pathology with pain on most days of the week for ≥ 3 months) were recruited via social media

from Victoria, Australia. Exclusion criteria included: (1) specific LBP diagnosis (fracture, cauda equina syndrome, cancer, spondyloarthropathies, infection); (2) previous spinal surgery or spine trauma (car accident or fall resulting in LBP); (3) structural scoliosis that required surgical consultation; (4) current radiculopathy symptoms (leg pain greater than LBP); (5) non-English speaking; (6) recent pregnancy or lactation (< 1 year) [24]; (7) past or current elite athlete (member of the National or State Institutes of Sports or national sport team) [25]; (8) absolute contraindications for exercise training [26] or high risk to exercise (Adult Pre-Exercise Screening System) [27], (9) recent running or sport involving running (> 1 per month in the last three months (10); recent lower limb injury (< 6 weeks) [28], and (11) no smartphone with cellular internet connection. All participants provided informed written consent before participation.

Intervention and control

The exercise training intervention was a pragmatic, digitally delivered run-walk interval program with progressively increasing running intervals over 12 weeks. After a two-minute test run, participants could self-select to start at stage one, two, or three, corresponding to intervals of 15, 30, or 45 s of running, followed by two minutes of walking (Supplement B). Participants pre-programmed the intervals into the Runkeeper mobile application (V.14.7, Anima Sana In Corpore Sano [ASICS] Runner App, Boston, USA). Participants were asked to complete three sessions per week in their local area, remain at the same level for at least one week, run at a slow, comfortable speed (approximately < 10 km/h), and complete between 6 and 10 repetitions per session. Using secure video conference (Zoom Video Communications, California, United States of America), participants consulted with a clinical exercise physiologist (weekly and fortnightly over the first 4 and 12 weeks, respectively) to discuss progress and plan for the subsequent week (whether to progress, regress, or stay at the same stage). All participants were advised to continue managing their LBP as usual (e.g., general practitioner consultation, over-the-counter pharmacotherapy, or manual therapy). Control participants were not prescribed an exercise program and were asked not to engage in any activity involving running. Control participants were offered the same 12-week exercise intervention with exercise physiology support after testing at 12 weeks.

Outcomes

Magnetic resonance imaging [29] was completed at baseline, six weeks, and 12 weeks, and running metric data were digitally collected weekly through a smart-phone application (Runkeeper). Electronic questionnaires via REDCap electronic data capture tools hosted at Deakin University (v13.8.2) [29, 30] were administered prior to the magnetic resonance imaging testing session, which captured participant demographics, habitual physical activity (International physical activity questionnaire) [31], average pain intensity (100-point visual analogue scale) [32], and self-reported disability (100-point Oswestry Disability Index) [33]. Height and weight were also measured at baseline.

Magnetic resonance imaging and blinded analysis

All scans were conducted using a 3 T Phillips scanner (Best, Netherlands; software release 5.7 2021-10−04) after midday to avoid diurnal variation [34]. Participants were asked to refrain from physical activity on testing days [35] and were required to sit for 20 min immediately before the scan. The scans used posterior coil, torso coil array anteriorly, and spinal coils capturing the entire lumbar spine from left to right with two magnetic resonance sequences, including sagittal spin-echo multi-echo (eight echo times; 15.75, 36.75, 57.75, 78.75, 99.75, 120.75, 141.75, and 162.75 ms) and T2 Sag spin-echo (one echo-time; 70.00 ms). Both scans are described in further detail in Supplement C.

IVD T2

Sagittal spin-echo multi-echo sequence data were processed in ImageJ using established methods [21]. Before processing, all images were blinded by assigning a randomly generated code to each participant scan, and one researcher (CLS) traced five pilot scans for review by a researcher experienced in tracing methods (SDT). The middle slice was identified via the location of the spinous process, and the region of interest was manually traced around the borders of the IVD without the endplate (Fig. 1). A custom-made plugin ("ROI Analyzer"; https://github.com/tjrantal/RoiAnalyzer and https://sites.google.com/site/daniellbelavy/home/roianalyser) replicated the region of interest across all echo times, which calculated the average IVD signal intensity at each echo. This process was completed for all available slices along the frontal axis, where the whole IVD was visible. Whole IVD T2 was calculated using R statistical software via the linear fit to the natural logarithm of signal intensity at each echo, which provided a quantitative value from the magnetic resonance relaxation times. The T2 value is influenced by IVD biochemical composition and water content, where higher values indicate higher hydration and better IVD health, which has demonstrated excellent reliability (ICC = 0.98) [21].

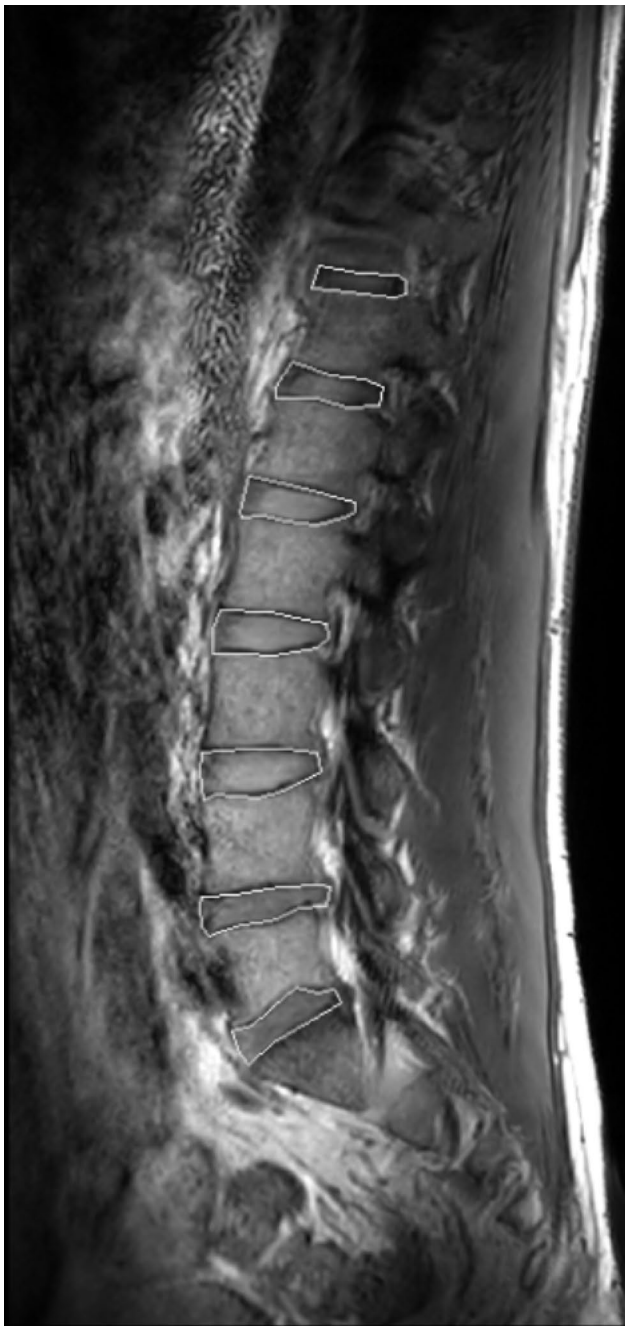

Fig. 1 One slice of traced intervertebral discs (IVD) using ImageJ software on lumbar spine magnetic resonance images

Pfirrmann grade

T2 Sag spin-echo sequence data were visually inspected and graded according to Pfirrmann grade from one (healthy) to five (severe degeneration), reflecting changes in IVD structure, nucleus to annulus distinction, signal intensity brightness, and height (Table 1) [36]. All images were blinded as described previously and were assessed at the middle slice after identifying the spinous process in duplicate by two independent clinical researchers (UHM, HRN) who had completed formal Pfirrmann grading training. All conflicts were resolved via discussion. For the analysis, baseline IVDs were classified into subgroups according to individual Pfirrmann grades and aggregated Pfirrmann score across spinal levels T11/T12 to L5/S1.

Table 1 Pfirrmann grade classification on magnetic resonance imaging, modified by Pfirrmann [36]

Grade	Classification
I	Homogeneous structure, distinct nucleus and annulus, with bright hyperintense white signal and a normal disc height.
II	Inhomogeneous structure with or without horizontal grey bands, distinct nucleus and annulus, with bright hyperintense white signal and normal disc height.
III	Inhomogeneous grey structure, less distinct nucleus and annulus, intermediate signal and disc height is normal or slightly decreased.
IV	Inhomogeneous grey to black structure, no nucleus and annulus distinction, intermediate to hypointense signal and disc height is normal or moderately decreased.
V	Inhomogeneous black structure, no nucleus and annulus distinction, hypointense signal and collapsed disc space.

