ARCHIVES OF THE MEDICINE AND CASE REPORTS
e-ISSN 2747-2051
hmpublisher.com/index.php/amcr · Vol. 7 No. 4 (2026) · pp. 387-406 · Article ID: 936
DOI: 10.37275/amcr.v7i4.936

SYSTEMATIC REVIEW AND META-ANALYSIS

Low Back Pain in Runners: A Systematic Review and Meta-Analysis of Prevalence, Incidence and the Share of Running-Related Injuries

Bagus Gede Krisna Astayogi[1*], I Nyoman Dharma Mukti Wisesa Jayamahe[1], I Wayan Agus Sandhika Nugraha[1]
[1] Medical Faculty, Universitas Mahasaraswati, Denpasar, Indonesia

ARTICLE INFO

Article history
Received: July 30, 2026
Revised: August 20, 2026
Accepted: September 3, 2026
Available Online: September 8, 2026
Published: September 9, 2026
Keywords
Athletic injuries; Low back pain; Prevalence; Running; Sports medicine
Corresponding author
Bagus Gede Krisna Astayogi
Email
astayogiortho@unmas.ac.id
ORCID
0009-0001-7907-8040
DOI
https://doi.org/10.37275/amcr.v7i4.936

ABSTRACT

Background: Running is widely practised, but the effect of habitual running on the lumbar spine remains contested. No quantitative synthesis of low back pain in runners existed, and published estimates ranged from below 1% to above 40%.

Objective: To estimate pooled prevalence and incidence of low back pain among runners, the low-back share of running-related injuries, heterogeneity, and small-study effects.

Methods: PubMed/MEDLINE, Europe PMC and Crossref were searched for primary studies reporting a low-back count against a defined denominator; reporting followed PRISMA 2020. Twenty-five studies from sixteen countries were included. Three denominator-specific outcomes were pooled separately using random-intercept logistic regression, and small-study effects were assessed with Peters regression.

Results: The low back accounted for 4.88% of running-related injuries (95% CI 3.29 to 7.17; k = 7). Pooled prevalence was 10.11% (95% CI 4.48 to 21.24; k = 10; 10,817 runners) and incidence was 2.44% (95% CI 1.17 to 5.04; k = 11; 9,682 runners). Log sample size explained 48.6% and 57.6% of between-study variance; estimates in cohorts of at least 500 runners were 4.07% and 1.32%. Apparent higher prevalence in elite and ultra-distance cohorts did not persist after adjustment for study size.

Conclusion: About one running injury in twenty is located at the low back, and approximately one runner in twenty-five reports low back pain in large cohorts. Small-study effects inflate published estimates; low back pain in a runner should not be attributed to running by default.

All authors have reviewed and approved the final version of the manuscript.

1. Introduction

Low back pain is the single largest contributor to years lived with disability worldwide, and its global burden has continued to rise with population growth and ageing.[1] Against that background, habitual running occupies an unusual position. It is one of the most accessible and most widely practised forms of vigorous physical activity, it is repeatedly recommended for cardiovascular and metabolic health, and yet the lumbar spine is one of the few anatomical regions for which the consequences of running remain genuinely contested among clinicians. Runners presenting to sports medicine, orthopaedic and physiotherapy clinics with low back pain are commonly advised either to reduce their mileage or to stop running altogether, and that advice is given in the absence of any pooled estimate of how often low back pain actually occurs in this population.[2]

The descriptive literature is not small. Running-injury epidemiology has produced large surveys and prospective cohorts for four decades, from clinic-based series[3] and mass-participation event surveys[4,5] to modern surveillance cohorts using standardised health-problem questionnaires.[6-8] Most of these studies report the anatomical distribution of injury, and most therefore contain a low back stratum. Alongside them sits a smaller set of studies designed specifically to measure low back pain in runners, including a national survey of 2,539 Italian runners and a post-race survey of 800 marathon competitors.[9,10] Yet the two literatures have never been brought together quantitatively, and the numbers they report are strikingly discordant: reported values range from 0.3% to well above 40%, a spread that cannot be reconciled by clinical reasoning alone.[2,11,12]

Part of that discordance is definitional. Studies differ in whether a case is any self-reported episode of low back pain or only an injury severe enough to interrupt training for a defined period; in whether the observation window is a single race, six weeks, twelve months or a lifetime; and in whether the anatomical label is the lumbar spine specifically or a composite region that also captures the buttock, sacrum, pelvis or hip.[6,13,14] A further and less widely appreciated source of discordance is structural. Some studies express the low back count as a proportion of runners, and others as a proportion of all recorded running-related injuries. These are different quantities with different denominators, and pooling them together produces a number that means nothing. Comparable problems have been described in other sports, where low back pain has now been synthesised quantitatively for soccer and appraised across athletic populations more broadly.[15-17]

A third source of discordance is arithmetic rather than clinical, and has attracted almost no attention. Estimates of occurrence derived from small convenience samples are systematically higher than those derived from large population surveys, because small runner cohorts are usually assembled in a clinic, a specialist race or a national squad and are therefore enriched for symptomatic athletes. If that gradient is strong enough, it will dominate any clinical moderator a synthesis might examine, and a review that reports subgroup differences without first testing for it risks presenting an artefact of sampling as a biological finding.

The novelty of this study lies in being the first quantitative synthesis of low back pain in runners: no meta-analysis of this outcome exists in this population, the only prior systematic review reported ranges rather than pooled estimates and pooled nothing, no previous work has separated the runner denominator from the injury denominator, which this review shows to be the structural error that makes the published literature appear irreconcilable, and no previous work has

tested whether the apparent clinical variation in this field survives adjustment for study size, which this review finds it does not. The aim of this study was to estimate the pooled prevalence and incidence of low back pain among runners and the share of running-related injuries located at the low back, to quantify and explain the heterogeneity between studies, and to determine how much of the variation in published estimates is attributable to study size rather than to any clinical characteristic of the runners themselves.

2. Methods

2.1 Design and reporting

A systematic review with meta-analysis of proportions was undertaken. The review was designed, conducted and reported in accordance with the Preferred Reporting Items for Systematic Reviews and Meta-Analyses 2020 statement.[18] The review was not prospectively registered. No ethical approval was required, because the analysis used aggregate data already published in the peer-reviewed literature and no individual participant data were obtained.