Body mass index

BMI was measured at the baseline testing session using standard height and weight measurements (stadiometer and scales). Participants were asked to remove footwear before testing. BMI was calculated at a later date as weight in kilograms divided by height in metres, squared (m^2).

Running moderators

Participants subjectively reported running surface during Zoom consultations as either grass, gravel, pavement, or trail. The dominant running surface was used for subgroup stratification. Running interval distance and speed were tracked via the Runkeeper application using Global Positioning System tracking. The cumulative running volume was calculated by summing the distance (metres) for each running interval. The mean running speed (km/h) was calculated as the average across all sessions over 12 weeks. Participants downloaded and accessed the free Runkeeper application using a generic email, for data security purposes, via their smartphone, and pre-programmed the interval times for the corresponding running stage. The Runkeeper application has been validated in the literature [37] using a Global Positioning System monitor and Garmin Forerunner 310XT watch (Garmin Ltd., Olathe, Kansas), demonstrating 96.8% accuracy in calculating running distance [38]. Cumulative running volume and mean running speed were stratified into IVD quartiles for analysis from baseline to

12 weeks. All running moderators were compared with the whole control sample.

Sample size

This sample size was based on calculations for the primary outcomes of IVD health, pain intensity, and disability [12] as per the published protocol [10].

Randomisation

Block randomisation with random block lengths (between two and six per block), developed using the 'blockrand' package in R (v4.1.2) [39], randomly assigned participants (1:1) to a waitlist control or an intervention with equal female and male participants. A member of the research team with no participant contact (SDT) employed the randomisation schedule, which involved constructing sequentially numbered, opaque, and sealed envelopes stratified by sex.

Blinding

Magnetic resonance technicians were blinded to treatment allocation and time point. Random codes were assigned to all scans, separate from the study ID, to blind the magnetic resonance image tracer (CLS) and Pfirrmann grade assessors (UHM, HRN) to treatment allocation and timepoint. Blinding was not possible for participants or exercise physiologist researchers administering the intervention, given the study design, which compared an exercise intervention to a waitlist control.

Statistical methods

All analyses were conducted using Stata (v17, StataCorp, College Station, Texas, United States of America). An intention-to-treat approach [40] using separate linear mixed models across subgroups, with random effects for participants and lumbar level, evaluated within- and between-group changes in IVD T2 over time. All linear mixed models employed restricted maximum likelihood estimations. Subgroups were classified by: (1) baseline Pfirrmann grade (normal, mild, moderate, severe, advanced) (2) aggregated Pfirrmann score from T11/T12 to L5/S1 (stratified in quartiles; range: 16–27 points), (3) sex (male, female), (4) baseline body mass index (stratified in quartiles; range: 19.3–53.4 kg/m^2), (5) running volume (stratified in quartiles; range: 1.8–109.8 km), (6) mean running speed (stratified in quartiles; range: 4.4–11.7 km/h), and (7) running surface (grass, gravel, pavement, trail). For each analysis of exercise intervention factors (running volume, speed, and surface), the whole control sample was used as the comparison group. Sensitivity analyses were performed by adjusting primary models for individual baseline characteristics. An α of 0.05 was adopted for all analyses. The statistical analysis and reporting are consistent with the Checklist for Statistical Assessment of Medical Papers (CHAMP) statement [41].

Results

Participant flow and recruitment

Participant recruitment (total $n=40$; female: $n=20$, male: $n=20$) was conducted from October 2022 to January 2023 (Fig. 2), and the study was conducted from December 2022 to May 2023.

Baseline and summary data

Descriptive characteristics at baseline were reported in Table 2. As reported elsewhere [12], mean (SD) intervention adherence was 70% (20%) of prescribed sessions. Mean (SD) Pfirrmann grade at baseline was 3.1 (0.7) and 2.94 (0.6) points in the intervention and control group, respectively. Baseline aggregated Pfirrmann score ranged from 18 to 24 and 16–27 in the intervention and control groups, respectively. In the intervention group, the most common running surface was pavement ($n = 14$). Mean (SD) running volume (excluding walking) from baseline to 12 weeks was 49.7 (30.8) km, which ranged from 1.8 to 109.8 km over the 12 weeks. Mean (SD) running speed from baseline to 12 weeks was 9.5 (1.7) km/h, which ranged from 4.4 to 11.7 km/h over the 12 weeks.

Numbers analysed

All participant data ($n = 40$) were analysed with an intention-to-treat approach as set a priori [10]. One control participant had missing magnetic resonance data at six and 12 weeks, however their baseline data were included in the analysis in line with an intention-to-treat approach.

Outcomes and Estimation

Patient factors

All IVD T2 changes stratified by baseline Pfirrmann grade are reported in Table 3; Fig. 3 and Supplement D. When IVDs were stratified by baseline Pfirrmann grade, IVDs with mild degeneration (Pfirrmann two) negatively moderated between-group IVD T2 changes at six weeks only. No other

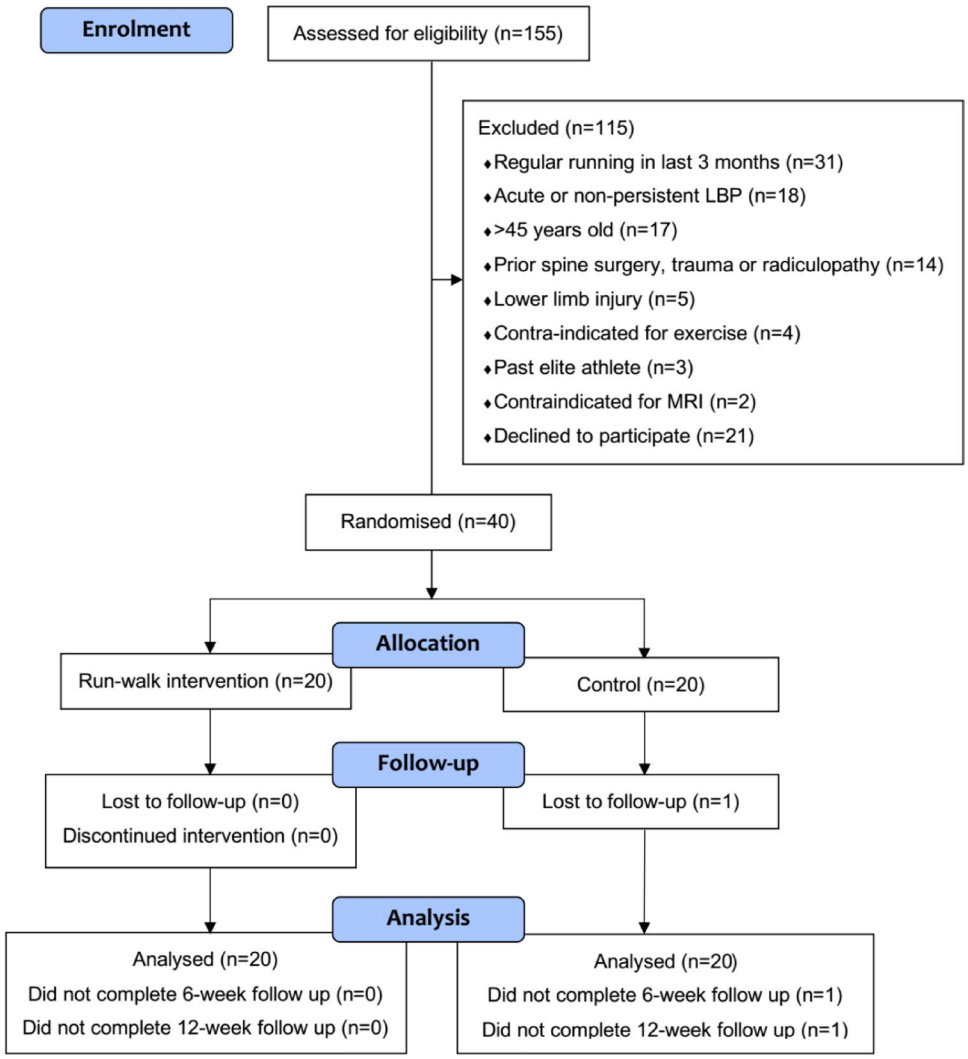

Fig. 2 Consort diagram of study flow

baseline Pfirrmann grades moderated between-group differences in IVD T2. Participant baseline aggregated Pfirrmann grade score (aggregated Pfirrmann) was stratified into quartiles, yielding four groups that represent increasing levels of degeneration across the spinal levels T11/T12 to L5/S1 (Table 3 and Supplement D). At six weeks, aggregated Pfirrmann < 21 (quartile one) negatively moderated between-group IVD T2. At 12 weeks, aggregated Pfirrmann ≥ 23 (quartile four) positively moderated between-group IVD T2. No other quartile or time point revealed between-group changes.