2.2 Eligibility criteria

Studies were eligible if they were peer-reviewed primary research articles reporting a count of runners with low back pain, lumbar pain or a low back injury against an explicitly defined denominator. The population was adult and adolescent runners of any competitive standard, including novice participants in structured beginner programmes, recreational and club runners, marathon and half-marathon competitors, ultra-trail runners and elite track and road athletes. Studies of non-running athletes, of mixed-sport cohorts in which the running subgroup could not be separated, and of clinical populations recruited for the treatment of established back pain rather than for their running exposure were excluded. Case reports, narrative reviews, systematic reviews, scoping reviews, editorials, conference abstracts, preprints, theses, animal experiments and modelling studies were excluded. No language or date restriction was applied.

2.3 Information sources and search

PubMed/MEDLINE was searched through the E-utilities interface and Europe PMC through its REST interface, the latter chosen because it indexes journals, preprint servers and theses outside MEDLINE and therefore captures regional sports medicine titles that a MEDLINE-only search misses. Search terms combined running vocabulary (runner, running, marathon, ultramarathon, trail running, jogging) with low back vocabulary (low back pain, lower back pain, lumbar, lumbopelvic, back injury, back pain) and, in the filtered passes, with epidemiological vocabulary (prevalence, incidence, epidemiology, cross-sectional, cohort, survey, questionnaire, injury location, anatomical site). The reference lists of the single prior synthesis of this literature and of all included reports were harvested. Crossref was used to resolve and confirm bibliographic records and was not used as a discovery source. Because the dedicated literature on low back pain in runners is small, the search was deliberately widened to running-injury epidemiology cohorts that report anatomical injury location, since these contain a low back stratum against a runner denominator even when low back pain is not the stated topic of the paper.

After the first analysis was complete the search was re-executed with deliberately broadened concept blocks and with the epidemiological filter removed, and separately across all sources indexed by Europe PMC, in order to test whether the reach of a two-database strategy was adequate for the completeness claim. The strategies, their yields and their contribution to the included set are given in Table 1. Three of the twenty-five included studies are published in journals that return no record in a MEDLINE-only search and were reached through Europe PMC and reference harvesting; as the contribution column of Table 1 records, no additional eligible study was identified by the extended pass.

2.4 Study selection and data extraction

Records were screened on title and abstract and then at full text. Screening, data extraction and source adjudication were performed independently by two members of the review team and reconciled by discussion, with the third member arbitrating where agreement could not be reached. No formal inter-rater agreement statistic was computed, which is a limitation relative to the prior synthesis in this field. For every study that passed full-text screening, the following were extracted: first author, year, country, design, sampling frame, the operational definition of a runner, the case definition of low back pain, the anatomical label used by the authors, the observation window, the numerator and the denominator, the sex distribution, mean age, competitive standard and the risk factors reported.

Every extracted percentage was recomputed from its own numerator and denominator, and any value that did not reconcile was returned to the source document and re-extracted rather than accepted. Bibliographic records were rebuilt field by field from the live PubMed record, with Crossref as fallback, so that no author string, volume or page range was transcribed by hand. Six studies required adjudication at source. In four of them the published percentage could not be reproduced from the numerator and denominator reported alongside it, and in each case the discrepancy proved to be a denominator problem rather than an arithmetic error. In one further study the running subgroup had to be separated from a mixed-sport school cohort, and the record inherited from the earlier synthesis was found to carry both an incorrect title and an incorrect denominator. These adjudications are reported in full in the results, because they determine the structure of the analysis.

2.5 Outcomes and the denominator taxonomy

The pre-specified primary outcome was the proportion of runners affected by low back pain. During adjudication it became clear that the source studies report that proportion against two structurally different denominators. Some divide the low back count by the number of runners studied, and others divide it by the number of running-related injuries recorded. The first answers the question of how many runners are affected; the second answers the question of, given that a runner is injured, how often the low back is the site. These are not interchangeable, and pooling them together yields a quantity with no interpretation. Studies were therefore assigned to a pool according to the denominator the source actually reports, giving three analyses: pooled prevalence among runners, pooled incidence of new low back pain among runners, and the pooled low back share of running-related injuries. Prevalence and incidence were never combined, since a period prevalence and an incidence over a defined observation window estimate different quantities. Assignment to the prevalence or the incidence pool followed the quantity the source measured rather than the word it used: a study entered the incidence pool where the outcome was a new episode arising during a defined observation period, and the prevalence pool where it was an

episode existing at, or recalled at, a single point. Two studies required a judgement under that rule, because both report a single-occasion measurement in a cross-sectional sample and the two sources use the terms in opposite senses; the 2013 Danish marathon survey was assigned to incidence because its outcome arose during the race period, and the 2002 bridge-run survey to prevalence because its outcome was current pain at the time of the survey.[19,20]

Table 1. Search strategies, yields and contribution to the included set.

Pass	Source	Concept blocks	Records	Contribution
Primary	PubMed/MEDLINE (E-utilities)	Running x low back x epidemiology, with design exclusions	135	Screened; 3 new studies identified with the Europe PMC pass
Primary	Europe PMC (REST)	Running x low back x epidemiology	149	Screened; reached the three non-MEDLINE journals in the included set
Primary	Reference harvesting	Prior synthesis and all included reports	22 studies	Carried forward to the verified source pool
Extended	PubMed/MEDLINE	Running x low back, broadened terms, no epidemiological filter	871	No additional eligible study
Extended	PubMed/MEDLINE	As above, with epidemiological filter and design exclusions	221	No additional eligible study
Extended	Europe PMC, all indexed sources	Running x low back x epidemiology, title and abstract	238	No additional eligible study
Extended	Europe PMC, sources outside MEDLINE only	As above	9	None eligible: 7 preprints, 1 conference abstract, 1 thesis
Supporting	Crossref REST API	Bibliographic resolution by DOI and by title	n/a	Record verification only; not a discovery source

The extended pass was added in response to peer review. It confirmed that no eligible peer-reviewed primary study had been missed, and its yields are reported so that the search can be reproduced.

The three pools are not disjoint. Three studies report both a runner denominator and an injury denominator and therefore contribute to a primary pool and to the secondary pool, under different denominators and with different numerators. No study appears twice within any single pool, so the pooled estimates are unaffected, but the three pool totals must not be summed, and any statement comparing one pool with another is made across partly overlapping evidence. The three studies concerned are identified in the characteristics table.

2.6 Risk of bias

Risk of bias was appraised with the Joanna Briggs Institute critical appraisal checklist for prevalence studies, covering the appropriateness of the sampling frame, the sampling method, sample size, description of setting and participants, coverage of the identified sample, validity and standardisation of outcome measurement, appropriateness of the statistical analysis and adequacy of the response rate. For the prospective and retrospective cohort designs the corresponding ROBINS-E domains were mapped onto the same nine-domain frame so that a single traffic-light display could be produced.[21] No included study was a randomised trial, so the Cochrane risk-of-bias tool for randomised trials was not applicable and was not used. An overall judgement of low risk, some concerns or high risk was assigned to each study.

2.7 Statistical analysis

Proportions were pooled using random-intercept logistic regression, that is a generalised linear mixed model with a logit link fitted by maximum likelihood. This model was chosen in preference to the Freeman-Tukey double-arcsine transformation because the denominators in this evidence base span 31 to 4,380 runners, a 140-fold range over which the double-arcsine transformation is unstable and back-transformation is ill-behaved; the mixed model also handles the single-event study in the incidence pool without a continuity correction. Study-level proportions are reported with exact binomial confidence intervals, whereas the pooled interval is a back-transformed interval from the mixed model; the two are constructed differently and their widths are not directly comparable. Heterogeneity is reported as the I^2 statistic with its 95% confidence interval, as τ^2 on the logit scale, and as a 95% prediction interval, the last of these being the interval within which the true proportion of a new comparable study would be expected to fall.

A pooled point estimate is reported for each pool notwithstanding the heterogeneity, for two reasons. First, no summary of this literature currently exists at all, and a point estimate with an honest prediction interval is more informative than the range of raw values that has stood in its place. Second,

and more importantly, the pooled estimate is required as the anchor for the meta-regression and small-study analyses that constitute this review's substantive contribution; it is reported for that purpose rather than as a definitive summary of the field, and it should always be read together with its prediction interval. That interval carries its own caveat: with τ^2 near two on the logit scale and ten studies, the logit-normal distributional assumption underlying it cannot meaningfully be tested, so an upper limit implying that three-quarters of runners might be affected should be read as a warning about the spread of this evidence base rather than as a plausible value.

Subgroup analyses were pre-specified for the recall window or observation period, the anatomical label (lumbar-specific versus composite regions), the case ascertainment method (any self-reported low back pain versus an injury qualifying under a time-loss definition), competitive standard, publication era, overall risk of bias and study design. Between-group differences were tested by the omnibus moderator test within a mixed-effects model, and the proportion of between-study variance explained is reported as R^2. Where a moderator produced a level containing fewer than two studies, the level-specific estimate is reported but the omnibus test is suppressed, because a between-group test across such a partition is not interpretable. Fourteen moderator tests were performed in total, seven in each primary pool; the nominal p values are not adjusted for multiplicity, and under a Bonferroni correction the threshold would be approximately 0.0036, which none of the nominally significant findings would meet. These analyses are therefore reported as exploratory throughout.

Univariable meta-regression was performed on publication year and on the natural logarithm of sample size, and multivariable models were fitted to determine whether any clinical moderator survived adjustment for study size; models were compared using the corrected Akaike information criterion. Publication era and publication year are treated as proxies rather than mechanisms in this evidence base, because the two pre-2000 studies in the prevalence pool are also the largest and third largest, so era and study size are collinear by construction.

Small-study effects were assessed with Peters regression of the log odds on the inverse of sample size. The Egger regression test is not valid for proportions, because the standard error of a proportion is intrinsically related to the proportion itself, and it was therefore not performed. Peters regression is generally recommended when at least ten studies are available; the prevalence pool contains exactly ten, so the test operates at the boundary of its recommended use and its result is interpreted for direction and significance rather than for the magnitude of its coefficient, which is on the scale of the reciprocal of sample size and is not interpretable in absolute terms. Funnel plots are presented on the logit scale with pseudo confidence contours.

Because small-study effects were substantial, two further procedures were run. The meta-regression on log sample size was used to generate predicted proportions at fixed sample sizes within the observed range, and the mixed model was refitted after restricting to cohorts above size thresholds. Both are sensitivity procedures for small-study effects and are conditional predictions under an assumed linear relationship on the logit scale; neither is a bias-corrected estimate of the population proportion, and they are labelled and interpreted accordingly throughout. Pre-specified sensitivity analyses excluded studies with composite anatomical labels, studies with derived or re-extracted numerators, studies published before 2000, studies at high risk of bias and studies of fewer than 100 participants, and a full leave-one-out analysis was run for both primary pools. Certainty of the evidence was rated using the GRADE approach adapted for prevalence data, starting at high for observational estimates of occurrence and rating down for risk of bias, inconsistency, indirectness, imprecision and publication bias.

3. Results

3.1 Study selection

The primary search identified 284 records across PubMed/MEDLINE and Europe PMC, which were screened alongside 22 studies carried forward from the verified source pool. Of the 284 records, 271 were excluded on title and abstract because the population was not a runner cohort, because no low back stratum was reported, or because the report was a review or case report. Thirteen reports were assessed at full text, of which ten were excluded: six reported no low back count against any denominator, two described cohorts already represented in the pool, and two reported lower-limb sites only. Three new studies met eligibility and were added: a survey of 697 participants in the Ljubljana Marathon, a cross-sectional survey of 167 professional endurance runners in the Rift Valley of Kenya, and a 30-week prospective surveillance cohort of South African trail runners.[8,12,22] The extended search added after peer review returned 871 records in PubMed/MEDLINE with broadened blocks and 238 in Europe PMC across all indexed sources, of which nine lay outside MEDLINE; those nine comprised seven preprints, one conference abstract and one doctoral thesis, none of which met the eligibility criteria. Twenty-five studies were therefore included in the review. The flow of records is set out in Figure 1.