IVD T2 changes stratified by sex are reported in Table 4 and Supplement D. Being male negatively moderated IVD T2 changes at six weeks and positively moderated IVD T2 changes at 12 weeks compared to control. When all participants were stratified into quartiles based on their baseline BMI, no group moderated the between-group changes in IVD T2 (Table 4 and Supplement D).

Intervention factors

The cumulative run volume distance of the intervention group was stratified into quartiles, and IVD T2 changes were compared with the entire control sample (Table 5 and Supplement D). Cumulative running volumes between 28.6 and 46.1 km positively moderated between-group IVD T2 at 12 weeks (Table 5 and Supplement D). Whereas cumulative running volumes less than 28.6 km or between 46.1 and 81.2 km, negatively moderated IVD T2 between-group, however, only at six weeks. Six- to 12-week running volumes less than 14.8 km positively moderated IVD T2 between-group at 12 weeks only (Supplement E). No other running volumes moderated between-group changes.

Mean running speeds from 10.5 to 11.7 km/h positively moderated between-group IVD T2 at 12 weeks (Table 5 and Supplement D). Whereas mean running speeds between 4.4 and 8.5 km/h negatively moderated IVD T2 at six weeks

Table 2 Descriptive characteristics of participants at baseline

	Intervention ($n=20$)	Control ($n=20$)
Age, years	33.6 (5.3)	32.2 (7.0)
Female, n (%)	10 (50)	10 (50)
Body mass index, kg/m^2	29.6 (6.9)	29.0 (7.5)
Current LBP duration, years	3.2 (2.8)	4.9 (5.8)
Pain, visual analogue scale (0–100)		
Current	30.8 (23.3)	40.1 (20.9)
Average, last seven days	33.5 (20.6)	46.0 (20.1)
Worst, last seven days	50.6 (22.5)	65.9 (17.9)
Disability, Oswestry Disability Index (0–100)	20.8 (8.5)	23.1 (9.7)
Mean IVD T2 (ms)	81.22 (10.34)	85.59 (12.85)
Mean Pfirrmann grade (1–5)	3.13 (0.66)	2.94 (0.63)
Pfirrmann grade at baseline (IVDs)		
No degeneration (grade 1)	0	0
Mild degeneration (grade 2)	13	29
Moderate degeneration (grade 3)	106	95
Severe degeneration (grade 4)	17	13
Advanced degeneration (grade 5)	4	3
Median Habitual Physical Activity (IPAQ; MET-minutes/week; IQR)	1814 (1303, 3250.5)	1512 (636, 3312)
Employment status, n (%)		
Employed	18 (90)	18 (90)
Unemployed	2 (10)	1 (5)
Homemaker	0 (0)	1 (5)
Retired	0 (0)	0 (0)
Smoking status, n (%)		
Current	2 (10)	0 (0)
Former	1 (5)	3 (15)
Never smoked	17 (85)	17 (85)

Data are mean (SD), count (percentage within-group), or median (IQR). IVD: Intervertebral disc; IPAQ: International Physical Activity Questionnaire – short form.

only. No other running speeds moderated between-group IVD T2 changes.

Running on grass positively moderated between-group IVD T2 at 12 weeks (Table 5 and Supplement D). No other running surface moderated IVD T2 changes.

Sensitivity analyses

Adjustment for individual baseline characteristics did not change the results of any primary analyses (Supplement F).

Harms

No study-related serious adverse events were reported. Nine study-related non-serious adverse events were reported in the intervention group: one exacerbation of LBP and seven lower limb complaints (calf and knee), which are reported in detail elsewhere [12].

Discussion

This is the first study to investigate whether patient or intervention factors moderated IVD T2 changes in a progressive run-walk intervention. The highest aggregated Pfirrmann scores and male sex positively moderated IVD T2 at 12 weeks. Mild IVD degeneration, the lowest aggregated Pfirrmann scores, and male sex negatively moderated the intervention effect at six weeks. When compared to the entire control sample, cumulative running volume between 28.6 and 46.1 km, mean running speeds from 10.50 to 11.7 km/h, and running on grass positively moderated effects at 12 weeks. When cumulative running volume was less than 28.6 km or between 46.1 and 81.2 km, and running speeds were between 4.4 and 8.46 km/h, there was a negative moderating effect at six weeks; however, no effects were observed at 12 weeks.

Our finding that only the highest aggregated Pfirrmann group (i.e., participants with the most multi-site degeneration) positively moderated the intervention effect on IVD T2 contrasts with much of the existing preclinical literature. In vitro studies of human IVD cells have found that degenerated IVDs are less responsive to mechanical loading than healthy IVDs [14]. Similarly, acute loading and cross-sectional imaging studies in humans have reported minimal or absent adaptive responses in degenerated IVDs compared to healthy controls [42, 43]. However, these studies are typically limited to short-term or acute responses, and there is a lack of longitudinal human data. Animal studies have shown that, over several days, mildly degenerated discs may upregulate anabolic genes in response to loading, while more severely degenerated discs may downregulate these genes [44]. Yet, these findings are based on interventions lasting only a few days and our 12-week data may be reflective of a positive chronic effect. This capacity for degenerated IVDs to adapt may reflect a 'diminishing returns' principle, whereby IVDs with lower baseline hydration have greater potential for rehydration and measurable improvement [45]. An alternate theory is that our findings may reflect the IVD vacuum phenomenon, where fissures in advanced degeneration allow gas and fluid pockets to form [46]. Higher T2 values may represent fluid intermittently replacing gas, a dynamic process that compromises structural integrity [47]. This phenomenon is best detected with computed tomography, which could be included in future studies to confirm this mechanism [47]. It is important to note that our sample of severely degenerated IVDs was small, and although we observed relatively positive changes in T2, these were not statistically significant compared to controls. Future studies with larger samples and longer follow-up are needed to confirm if there is a positive chronic effect in degenerated IVDs.

Table 3 Changes in IVD T2 by baseline degeneration grade and aggregated Pfirrmann score quartiles across spinal levels T11/T12 to L5/S1 over time at the IVD level

Variable	Intervention			Control			Group-by-time	
	Mean (SE)	Mean (95% CI) Δ from baseline	*P*	Mean (SE)	Mean (95% CI) Δ from baseline	*P*	β (95% CI)	*P*
Baseline Pfirrmann grade								
Mild degeneration		***N*=13 IVDs**			***N*=29 IVDs**			
Whole IVD T2								
Baseline	87.07 (3.33)			100.14 (2.23)				
6-week	86.81 (3.33)	−0.26 (−4.24, 3.72)	0.898	104.67 (2.23)	4.54 (1.87, 7.20)	**0.001**	−4.80 (−9.59, −0.01)	**0.050**
12-week	87.36 (3.33)	0.29 (−3.69, 4.27)	0.886	98.88 (2.23)	−1.26 (−3.92, 1.40)	0.354	1.55 (−3.24, 6.34)	0.526
Moderate degeneration		***N*=106 IVDs**			***N*=95 IVDs**			
Whole IVD T2								
Baseline	82.75 (1.16)			83.09 (1.20)				
6-week	82.86 (1.16)	0.11 (−0.88, 1.11)	0.822	82.97 (1.21)	−0.12 (−1.21, 0.97)	0.832	0.23 (−1.24, 1.71)	0.758
12-week	82.74 (1.16)	−0.01 (−1.00, 0.99)	0.991	82.52 (1.21)	−0.57 (−1.66, 0.52)	0.304	0.56 (−0.91, 2.04)	0.453
Severe degeneration		***N*=17 IVDs**			***N*=13 IVDs**			
Whole IVD T2								
Baseline	70.01 (1.75)			74.90 (2.01)				
6-week	69.24 (1.75)	−0.77 (−3.26, 1.72)	0.546	73.55 (2.01)	−1.35 (−4.20, 1.50)	0.354	0.58 (−3.21, 4.37)	0.764
12-week	68.07 (1.75)	−1.94 (−4.44, 0.55)	0.127	71.78 (2.01)	−3.12 (−5.97, −0.27)	**0.032**	1.18 (−2.61, 4.97)	0.542
Advanced degeneration		***N*=4 IVDs**			***N*=3 IVDs**			
Whole IVD T2								
Baseline	78.91 (6.82)			74.95 (7.87)				
6-week	79.63 (6.82)	0.71 (−9.51, 10.94)	0.891	77.47 (7.87)	2.52 (−9.29, 14.32)	0.676	−1.80 (−17.42, 13.82)	0.821
12-week	83.94 (6.82)	5.03 (−5.20, 15.25)	0.335	76.41 (7.87)	1.45 (−10.35, 13.26)	0.810	3.58 (−12.04, 19.20)	0.654
Baseline aggregated Pfirrmann sum in quartiles								
Quartile 1: <21		***N*=21 IVDs**			***N*=56 IVDs**			
Whole IVD T2								
Baseline	78.79 (3.07)			93.33 (2.07)				
6-week	77.63 (3.07)	−1.17 (−3.98, 1.65)	0.417	95.69 (2.07)	2.36 (0.63, 4.08)	**0.007**	−3.53 (−6.83, −0.22)	**0.036**
12-week	78.95 (3.07)	0.15 (−2.67, 2.97)	0.916	92.24 (2.07)	−1.09 (−2.81, 0.64)	0.216	1.24 (−2.06, 4.54)	0.461
Quartile 2: 21		***N*=42 IVDs**			***N*=42 IVDs**			
Whole IVD T2								
Baseline	87.15 (1.73)			80.92 (1.73)				
6-week	86.69 (1.73)	−0.46 (−2.14, 1.21)	0.587	80.73 (1.77)	−0.18 (−2.01, 1.65)	0.846	−0.28 (−2.76, 2.20)	0.822
12-week	85.35 (1.73)	−1.80 (−3.48, −0.12)	**0.035**	79.63 (1.77)	−1.28 (−3.11, 0.55)	0.169	−0.52 (−3.00, 1.96)	0.682
Quartile 3: 22		***N*=42 IVDs**			***N*=21 IVDs**			
Whole IVD T2								

Table 3 (continued)

Variable	Intervention			Control			Group-by-time	
	Mean (SE)	Mean (95% CI) Δ from baseline	*P*	Mean (SE)	Mean (95% CI) Δ from baseline	*P*	β (95% CI)	*P*
Baseline	81.45 (1.71)			85.42 (2.10)				
6-week	81.03 (1.71)	−0.41 (−1.82, 1.00)	0.566	84.84 (2.10)	−0.58 (−2.57, 1.41)	0.569	0.17 (−2.27, 2.61)	0.894
12-week	80.54 (1.71)	−0.90 (−2.31, 0.50)	0.208	85.02 (2.10)	−0.40 (−2.39, 1.60)	0.697	−0.51 (−2.95, 1.93)	0.683
Quartile 4: ≥23		***N*=35 IVDs**			***N*=21 IVDs**			
Whole IVD T2								
Baseline	75.74 (1.48)			74.47 (1.86)				
6-week	77.45 (1.48)	1.71 (−0.25, 3.67)	0.087	74.22 (1.86)	−0.24 (−2.77, 2.29)	0.851	1.95 (−1.25, 5.15)	0.231
12-week	78.62 (1.48)	2.88 (0.92, 4.83)	**0.004**	73.92 (1.86)	−0.55 (−3.07, 1.98)	0.672	3.42 (0.22, 6.62)	**0.036**

Data are sample size, estimated marginal mean (SE), within-group mean change (Δ) from baseline (95% CI), within-group P-value, group-by-time β coefficient (95% CI) and group-by-time P-value from linear mixed models. Mild degeneration=Pfirrmann grade 2, moderation degeneration=Pfirrmann grade 3, Severe degeneration=Pfirrmann grade 4 and advanced degeneration=Pfirrmann grade 5. One control group participant did not complete magnetic resonance outcome measures at the six weeks and twelve weeks (n=133). Observed mean (SD) are presented in Supplement D, Table B and C. IVD: intervertebral disc

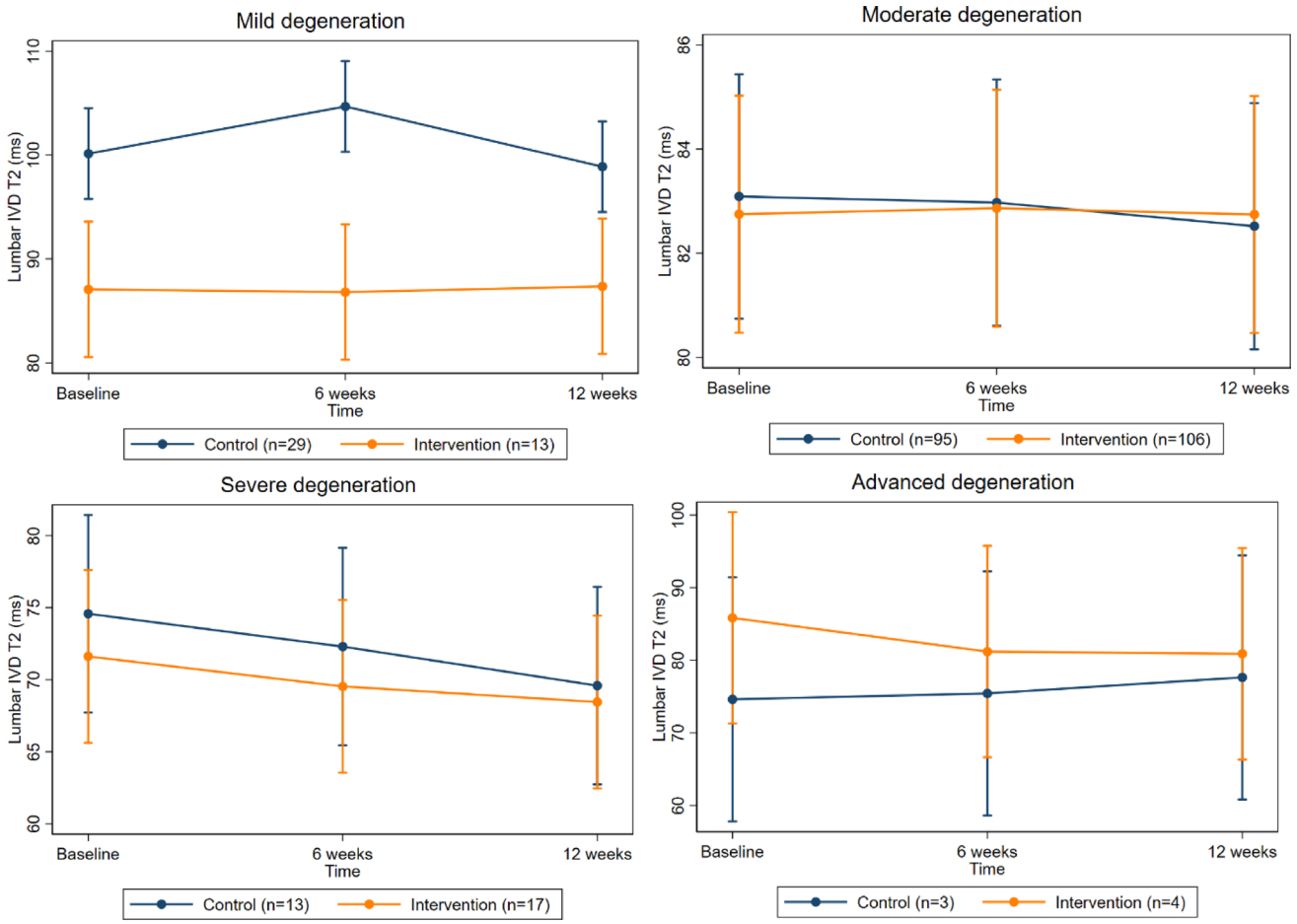

Fig. 3 Raw mean (SD) IVD T2 at baseline, six weeks and 12 weeks stratified by Pfirrmann grade degeneration. N as number of IVDS. IVD: intervertebral disc

Table 4 Changes in IVD T2 by sex and baseline body mass index over time at the IVD level