3.2 Characteristics of the included studies

The 25 included studies were published between 1981 and 2021 and were conducted in sixteen countries on five continents. Ten studies contributed to the prevalence pool, comprising 10,817 runners and 928 cases; eleven contributed to the incidence pool, comprising 9,682 runners and 162 cases; and seven contributed to the secondary pool, comprising 3,503 running-related injuries and 164 low back injuries. The two primary pools draw on disjoint sets of studies, so the review summarises 20,499 runners across 21 studies with a runner denominator; the four studies that report only an injury denominator contribute no runner count. Three studies appear in both a primary and the secondary pool under different denominators and are flagged as such in the table; the three totals are therefore not additive. Designs were mixed: thirteen prospective or retrospective cohorts, ten cross-sectional surveys and two clinic series. Denominators ranged from 31 adolescent elite runners to 4,380 marathon entrants, a 140-fold spread. The prevalence pool comprised ten studies,[4,5,9,11,12,14,20,22-24] the incidence pool eleven,[6,7,10,13,19,25-30] and the secondary injury-denominator pool seven.[3,8,22,28,30-32] Study-level characteristics, counts and exact binomial confidence intervals are given in Table 2.

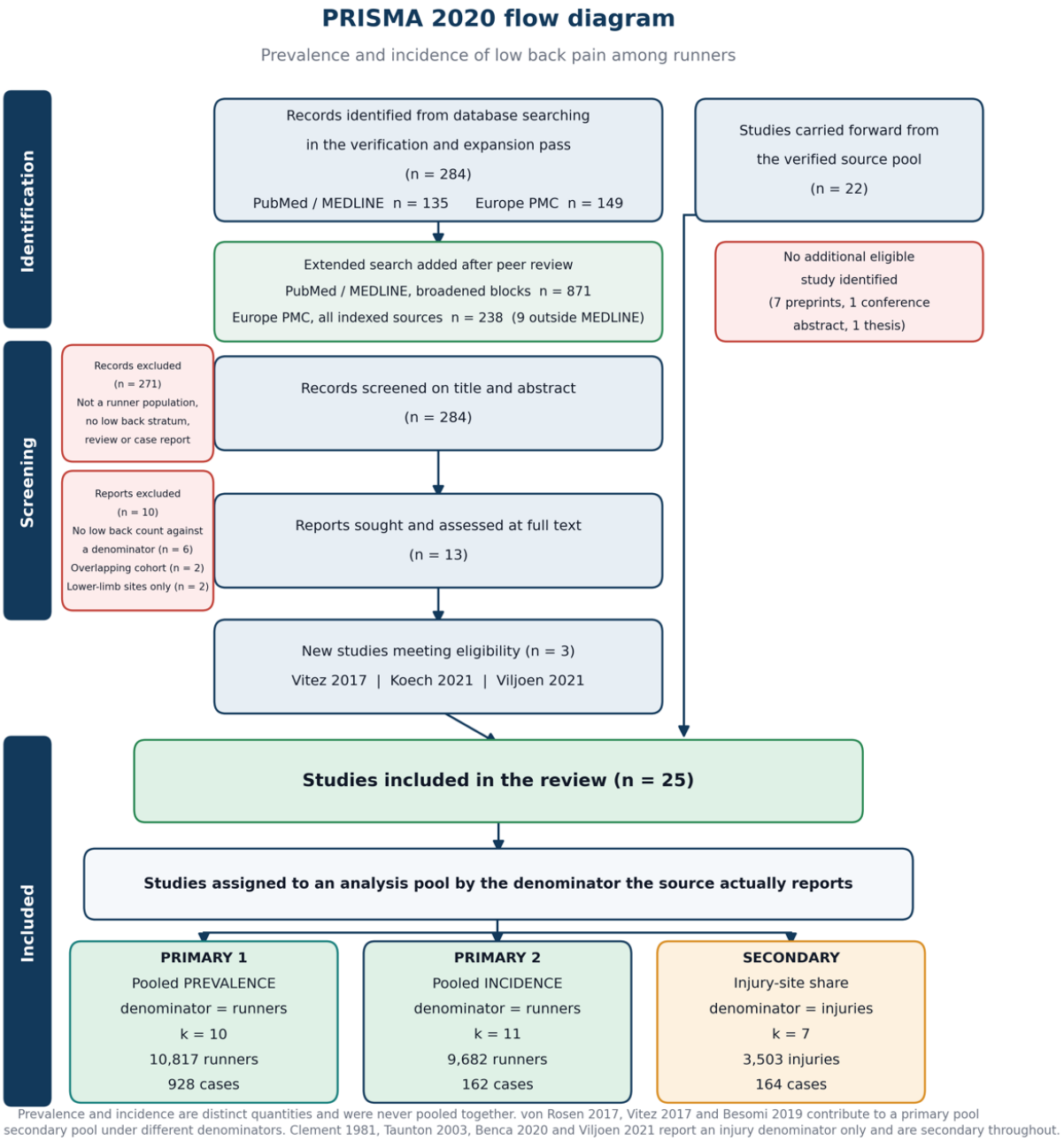

Figure 1. Study selection, shown as a PRISMA 2020 flow diagram. The three analysis pools are defined by the denominator each source study actually reports, not by the question the authors set out to answer. Three studies contribute to a primary pool and to the secondary pool under different denominators, and four studies report an injury denominator only. The extended search added after peer review identified no additional eligible study.

The operational definition of a runner varied by two orders of magnitude of exposure, from someone running two miles three days a week to a professional training on up to 160 km a week. Case ascertainment differed systematically as well: four studies counted any self-reported episode of low back pain, while the remainder required an injury qualifying under a time-loss definition, most often an episode severe enough to reduce or interrupt running for at least one day, one week or two weeks.[9,10,13,19,20,24] Six studies used a composite anatomical label that also captured the buttock, sacrum, pelvis or hip rather than a lumbar-specific label.[4-6,14,27,29] Observation windows ranged from a single race to two years of prospective surveillance, as the window column of Table 2 shows. Because the entry criterion and the case definition together determine what a pooled proportion means, both are set out for every study in Table 3.

Table 2. Characteristics, counts and study-level proportions of the 25 included studies, grouped by analysis pool.