Variable	Intervention			Control			Group-by-time	
	Mean (SE)	Mean (95% CI) Δ from baseline	*P*	Mean (SE)	Mean (95% CI) Δ from baseline	*P*	β (95% CI)	*P*
Sex								
Male		***N*=70 IVDs**			***N*=70 IVDs**			
Whole IVD T2								
Baseline	79.26 (1.80)			83.95 (1.80)				
6-week	81.84 (1.82)	−0.22 (−1.61, 1.18)	0.761	79.47 (1.80)	2.24 (0.78, 3.71)	**0.003**	−2.46 (−4.48, −0.44)	**0.017**
12-week	79.62 (1.80)	0.14 (−1.25, 1.54)	0.842	86.20 (1.82)	−2.12 (−3.58, −0.65)	**0.005**	2.26 (0.24, 4.28)	**0.029**
Female		***N*=70 IVDs**			***N*=70 IVDs**			
Whole IVD T2								
Baseline	83.39 (1.56)			87.22 (1.56)				
6-week	87.34 (1.56)	0.19 (−1.09, 1.48)	0.767	83.19 (1.56)	−0.46 (−1.75, 0.82)	0.480	0.66 (−1.16, 2.48)	0.478
12-week	82.91 (1.56)	−0.28 (−1.57, 1.01)	0.669	86.76 (1.56)	0.11 (−1.17, 1.40)	0.861	−0.40 (−2.21, 1.42)	0.670
Baseline body mass index stratified by quartiles								
Quartile 1: < 25.21		***N*=21 IVDs**			***N*=49 IVDs**			
Whole IVD T2								
Baseline	93.28 (3.02)			93.89 (1.98)				
6-week	90.80 (1.98)	−0.82 (−3.63, 1.99)	0.569	94.10 (3.02)	1.31 (−0.53, 3.15)	0.163	−2.13 (−5.48, 1.23)	0.215
12-week	90.40 (3.02)	−3.70 (−6.51, −0.89)	**0.010**	95.20 (1.98)	−3.10 (−4.93, −1.26)	**0.001**	−0.60 (−3.96, 2.76)	0.725
Quartile 2: 25.21–27.56		***N*=63 IVDs**			***N*=7 IVDs**			
Whole IVD T2								
Baseline	78.50 (1.21)			74.45 (3.54)				
6-week	72.41 (3.54)	0.29 (−0.96, 1.54)	0.654	78.22 (1.21)	−1.35 (−5.10, 2.40)	0.481	1.63 (−2.32, 5.59)	0.418
12-week	78.03 (1.21)	−0.19 (−1.44, 1.06)	0.765	73.10 (3.54)	−2.04 (−5.79, 1.71)	0.287	1.85 (−2.11, 5.80)	0.360
Quartile 3: 27.56–33.68		***N*=28 IVDs**			***N*=42 IVDs**			
Whole IVD T2								
Baseline	80.61 (1.94)			83.59 (1.58)				
6-week	83.59 (1.62)	−0.23 (−2.12, 1.66)	0.814	80.84 (1.94)	0.21 (−1.47, 1.89)	0.805	−0.44 (−2.97, 2.09)	0.734
12-week	80.91 (1.94)	0.08 (−1.81, 1.97)	0.937	83.80 (1.62)	0.00 (−1.68, 1.68)	0.999	0.08 (−2.45, 2.60)	0.954
Quartile 4: >33.68		***N*=28 IVDs**			***N*=42 IVDs**			
Whole IVD T2								
Baseline	79.41 (2.14)			79.76 (1.93)				
6-week	80.76 (1.93)	0.14 (−2.01, 2.30)	0.896	79.27 (2.14)	1.15 (−0.61, 2.91)	0.200	−1.01 (−3.79, 1.78)	0.479
12-week	82.05 (2.14)	2.78 (0.62, 4.94)	**0.012**	80.91 (1.93)	1.00 (−0.76, 2.76)	0.264	1.78 (−1.01, 4.56)	0.211

Data are sample size, estimated marginal mean (SE), within-group mean change (Δ) from baseline (95% CI), within-group P-value, group-by-time β coefficient (95% CI) and group-by-time P-value from linear mixed models. Mild degeneration = Pfirrmann grade 2, moderation degeneration = Pfirrmann grade 3, Severe degeneration = Pfirrmann grade 4 and advanced degeneration = Pfirrmann grade 5. One control group participant did not complete magnetic resonance outcome measures at the six weeks and twelve weeks (n = 133). Observed mean (SD) are presented in Supplement D, Table B. IVD: intervertebral disc.

A negative moderating effect was observed at six weeks in the smallest aggregated Pfirrmann group, mildly degenerated IVDs (grade two) and in male participants. This effect was not sustained at 12 weeks, which suggests these changes may reflect the acute effects of starting a running intervention or reflect a type I error. The between-group changes in male participants were explained by within-control group changes, which also suggests that within-group variation rather than an intervention effect may partially explain this result. In contrast, previous in vitro acute loading studies reveal that when exposed to physical loading, healthy IVDs improve health (increased hydration), and degenerated IVDs do not [14]. This contrasts with the current study, where the healthiest IVDs decreased health (decreased hydration) within six weeks, whereas more degenerated IVDs did not. Animal data have shown the efficacy of cyclic loading in

Table 5 Changes in IVD T2 stratified by intervention running volume spread into quartiles across spinal levels T11/T12 to L5/S1 over time

Variable	Intervention			Control			Group-by-time	
	Mean (SE)	Mean (95% CI) Δ from baseline	*P*	Mean (SE)	Mean (95% CI) Δ from baseline	*P*	β (95% CI)	*P*
Run volume from baseline to 12 weeks								
Quartile 1: <28.61 km		***N*=35 IVDs**			***N*=140 IVDs**			
Whole IVD T2								
Baseline	78.39 (2.24)			85.59 (1.22)				
6-week	76.74 (2.24)	−1.65 (−3.63, 0.33)	0.102	86.42 (1.23)	0.84 (−0.18, 1.85)	0.107	−2.48 (−4.71, −0.26)	**0.029**
12-week	77.39 (2.24)	−1.00 (−2.98, 0.97)	0.320	84.66 (1.23)	−0.93 (−1.94, 0.09)	0.074	−0.08 (−2.30, 2.15)	0.944
Quartile 2: 28.61–46.08 km		***N*=35 IVDs**			***N*=140 IVDs**			
Whole IVD T2								
Baseline	76.73 (2.23)			85.59 (1.24)				
6-week	78.80 (2.23)	2.07 (0.03, 4.11)	**0.047**	86.43 (1.24)	0.84 (−0.21, 1.89)	0.115	1.23 (−1.07, 3.52)	0.294
12-week	80.61 (2.23)	3.88 (1.84, 5.92)	**<0.001**	84.67 (1.24)	−0.92 (−1.97, 0.13)	0.085	4.80 (2.51, 7.10)	**<0.001**
Quartile 3: 46.08–81.24 km		***N*=35 IVDs**			***N*=140 IVDs**			
Whole IVD T2								
Baseline	84.62 (2.25)			85.59 (1.22)				
6-week	83.09 (2.25)	−1.53 (−3.46, 0.39)	0.118	86.42 (1.23)	0.83 (−0.15, 1.82)	0.098	−2.37 (−4.53, −0.21)	**0.032**
12-week	83.66 (2.25)	−0.97 (−2.89, 0.95)	0.324	84.66 (1.23)	−0.93 (−1.92, 0.05)	0.064	−0.04 (−2.20, 2.12)	0.973
Quartile 4: >81.24 km		***N*=35 IVDs**			***N*=140 IVDs**			
Whole IVD T2								
Baseline	85.59 (2.19)			85.59 (1.15)				
6-week	86.66 (2.19)	1.07 (−0.98, 3.12)	0.306	86.43 (1.15)	0.84 (−0.21, 1.89)	0.117	0.23 (−2.07, 2.54)	0.844
12-week	83.40 (2.19)	−2.19 (−4.24, −0.13)	**0.037**	84.67 (1.15)	−0.92 (−1.97, 0.13)	0.086	−1.27 (−3.57, 1.04)	0.282
Mean running speed from baseline to 12 weeks								
Quartile 1:< 8.46 km/h		***N*=35 IVDs**			***N*=140 IVDs**			
Whole IVD T2								
Baseline	81.43 (2.25)			85.59 (1.20)				
6-week	79.93 (2.25)	−1.50 (−3.50, 0.49)	0.140	86.42 (1.20)	0.83 (−0.19, 1.86)	0.110	−2.33 (−4.58, −0.09)	**0.041**
12-week	81.21 (2.25)	−0.22 (−2.22, 1.77)	0.827	84.66 (1.20)	−0.93 (−1.95, 0.10)	0.075	0.71 (−1.54, 2.95)	0.538
Quartile 2: 8.46–9.84 km/h		***N*=35 IVDs**			***N*=140 IVDs**			
Whole IVD T2								
Baseline	79.20 (2.27)			85.59 (1.26)				

Table 5 (continued)

Variable	Intervention			Control			Group-by-time	
	Mean (SE)	Mean (95% CI) Δ from baseline	*P*	Mean (SE)	Mean (95% CI) Δ from baseline	*P*	β (95% CI)	*P*
6-week	79.81 (2.27)	0.61 (−1.34, 2.55)	0.541	86.42 (1.27)	0.83 (−0.17, 1.83)	0.102	−0.22 (−2.41, 1.96)	0.841
12-week	80.08 (2.27)	0.87 (−1.08, 2.82)	0.381	84.66 (1.27)	−0.93 (−1.93, 0.07)	0.068	1.80 (−0.39, 3.99)	0.107
Quartile 3: 9.84–10.50 km/h		***N*=35 IVDs**			***N*=140 IVDs**			
Whole IVD T2								
Baseline	81.42 (2.24)			85.59 (1.25)				
6-week	81.82 (2.24)	0.40 (−1.63, 2.43)	0.699	86.43 (1.26)	0.84 (−0.20, 1.88)	0.113	−0.44 (−2.72, 1.84)	0.706
12-week	78.95 (2.24)	−2.47 (−4.50, −0.45)	**0.017**	84.67 (1.26)	−0.92 (−1.96, 0.12)	0.082	−1.55 (−3.83, 0.73)	0.182
Quartile 4: > 10.50 km/h		***N*=35 IVDs**			***N*=140 IVDs**			
Whole IVD T2								
Baseline	83.28 (2.19)			85.59 (1.14)				
6-week	83.73 (2.19)	0.45 (−1.62, 2.52)	0.670	86.43 (1.15)	0.84 (−0.22, 1.90)	0.120	−0.39 (−2.72, 1.94)	0.742
12-week	84.83 (2.19)	1.54 (−0.53, 3.62)	0.144	84.67 (1.15)	−0.92 (−1.98, 0.14)	0.090	2.46 (0.14, 4.79)	**0.038**
Running surface from baseline to 12 weeks								
Grass		***N*=21 IVDs**			***N*=140 IVDs**			
Whole IVD T2								
Baseline	88.50 (2.86)			85.59 (1.11)				
6-week	84.66 (1.11)	1.05 (−1.59, 3.68)	0.436	87.46 (2.86)	0.84 (−0.21, 1.88)	0.117	0.21 (−2.62, 3.04)	0.884
12-week	89.46 (2.86)	2.01 (−0.63, 4.64)	0.135	86.42 (1.11)	−0.93 (−1.97, 0.12)	0.083	2.93 (0.10, 5.76)	**0.043**
Gravel		***N*=7 IVDs**			***N*=140 IVDs**			
Whole IVD T2								
Baseline	82.93 (5.02)			85.59 (1.12)				
6-week	84.67 (1.13)	−2.08 (−6.85, 2.69)	0.393	85.01 (5.02)	0.84 (−0.25, 1.93)	0.132	−2.92 (−7.82, 1.98)	0.243
12-week	85.32 (5.02)	0.31 (−4.46, 5.08)	0.899	86.43 (1.13)	−0.92 (−2.01, 0.17)	0.099	1.23 (−3.67, 6.13)	0.622
Pavement		***N*=98 IVDs**			***N*=140 IVDs**			
Whole IVD T2								
Baseline	79.70 (1.59)			85.59 (1.44)				
6-week	84.67 (1.45)	−0.18 (−1.34, 0.98)	0.759	79.88 (1.59)	0.84 (−0.15, 1.83)	0.097	−1.02 (−2.54, 0.50)	0.189
12-week	79.33 (1.59)	−0.55 (−1.70, 0.61)	0.355	86.43 (1.45)	−0.92 (−1.91, 0.07)	0.068	0.38 (−1.15, 1.90)	0.629
Trail		***N*=14 IVDs**			***N*=140 IVDs**			
Whole IVD T2								
Baseline	81.13 (3.50)			85.59 (1.15)				
6-week	84.66 (1.16)	0.63 (−2.56, 3.81)	0.700	80.51 (3.50)	0.83 (−0.20, 1.87)	0.113	−0.21 (−3.56, 3.14)	0.903
12-week	80.47 (3.50)	−0.04 (−3.22, 3.15)	0.981	86.42 (1.16)	−0.93 (−1.96, 0.11)	0.079	0.89 (−2.46, 4.24)	0.603

Data are sample size, estimated marginal mean (SE), within-group mean change (Δ) from baseline (95% CI), within-group P-value, group-by-time β coefficient (95% CI) and group-by-time P-value from linear mixed models. One control group participant did not complete magnetic resonance outcome measures at six weeks and twelve weeks ($n=133$). Observed mean (SD) are presented in Supplement D, Table D–F. IVD: intervertebral disc.

preventing the progression of IVD degeneration [48]. Rabbits with healthy IVDs (Pfirrmann grade one) underwent experimentally induced IVD degeneration using established methods [48], which typically progresses over 24 weeks [49]. After eight weeks, a subset of the rabbits was exposed to slow-rate cyclic loading for two hours per day, on five days per week, and at a loading rate of 0.5 cycles per second [49]. After a further eight weeks, the average Pfirrmann grades had progressed to grade five in the control group and only to grade 2.1 in the cyclic loading group. This animal study suggests a protective effect of cyclic loading against the degenerative process, and running may be a viable intervention for preventing IVD degeneration [49]. Thus, our observations in mildly degenerated IVDs at six weeks were unexpected and may be explained by another mechanism beyond running. Moreover, the absence of adverse changes at 12 weeks suggests that prolonged exposure to running may be beneficial for IVD health, reinforcing the rationale for evaluating longer-term follow-up.

Running volume positively moderated IVD T2 changes only when between 28.9 and 46.1 km over 12 weeks, suggesting evidence of an optimal volume of running for IVD health. Cell-based studies suggest a U-shaped relationship between physical loading and IVD adaptation, with optimal loading volumes promoting beneficial changes and loading factors above and below the optimal range leading to detrimental changes [7]. The optimal duration of loading was previously estimated between 30 min and eight hours per day [7]. Whereas the current study showed evidence of running volumes that are likely optimal for human IVDs. When divided by twelve weeks, the volumes equate to a weekly average running distance of 2.4–3.8 km, over two to three sessions per week. Further, the lowest running volume (< 4.81 km) from 6−12 weeks was the only subgroup to positively moderate IVD T2 changes. Thus, a much lower running volume than previously reported may be sufficient and perhaps more optimal for IVD health. Without other prospective trials to compare with, these observations are difficult to interpret meaningfully. Cross-sectional studies have demonstrated that long-distance running for more than 5 years (> 50 km per week) was associated with better IVD health [50, 51]. The running volume in the current study is not comparable to 50 km per week; however, we suggest that the optimal volume of running increases with more prolonged exposure to the running intervention. This theory is supported by the negative T2 changes observed in two subgroups (< 28.6 km and 46.1 − 81.2 km) at six weeks, given that they were not present at 12 weeks, indicating a diminishing effect with prolonged exposure. In summary, the conservative running volumes in the current study, which led to increased IVD T2 (28.9–46.1 km), may be optimal when first starting a running program and may increase over time (> 12 weeks).

In the current study, participants running at the highest mean speeds from 10.5 to 11.7 km/h and those running on grass revealed increased IVD T2 at 12 weeks. Cell models suggest that slower cyclic motions, equivalent to jogging speeds of 7.2 km/h, positively moderate optimal loading for IVD health [7], which is further supported given that rapid or sudden loading appeared detrimental [6, 52]. The highest speeds in the current study exceed those previously proposed as optimal; however, prior recommendations are based on in vitro or animal model data [8, 9]. Faster running speeds correlate with higher running cadence [53]. Mechanistically, higher running cadence is associated with a decreased risk of running-related injuries through decreased biomechanical loads [54], which may also be the case for optimising mechanical loads absorbed by the IVDs. Although, this remains speculative without directly measuring cadence with accelerometry. Similarly, running on grass is associated with a lower biomechanical load compared to running on concrete [20, 55], which may be more optimal for IVD loading [20]. Participants running at slower speeds (4.4–8.5 km/h) revealed decreased IVD T2. Slow recreational runners exhibit considerable variation in running gait stance time [56], which increases time under load and likely reduces the number of cyclic loading impacts. However, the decreased T2 was only revealed at six weeks, which potentially demonstrates the acute effects of starting a run program, with the longer-term effects by 12 weeks offsetting any early negative changes. A caveat worth consideration is that the mean running speed was moderately correlated with 12-week running volume (Supplement F). This indicates that those who ran further tended to run faster as well. Thus, the combination of both factors may have precluded the observation of significant moderating effects. The overall changes in IVD T2 stratified by the running speed and grass surface were less than the IVD T2 changes in the volume subgroups, which suggests speed and surface may be weaker moderators than running volume. Nonetheless, running at faster speeds than previously suggested and running on grass warrants further investigation.

The clinical implications of this study require reference to pain and feasibility outcomes, reported elsewhere [12]. We previously showed this intervention reduced pain by fifteen points on a 0−100 visual analogue scale compared to control and there was zero intervention attrition at 12 weeks. However, the mechanisms underpinning these changes remain unknown. LBP is multifactorial and only a portion of individuals (40%) [22] are considered impacted by IVD-related pain; hence, whether improving IVD T2 values will have implications for improving pain remains unclear. A longer intervention with a larger sample could reveal more

pronounced IVD T2 changes, enabling mediation analysis to test whether these mechanisms explain changes in subjective low back pain. To better understand the impact of moderating factors, future studies could define subgroups a priori and measure moderators using more robust methods. For example, using Dual-Energy X-ray Absorptiometry for body composition and accelerometry for habitual physical activity and intervention running metrics, including volume, speed and incline.

Our study was strengthened by a robust single-blinded (data processing blinded to allocation and time point) randomised controlled trial design and using magnetic resonance imaging, which is the gold standard for measuring IVD health, and has excellent long-term reliability for the primary outcome of IVD T2 [21]. However, there are several limitations. First, the quartile subgroups for baseline degeneration and intervention cumulative running volume, mean running speed, and running surface resulted in small samples, which may have markedly reduced statistical power. However, our analyses included IVDs rather than participants, which markedly increased the overall sample size. Second, our observations are generalisable only to individuals aged 18–45 experiencing non-specific CLBP. Third, intradiscal pressures were not directly measured. Any moderating effects attributed to influencing IVD health remain inferential and based on biomechanical principles. Fourth, participants in the control group were instructed to refrain from running but were not restricted from engaging in other forms of physical activity, which may have confounded IVD changes. Fifth, IVD T2 was measured across the whole IVD rather than region-specific, which may obscure localised changes. Finally, clinically meaningful changes in IVD T2 remain undefined, and a statistically significant difference may not necessarily reflect physiological or therapeutic relevance.