Study	Country	Design	Pool	Total n (share of pool %)	Cases	Proportion % (95% CI)	Window	Region label	Risk of bias
Marti B 1988	Switzerland	Cross-sectional	Prevalence	4358 (40.3)	30	0.69 (0.46-0.98)	12 months	Composite	Some concerns
Walter SD 1989	Canada	Prospective cohort	Prevalence	1288 (11.9)	56	4.35 (3.30-5.61)	12 months	Composite	Some concerns
Woolf SK 2002	USA	Cross-sectional	Prevalence	436 (4.0)	59	13.53 (10.46-17.11)	Point	Lumbar	Some concerns
Chang WL 2012	Taiwan	Cross-sectional	Prevalence	893 (8.3)	29	3.25 (2.19-4.63)	Lifetime	Lumbar	Some concerns
Ellapen TJ 2013	South Africa	Retrospective descriptive	Prevalence	200 (1.8)	28	14.00 (9.51-19.59)	12 months	Composite	High risk
Malliaropoulos N 2015	Greece	Cross-sectional	Prevalence	40 (0.4)	17	42.50 (27.04-59.11)	Lifetime	Lumbar	High risk
Teixeira RN 2016	Brazil	Cross-sectional	Prevalence	199 (1.8)	28	14.07 (9.56-19.69)	12 months	Composite	Low risk
Vitez L 2017 [d]	Slovenia	Retrospective cohort	Prevalence	697 (6.4)	30	4.30 (2.92-6.09)	Lifetime	Lumbar	Some concerns
Koech RC 2021	Kenya	Cross-sectional	Prevalence	167 (1.5)	78	46.71 (38.96-54.57)	12 months	Lumbar	Some concerns
Maselli F 2021	Italy	Cross-sectional	Prevalence	2539 (23.5)	573	22.57 (20.95-24.24)	12 months	Lumbar	Low risk
Bach DK 1985	USA	Prospective cohort	Incidence	45 (0.5)	10	22.22 (11.20-37.09)	12 months	Lumbar	High risk
Lysholm J 1987	Sweden	Prospective cohort	Incidence	60 (0.6)	3	5.00 (1.04-13.92)	12 months	Lumbar	High risk
Buist I 2010	Netherlands	Prospective cohort	Incidence	629 (6.5)	31	4.93 (3.37-6.92)	8 weeks	Composite	Low risk
Rasmussen CH 2013	Denmark	Retrospective cohort	Incidence	662 (6.8)	3	0.45 (0.09-1.32)	Race period	Lumbar	Some concerns
Kluitenberg B 2015	Netherlands	Prospective cohort	Incidence	1696 (17.5)	6	0.35 (0.13-0.77)	6 weeks	Composite	Low risk
van der Worp MP 2016	Netherlands	Prospective cohort	Incidence	373 (3.9)	10	2.68 (1.29-4.88)	12 weeks	Composite	Low risk
von Rosen P 2017 [d]	Sweden	Prospective cohort	Incidence	31 (0.3)	1	3.23 (0.08-16.70)	52 weeks	Lumbar	Some concerns
Messier SP 2018	USA	Prospective cohort	Incidence	300 (3.1)	18	6.00 (3.59-9.32)	2 years	Composite	Low risk
Besomi M 2019 [d]	Chile	Cross-sectional	Incidence	4380 (45.2)	31	0.71 (0.48-1.00)	Race period	Lumbar	Low risk
Dallinga J 2019	Netherlands	Prospective cohort	Incidence	706 (7.3)	13	1.84 (0.98-3.13)	12 weeks	Lumbar	Low risk
Wu B 2021	China	Cross-sectional	Incidence	800 (8.3)	36	4.50 (3.17-6.18)	Race period	Lumbar	Some concerns
Clement DB 1981	Canada	Clinic series	Injury-site share	1819 (51.9)	68	3.74 (2.91-4.72)	2 years	Lumbar	High risk
Taunton JE 2003	Canada	Prospective cohort	Injury-site share	249 (7.1)	4	1.61 (0.44-4.06)	13 weeks	Lumbar	Some concerns
von Rosen P 2017 [d]	Sweden	Prospective cohort	Injury-site share	36 (1.0)	1	2.78 (0.07-14.53)	52 weeks	Lumbar	Some concerns
Vitez L 2017 [d]	Slovenia	Retrospective cohort	Injury-site share	375 (10.7)	30	8.00 (5.46-11.22)	Lifetime	Lumbar	Some concerns
Besomi M 2019 [d]	Chile	Cross-sectional	Injury-site share	620 (17.7)	31	5.00 (3.42-7.02)	Race period	Lumbar	Low risk
Benca E 2020	Austria	Clinic series	Injury-site share	199 (5.7)	20	10.05 (6.25-15.09)	Clinic series	Lumbar	High risk
Viljoen CT 2021	South Africa	Prospective cohort	Injury-site share	205 (5.9)	10	4.88 (2.36-8.79)	30 weeks	Composite	Low risk

The figure in brackets after the denominator is the study's percentage share of its own pool denominator. Within the secondary pool the denominator is injury events for Clement 1981, Taunton 2003, von Rosen 2017, Benca 2020 and Viljoen 2021, and injured runners for Vitez 2017 and Besomi 2019. [d] contributes to a primary pool and to the secondary pool under different denominators; the three pool totals must not be summed. Region label: lumbar = a lumbar-specific anatomical label; composite = a region that also captures the buttock, sacrum, pelvis or hip. Total n is the denominator: runners for the prevalence and incidence pools, and running-related injuries for the secondary pool.

Table 3. Operational definition of a runner, reported training exposure and case definition, for each included study.