Conclusion

Intervention factors may have positively moderated IVD T2 changes in a 12-week run-walk intervention in adults with CLBP. Any negative moderating effects were observed only at six weeks, suggesting the effects may be attenuated by longer-term loading, which warrants further investigation. Our data suggests a weekly average running distance of 2.4–3.8 km in the short term, running speeds at approximately 11 km/h, and running on grass may optimise IVD health; however, these recommendations remain speculative. Future studies with larger subgroups across different degeneration levels would help determine whether mechanotransduction through running can play either a preventive or therapeutic role in improving IVD health.

Supplementary Information The online version contains supplementary material available at https://doi.org/10.1007/s00586-026-09759-7.

Acknowledgements The authors thank the participants for taking part in the study, the staff at Imaging @ Olympic Park (Melbourne, Australia) for facilitating study conduct, and the broader ASTEROID study research team.

Author contributions Author contributions: Conceptualisation: SDT, DLB, UHM, PJO; Data curation: CLS, CN, UHM, HRN; Formal Analysis: CLS; Funding acquisition: DLB, UHM, DS, PJO; Investigation: CLS, UHM, HRN, PJO; Methodology: All; Project administration: CLS, CN, SDT, PJO; Resources: PJO; Software: CLS, PJO; Supervision: CTM, NLM, DS, PJO; Validation: PJO; Visualisation: CLS, PJO; Writing – original draft: CLS; Writing – review & editing: All.

Funding Open Access funding enabled and organized by CAUL and its Member Institutions

Data availability The datasets used and analysed during the current study are available from the corresponding author on reasonable request.

Declarations

Competing interests The authors declare no competing interests.

References

1. Brinjikji W, Diehn FE, Jarvik JG et al (2015) MRI findings of disc degeneration are more prevalent in adults with low back pain than in asymptomatic controls: a systematic review and meta-analysis. AJNR Am J Neuroradiol 36:2394–2399. https://doi.org/10.3174/ajnr.A4498
2. Diwan AD, Melrose J (2022) Intervertebral disc degeneration and how it leads to low back pain. JOR Spine 6:e1231. https://doi.org/10.1002/jsp2.1231
3. Wang X, Wanyan P, Tian JH, Hu L (2015) Meta-analysis of randomized trials comparing fusion surgery to non-surgical treatment for discogenic chronic low back pain. J Back Musculoskelet Rehabil 28:621–627. https://doi.org/10.3233/BMR-140571
4. Lewin AM, Fearnside M, Kuru R et al (2021) Rates, costs, return to work and reoperation following spinal surgery in a workers' compensation cohort in new South Wales, 2010–2018: a cohort study using administrative data. BMC Health Serv Res 21:1–10. https://doi.org/10.1186/s12913-021-06900-8

5. Bovonratwet P, Ottesen TD, Gala RJ et al (2018) Outpatient elective posterior lumbar fusions appear to be safely considered for appropriately selected patients. Spine J 18:1188–1196. https://doi.org/10.1016/j.spinee.2017.11.011
6. Belavỳ DL, Albracht K, Bruggemann G-P et al (2016) Can exercise positively influence the intervertebral disc? Sports Med 46:473–485. https://doi.org/10.1007/s40279-015-0444-2
7. Chan SC, Ferguson SJ, Gantenbein-Ritter B (2011) The effects of dynamic loading on the intervertebral disc. Eur Spine J 20:1796–1812. https://doi.org/10.1007/s00586-011-1827-1
8. Säämänen A-M, Puustjärvi K, Ilves K et al (1993) Effect of running exercise on proteoglycans and collagen content in the intervertebral disc of young dogs. Int J Sports Med 14:48–51. https://doi.org/10.1055/s-2007-1021145
9. Brisby H, Wei AQ, Molloy T et al (2010) The effect of running exercise on intervertebral disc extracellular matrix production in a rat model. Spine 35:1429–1436. https://doi.org/10.1007/s40279-015-0444-2
10. Tagliaferri SD, Belavý DL, Bowe SJ et al (2023) Assessing safety and treatment efficacy of running on intervertebral discs (ASTEROID) in adults with chronic low back pain: protocol for a randomised controlled trial. BMJ Open Sport Exerc Med 9:e001524. https://doi.org/10.1136/bmjsem-2022-001524
11. Owen PJ, Miller CT, Rantalainen T et al (2020) Exercise for the intervertebral disc: a 6-month randomised controlled trial in chronic low back pain. Eur Spine J 29:1887–1899. https://doi.org/10.1007/s00586-020-06379-7
12. Neason C, Samanna CL, Tagliaferri SD et al (2025) Running is acceptable and efficacious in adults with non-specific chronic low back pain: the ASTEROID randomised controlled trial. Br J Sports Med 59:99–108. https://doi.org/10.1136/bjsports-2024-108245
13. Zhang W, Wang H, Yuan Z et al (2021) Moderate mechanical stimulation rescues degenerative annulus fibrosus by suppressing caveolin-1 mediated pro-inflammatory signaling pathway. Int J Biol Sci 17:1395. https://doi.org/10.7150/ijbs.57774
14. Le Maitre CL, Frain J, Fotheringham AP et al (2008) Human cells derived from degenerate intervertebral discs respond differently to those derived from non-degenerate intervertebral discs following application of dynamic hydrostatic pressure. Biorheology 45:563–575. https://doi.org/10.3233/BIR-2008-0498
15. Coppock JA, Danyluk ST, Englander ZA et al (2021) Increasing BMI increases lumbar intervertebral disc deformation following a treadmill walking stress test. J Biomech 121:110392. https://doi.org/10.1016/j.jbiomech.2021.110392
16. Stott B, Driscoll M (2024) Development and evaluation of sex-specific thoracolumbar spine finite element models to study spine biomechanics. Med Biol Eng Comput 62:1191–1199. https://doi.org/10.1007/s11517-023-03003-w
17. Jiang X, Bíró I, Sárosi J et al (2024) Comparison of ground reaction forces as running speed increases between male and female runners. Front Bioeng Biotechnol 12:p.1378284. https://doi.org/10.3389/fbioe.2024.1378284
18. Nahhas Rodacki C, Monteiro CA, Paulo AC et al (2024) Cadence matters: influence of cadence on spinal load during running. Gait Posture 113:528–533. https://doi.org/10.1016/j.gaitpost.2024.07.298
19. Castillo ER, Lieberman DE (2018) Shock attenuation in the human lumbar spine during walking and running. J Exp Biol 221:jeb177949. https://doi.org/10.1242/jeb.177949
20. Ferro-Sánchez A, Martín-Castellanos A, de la Rubia A et al (2023) An analysis of running impact on different surfaces for injury prevention. Int J Environ Res Public Health 20:6405. https://doi.org/10.3390/ijerph20146405
21. Belavý DL, Owen PJ, Armbrecht G et al (2021) Quantitative assessment of the lumbar intervertebral disc via T2 shows excellent long-term reliability. PLoS One 16:e0249855. https://doi.org/10.1371/journal.pone.0249855
22. Zhang X, Yang L, Gao F et al (2015) Comparison of T1ρ and T2* relaxation mapping in patients with different grades of disc degeneration at 3T MR. Med Sci Monit 21:1934. https://doi.org/10.12659/MSM.894406
23. Kirnaz S, Capadona C, Lintz M et al (2021) Pathomechanism and biomechanics of degenerative disc disease: features of healthy and degenerated discs. Int J Spine Surg 15:10–25. https://doi.org/10.14444/8052
24. Inge P, Orchard JJ, Purdue R, Orchard JW (2022) Exercise after pregnancy. Aust J Gen Pract 51:117–121. https://doi.org/10.31128/AJGP-09-21-6181
25. Kujala UM (2021) Fitter, healthier and stronger? Many factors influence elite athletes' long-term health. Br J Sports Med 55:77–78. https://doi.org/10.1136/bjsports-2020-102239
26. de Rooij M, Steultjens MPM, Avezaat E et al (2013) Restrictions and contraindications for exercise therapy in patients with hip and knee osteoarthritis and comorbidity. Phys Ther Rev 18:101–111. https://doi.org/10.1179/1743288X12Y.0000000056
27. Norton K, Norton L (2011) Pre-exercise screening. Guide Aust Adult Pre-Exerc Screen Syst Exerc Sports Sci Aust
28. Toohey LA, Drew MK, Cook JL et al (2017) Is subsequent lower limb injury associated with previous injury? A systematic review and meta-analysis. Br J Sports Med 51:1670–1678. https://doi.org/10.1136/bjsports-2017-097500
29. Harris PA, Taylor R, Thielke R et al (2009) Research electronic data capture (REDCap)—a metadata-driven methodology and workflow process for providing translational research informatics support. J Biomed Inf 42:377–381. https://doi.org/10.1016/j.jbi.2008.08.010
30. Harris PA, Taylor R, Minor BL et al (2019) The REDCap consortium: Building an international community of software platform partners. J Biomed Inf 95:103208. https://doi.org/10.1016/j.jbi.2019.103208
31. Craig CL, Marshall AL, Sjöström M et al (2003) International physical activity questionnaire: 12-country reliability and validity. Med Sci Sports Exerc 35:1381–1395
32. Ohnhaus EE, Adler R (1975) Methodological problems in the measurement of pain: a comparison between the verbal rating scale and the visual analogue scale. Pain 1:379–384. https://doi.org/10.1016/0304-3959(75)90075-5
33. Fairbank JC, Pynsent PB (2000) The oswestry disability index. Spine 25:2940–2953. https://doi.org/10.1097/00007632-200011150-00017
34. Liu C, Ran J, Morelli JN et al (2023) Determinants of diurnal variation in lumbar intervertebral discs and paraspinal muscles: a prospective quantitative magnetic resonance imaging study. Eur J Radiol 160:110712. https://doi.org/10.1016/j.ejrad.2023.110712
35. Dimitriadis A, Papagelopoulos P, Smith F et al (2011) Intervertebral disc changes after 1 h of running: a study on athletes. J Int Med Res 39:569–579. https://doi.org/10.1177/147323001103900226
36. Pfirrmann CW, Metzdorf A, Zanetti M et al (2001) Magnetic resonance classification of lumbar intervertebral disc degeneration. Spine 26:1873–1878. https://doi.org/10.1097/00007632-200109010-00011
37. Martinez-Nicolas A, Muntaner-Mas A, Ortega FB (2017) Runkeeper: a complete app for monitoring outdoor sports. Br J Sports Med 51:1560–1561. https://doi.org/10.1136/bjsports-2016-096678
38. Adamakis M (2017) Comparing the validity of a GPS monitor and a smartphone application to measure physical activity. J Mob Technol Med 6:28–38. https://doi.org/10.7309/JMTM.6.2.4
39. Snow G (2020) Blockrand: randomization for block random clinical trials