Study	Definition of a runner used by the authors	Reported exposure	Case definition	Ascertainment
Marti B 1988	Participant in a mass-participation 16 km running event	Median 25 km per week	Self-reported jogging injury in the previous 12 months, graded I to III	Injury-based
Walter SD 1989	Registered entrant of a 4 to 22.4 km road race	[NR]	Self-reported running-related injury	Injury-based
Woolf SK 2002	Any standard, from novice to competitive athlete	[NR]	Self-reported low back pain	Symptom-based
Chang WL 2012	Running at least 30 minutes twice weekly	[NR]	Self-reported running injury; severity not defined	Injury-based
Ellapen TJ 2013	Has run at least a half marathon (21.1 km)	Mean 12.2 years of road running	Distress or agony preventing physical activity for at least 24 hours	Injury-based
Malliaropoulos N 2015	Ultra-trail runner per the ITRA definition; distance beyond 42.195 km	[NR]	Symptoms severe enough to forgo training for at least 1 day or to quit a race	Injury-based
Teixeira RN 2016	Competing at national or international level	Up to 160 km per week	Self-reported running-related musculoskeletal pain in the previous 12 months	Symptom-based
Vitez L 2017	Registered participant of the 18th Ljubljana Marathon (10 km, half or full)	[NR]	Self-reported running injury causing absence of 2 weeks or more	Injury-based
Koech RC 2021	Professional endurance runner competing locally and internationally	[NR]	Musculoskeletal injury of the lower limb or back causing more than 48 hours out of running	Injury-based
Maselli F 2021	Registered member of a national running association	[NR]	Self-reported low back pain in the previous 12 months	Symptom-based
Bach DK 1985	Runs 18 or more miles per week, regularly, for at least 1 year	At least 29 km per week	Injury severe enough to temporarily stop running	Injury-based
Lysholm J 1987	Previous running training experience of 7 hours per week or more	7 or more hours per week	Injury markedly hampering training or competition for at least 1 week	Injury-based
Buist I 2010	Self-categorised novice, returning or regular runner	10 to 60 minutes per session, 8-week programme	Musculoskeletal pain of the lower limb or back restricting running for at least 1 day	Injury-based
Rasmussen CH 2013	Completion of the H.C. Andersen marathon	Stratified by weekly volume before the race	Running injury reducing distance, speed, duration or frequency for at least 14 days	Injury-based
Kluitenberg B 2015	Novice runner enrolled in a Start to Run programme	6-week supervised programme	Lower extremity or back complaint attributed to running, hampering three consecutive sessions	Injury-based
van der Worp MP 2016	Adult woman registered for the Marikenloop 5 km or 10 km event	[NR]	Running-related pain of the low back or lower extremity restricting running for at least 1 day	Injury-based
von Rosen P 2017	Elite adolescent endurance athlete in the running discipline	Sport-school training load	Any physical complaint affecting normal training or competition (OSTRC)	Injury-based
Messier SP 2018	Runs a minimum of 5 miles per week	At least 8 km per week	Overuse running injury graded 1 to 3 by the Marti method	Injury-based
Besomi M 2019	Competed in one of three race distances (10, 21 or 42 km)	Stratified by pre-competition weekly volume	Injury caused by running, reducing running for at least 7 days	Injury-based
Dallinga J 2019	Registered for the 8 km or 16 km distance	12-week preparation period	Any physical complaint causing at least one week of training loss	Injury-based
Wu B 2021	Ran a half or full marathon in the index event	[NR]	Self-reported low back pain after the competition	Symptom-based
Clement DB 1981	Running at least 2 miles (3 km) three days per week at the time of injury	At least 3 km, three days per week	Physician-diagnosed running-related injury	Injury-based
Taunton JE 2003	Novice or intermediate participant in a 13-week structured clinic	13-week structured programme	Injury recorded during the 13-week programme	Injury-based
Benca E 2020	Patient presenting to a sports medicine clinic with a running-related injury	[NR]	Clinician-diagnosed running-related injury	Injury-based
Viljoen CT 2021	Trail runner followed biweekly for 30 weeks with a modified OSTRC health-problem questionnaire	Recorded biweekly as training hours	Any running-related health problem reported on the OSTRC questionnaire	Injury-based

Ascertainment: symptom-based = any self-reported episode of low back pain; injury-based = an episode qualifying under the study's time-loss or clinical injury definition. [NR] = not reported.

The counts and proportions that resulted from that adjudication are the ones shown in Table 2. Six studies required adjudication against the source document because the published percentage could not be reproduced from its own numerator and denominator. In four the discrepancy was structural rather than arithmetic. The 1981 clinic survey reports 68 low back injuries among 1,819 injuries recorded in 1,650 attendees, and its published figure of 3.7% reproduces exactly when the

e-ISSN 2747-2051
DOI: 10.37275/amcr.v7i4.936

denominator is injuries; it was therefore moved to the secondary pool.[3] The 2003 Vancouver clinic cohort reports its low back figure as a share of injuries, and the runner-denominator numerator could not be recovered from an image-only source table, so it was excluded from the primary analysis and carried in the secondary pool as an approximate value.[31] The ultra-trail survey was found to be correct as extracted at 17 of 40 runners, and the lower value quoted for it in the earlier synthesis is a misreading of the source figure.[11] The 2019 Chilean survey measures injury sustained during the race and was reclassified from prevalence to incidence.[30] In the adolescent elite cohort the running subgroup proved to be 31 athletes rather than the 189 recorded in the earlier synthesis, a denominator wrong by a factor of six, and the study also carried an incorrect title there; the numerator was re-extracted as one low back injury.[28] The largest dedicated survey was confirmed at 573 of 2,539 runners.[9]

3.3 Risk of bias

Of the 25 studies, nine were judged at low risk of bias, ten raised some concerns and six were at high risk. Within the prevalence pool two studies were at low risk, six raised some concerns and two were at high risk; within the incidence pool six were at low risk, three raised some concerns and two were at high risk. Reading down the domain columns of Figure 2, the domains that most often fell short were coverage of the identified sample and the adequacy of the response rate, both of which are chronic weaknesses of race-day and postal surveys, and the validity of outcome measurement, which was unclear wherever low back pain was captured by an unvalidated bespoke questionnaire. The two clinic series were judged at high risk on sampling frame and sampling method, since patients presenting to a sports medicine clinic are not a sample of runners. The domain-level appraisal is displayed in Figure 2.

Figure 2. Risk of bias across the 25 included studies, appraised with the Joanna Briggs Institute prevalence checklist and with ROBINS-E domains mapped for the cohort designs. No included study was a randomised trial.

3.4 Pooled estimates

The pooled share of running-related injuries located at the low back was 4.88% (95% CI 3.29 to 7.17) across seven studies and 3,503 injuries, with I^2 of 79.7% and a prediction interval of 1.40% to 15.63%. This was by a wide margin the most consistent of the three quantities, although 1,819 of those 3,503 injuries, 51.9%, come from a single clinic series rated at high risk of bias. Pooled prevalence of low back pain among runners was 10.11% (95% CI 4.48 to 21.24) across ten studies and 10,817 runners, and pooled incidence of new low back pain was 2.44% (95% CI 1.17 to 5.04) across eleven studies and 9,682 runners. Both primary pools showed extreme heterogeneity, at 99.0% (95% CI 98.1 to 99.7) and 94.9% (95% CI 89.7 to 98.5), and their prediction intervals spanned 0.38% to 76.88% and 0.15% to 29.89% respectively. Neither pooled point estimate should be quoted without its prediction interval. The three pools are given in Table 4 and the study-level estimates in Figure 3.

e-ISSN 2747-2051
DOI: 10.37275/amcr.v7i4.936

Table 4. Pooled estimates for the three analysis pools. Prevalence and incidence were never combined, because they estimate different quantities.

Analysis	k	Total n	Cases	Pooled % (95% CI)	95% prediction interval (%)	I^2 (%, 95% CI)	τ^2	Q
Primary 1: Prevalence among runners	10	10817	928	10.11 (4.48 to 21.24)	0.38 to 76.88	99.0 (98.1 to 99.7)	1.9580	779.21
Primary 2: Incidence among runners	11	9682	162	2.44 (1.17 to 5.04)	0.15 to 29.89	94.9 (89.7 to 98.5)	1.4233	169.98
Secondary: Low back share of injuries	7	3503	164	4.88 (3.29 to 7.17)	1.40 to 15.63	79.7 (48.2 to 97.2)	0.2061	28.94

The two primary pools draw on disjoint sets of studies and together summarise 20,499 runners across 21 studies. All models are random-intercept logistic regressions fitted by maximum likelihood. Q was significant at $p < 0.0001$ in every pool. τ^2 is on the logit scale. Study-level intervals in the accompanying figure are exact binomial intervals, whereas the pooled interval is back-transformed from the mixed model; their widths are not directly comparable.

Figure 3. Forest plots for the three pools. Panel A, prevalence among runners; panel B, incidence of new low back pain among runners; panel C, the share of running-related injuries located at the low back. Marker area is proportional to the square root of the denominator. Study-level bars are exact binomial intervals; red bars are 95% prediction intervals, and in panel A the upper prediction limit lies beyond the plotted axis.

The contrast is visible in Figure 3: the individual estimates in panel C sit within a narrow band, whereas those in panel A span almost the whole plotted axis. That the injury-site share is the most stable of the three quantities is itself informative. It suggests that what varies between studies is not where running injuries occur but how runners and their injuries are counted, and it is the reason this review treats the injury-site share as its

most transportable result despite its status as a secondary outcome.

3.5 Study size dominates both primary pools

Peters regression of the log odds on the inverse of sample size was strongly significant in both primary pools, with coefficients of 74.411 (SE 11.561; t = 6.436) for prevalence and 103.472 (SE 14.570; t = 7.102) for incidence, both at p < 0.0001. The coefficients are on the scale of the reciprocal of sample size and are not interpretable in absolute terms; only their sign and significance are used here. Smaller studies systematically reported higher proportions. The prevalence pool contains exactly ten studies, which is the lower boundary at which Peters regression is recommended, so this result is reported with that caveat. The secondary pool contains seven studies and was not tested. Funnel plots on the logit scale are shown in Figure 4.

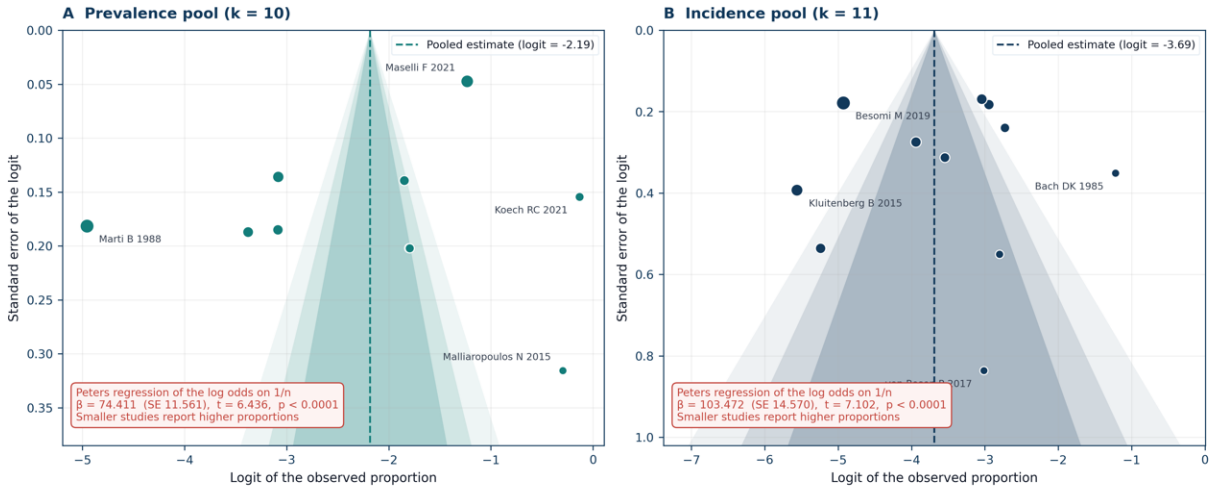

Figure 4. Contour-enhanced funnel plots on the logit scale for the two primary pools, with pseudo 90%, 95% and 99% confidence contours around the pooled estimate and the Peters regression result annotated. The Egger test is not valid for proportions and was not used.

Two sensitivity procedures were run to characterise the gradient. Predicted proportions were generated from the meta-regression on log sample size at fixed denominators within the observed range, and the mixed model was refitted after restricting to larger cohorts. The steepness of that relationship is what Figure 5 panel A displays. Modelled at 100 runners, prevalence was 28.88%; at 1,000 runners it was 6.30%; and at 2,500 runners it was 3.18% (95% CI 1.16 to 8.46). Restricting the model to the five cohorts of at least 500 runners gave 4.07% (95% CI 1.46 to 10.82), and to the three cohorts of at least 1,000 runners gave 4.32% (95% CI 0.79 to 20.36), the latter interval being too wide to be informative. The corresponding incidence figures were 6.71% at 100 runners, 1.59% at 1,000, 0.88% at 2,500 and 1.32% (95% CI 0.56 to 3.11) when restricted to cohorts of at least 500. These are conditional predictions and restricted refits, not bias-corrected estimates; what they establish is the direction and steepness of the gradient, not a corrected value for the population proportion. Results are given in Table 5 and the relationship is displayed in Figure 5.