40. Tripepi G, Chesnaye NC, Dekker FW et al (2020) Intention to treat and per protocol analysis in clinical trials. Nephrology 25:513–517. https://doi.org/10.1111/nep.13709
41. Mansournia MA, Collins GS, Nielsen RO et al (2021) A checklist for statistical assessment of medical oapers (the CHAMP statement): explanation and elaboration. Br J Sports Med 55:1009–1017. https://doi.org/10.1136/bjsports-2020-103652
42. Torén L, Lagerstrand K, Waldenberg C et al (2024) Different load-induced alterations in intervertebral discs between low back pain patients and controls: a T2-map study. Spine. https://doi.org/10.1097/BRS.0000000000005028
43. Hebelka H, Torén L, Lagerstrand K, Brisby H (2018) Axial loading during MRI reveals deviant characteristics within posterior IVD regions between low back pain patients and controls. Eur Spine J 27:2840–2846. https://doi.org/10.1007/s00586-018-5774-y
44. Takeoka Y, Kanda Y, Kang JD, Mizuno S (2023) Regenerative capability of human nucleus pulposus cells in degenerated disc under hydrostatic pressure mimicking physiologically relevant intradiscal pressure in vitro. Spine 48:728–736. https://doi.org/10.1097/brs.0000000000004530
45. Mitchell UH, Beattie PF, Bowden J et al (2017) Age-related differences in the response of the L5-S1 intervertebral disc to spinal traction. Musculoskelet Sci Pract 31:1–8. https://doi.org/10.1016/j.msksp.2017.06.004
46. Kanna RM, Hajare S, Thippeswamy PB et al (2022) Advanced disc degeneration, bi-planar instability and pathways of peri-discal gas suffusion contribute to pathogenesis of intradiscal vacuum phenomenon. Eur Spine J 31:755–763. https://doi.org/10.1007/s00586-022-07122-0
47. D'Anastasi M, Birkenmaier C, Schmidt GP et al (2011) Correlation between vacuum phenomenon on CT and fluid on MRI in degenerative disks. AJR Am J Roentgenol 197:1182–1189. https://doi.org/10.2214/AJR.10.6359
48. Sobajima S, Kompel JF, Kim JS et al (2005) A slowly progressive and reproducible animal model of intervertebral disc degeneration characterized by MRI, X-ray, and histology. Spine 30:15–24. https://doi.org/10.1097/01.brs.0000148048.15348.9b
49. Capogna EA, Brown E, Walrath E et al (2022) ISSLS prize in bioengineering science 2022: low rate cyclic loading as a therapeutic strategy for intervertebral disc regeneration. Eur Spine J 31:1088–1098. https://doi.org/10.1007/s00586-022-07239-2
50. Belavỳ DL, Quittner MJ, Ridgers N et al (2017) Running exercise strengthens the intervertebral disc. Sci Rep 7:1–8. https://doi.org/10.1038/srep45975
51. Mitchell UH, Bowden JA, Larson RE et al (2020) Long-term running in middle-aged men and intervertebral disc health, a cross-sectional pilot study. PLoS ONE 15:e0229457. https://doi.org/10.1371/journal.pone.0229457
52. Kraft CN, Pennekamp PH, Becker U et al (2009) Magnetic resonance imaging findings of the lumbar spine in elite horseback riders: correlations with back pain, body mass index, trunk/leg-length coefficient, and riding discipline. Am J Sports Med 37:2205–2213. https://doi.org/10.1177/0363546509336927
53. Leacox A, Fashingbauer L, Ferguson T et al (2025) The effect of running speed on cadence and running kinetics. Int J Sports Phys Ther 20:957–963. https://doi.org/10.26603/001c.140544
54. Schubert AG, Kempf J, Heiderscheit BC (2014) Influence of stride frequency and length on running mechanics: a systematic review. Sports Health 6:210–217. https://doi.org/10.1177/1941738113508544
55. Tessutti V, Ribeiro AP, Trombini-Souza F, Sacco ICN (2012) Attenuation of foot pressure during running on four different surfaces: asphalt, concrete, rubber, and natural grass. J Sports Sci 30:1545–1550. https://doi.org/10.1080/02640414.2012.713975
56. Bonnaerens S, Fiers P, Galle S et al (2021) Relationship between duty factor and external forces in slow recreational runners. BMJ Open Sport Exerc Med 7:e000996. https://doi.org/10.1136/bmjsem-2020-000996

Authors and Affiliations

Claire L Samanna[1,2] · Christopher Neason[1,2] · Scott D Tagliaferri[3,4] · Daniel L Belavý[5] · Ulrike H Mitchell[5] · Hanna Rae Nez[6] · Paul Buntine[7,8] · Clint T Miller[9] · David Scott[8,9] · Niamh L Mundell[9] · Matthew J Clarkson[10] · Patrick J Owen[1,7]

✉ Claire L Samanna
claire.samanna@monash.edu

Christopher Neason
christopher.neason@monash.edu

Scott D Tagliaferri
scott.tagliaferri@unimelb.edu.au

Daniel L Belavý
daniel.belavy@hs-bochum.de

Ulrike H Mitchell
rike@byu.edu

Hanna Rae Nez
hannaraenez@gmail.com

Paul Buntine
paul.buntine@monash.edu

Clint T Miller
c.miller@deakin.edu.au

David Scott
d.scott@deakin.edu.au

Niamh L Mundell
niamh.mundell@deakin.edu.au

Matthew J Clarkson
matthew.clarkson@vu.edu.au

Patrick J Owen
patrick.owen@easternhealth.org.au

1 Eastern Health Clinical School, Monash University, Box Hill, Melbourne, Australia

2 Institute for Physical Activity and Nutrition, School of Exercise and Nutrition Sciences, Deakin University, VictoriaGeelong, Australia

3 Centre for Youth Mental Health, University of Melbourne, Melbourne, Australia

4 Orygen, Parkville, Australia

5 Department of Exercise Sciences, Brigham Young University, Provo, United States

6 Integrated Pain Specialists, Utah, United States

7 Eastern Health Emergency Medicine Program, Eastern Health, Box Hill, Australia

8 School of Clinical Sciences at Monash Health, Monash University, Melbourne, Australia

9 Institute for Physical Activity and Nutrition, Deakin University, Burwood, Australia

10 Institute for Health and Sport, Victoria University, Melbourne, Australia