Table 5. Small-study effects and size-related sensitivity procedures.

Pool	k	Test or procedure	Beta	SE	t	p	Result
Prevalence	10	Peters regression (log odds on 1/n, FE)	74.411	11.561	6.436	< 0.0001	Evidence of small-study effects
Incidence	11	Peters regression (log odds on 1/n, FE)	103.472	14.570	7.102	< 0.0001	Evidence of small-study effects
Injury-site share	7	Peters regression	NA	NA	NA	NA	Not performed: fewer than 10 studies
Prevalence		Size-adjusted estimate					Modelled at n = 100: 28.88% (12.41 to 53.76)
Prevalence		Size-adjusted estimate					Modelled at n = 500: 10.36% (5.61 to 18.34)
Prevalence		Size-adjusted estimate					Modelled at n = 1,000: 6.30% (3.10 to 12.39)
Prevalence		Size-adjusted estimate					Modelled at n = 2,500: 3.18% (1.16 to 8.46)
Prevalence		Size-adjusted estimate					Modelled at n = 4,380: 2.08% (0.60 to 6.96)
Prevalence		Size-adjusted estimate					GLMM, studies n>=500 (k=5): 4.07% (1.46 to 10.82)
Prevalence		Size-adjusted estimate					GLMM, studies n>=1000 (k=3): 4.32% (0.79 to 20.36)
Incidence		Size-adjusted estimate					Modelled at n = 100: 6.71% (3.18 to 13.58)
Incidence		Size-adjusted estimate					Modelled at n = 500: 2.47% (1.46 to 4.13)
Incidence		Size-adjusted estimate					Modelled at n = 1,000: 1.59% (0.86 to 2.90)
Incidence		Size-adjusted estimate					Modelled at n = 2,500: 0.88% (0.38 to 2.03)
Incidence		Size-adjusted estimate					Modelled at n = 4,380: 0.61% (0.22 to 1.69)
Incidence		Size-adjusted estimate					GLMM, studies n>=500 (k=6): 1.32% (0.56 to 3.11)

Peters regression of the log odds on the inverse of sample size; its coefficient is not interpretable in absolute terms. Size-related procedures are predictions from the meta-regression on log sample size within the observed range of denominators, or refits of the random-intercept logistic model restricted to larger cohorts. Neither is a bias-corrected estimate.

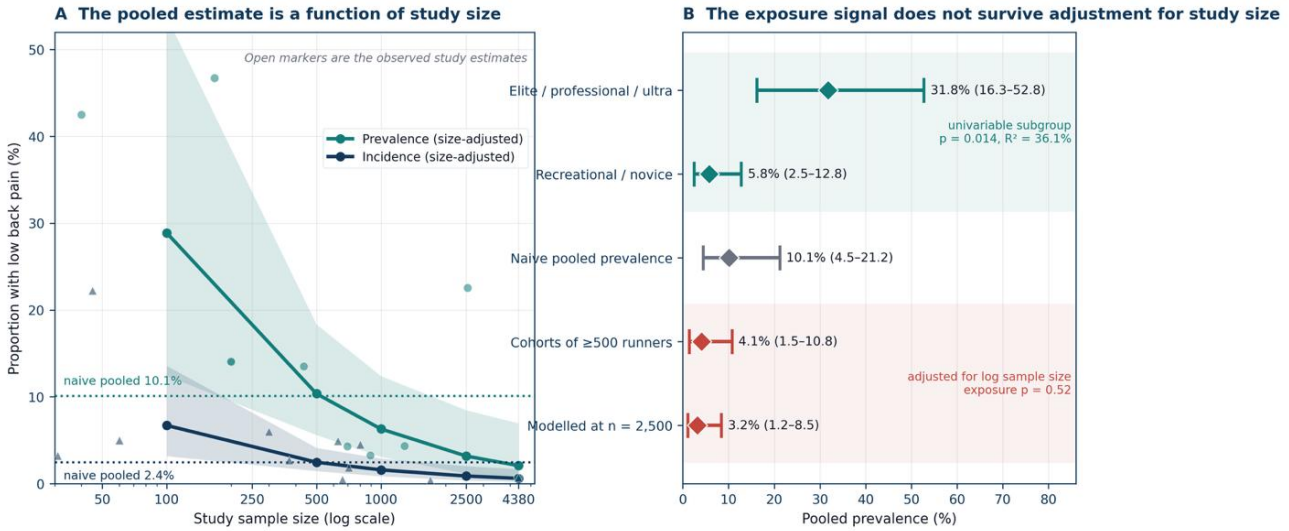

Figure 5. Study size and the reported proportion. Panel A shows the predicted proportion with its 95% confidence band against sample size on a logarithmic scale, with the observed study estimates plotted as open markers and the naive pooled values as dotted lines. Panel B shows what happens to the apparent effect of competitive standard once study size is adjusted for.

3.6 Subgroup analyses

Seven moderators were examined in each primary pool, giving fourteen exploratory tests in total; the nominal p values below are unadjusted, and under a Bonferroni correction for fourteen comparisons the threshold would be approximately 0.0036. In the prevalence pool, elite, professional and ultra-distance cohorts reported a pooled prevalence of 31.78% (95% CI 16.26 to 52.78) against 5.81% (95% CI 2.52 to 12.82) in recreational and novice cohorts, a nominal between-group difference at p = 0.014 explaining 36.1% of between-study variance. Publication era was also nominally significant at p = 0.013, but in this evidence base the two pre-2000 studies are the largest and third largest in the pool, so era is a proxy for study size rather than a temporal trend. Neither the recall window, the anatomical label, the case ascertainment method, the risk-of-bias judgement nor the study design produced a significant difference. In the incidence pool the observation window behaved as it should, with studies following runners for 26 weeks or longer recording 7.61% against 1.46% for shorter windows (p = 0.013); this is a positive control on the analysis rather than a discovery. Reading the ascertainment rows of Table 6, symptom-based ascertainment gave a higher prevalence than injury-based ascertainment, 16.82% against 8.17%, which is the expected direction, but the difference was not significant. Full results are given in Table 6.

3.7 Meta-regression

Univariable meta-regression identified the natural logarithm of sample size as the strongest single moderator in both primary pools. In the prevalence pool each unit increase in log sample size was associated with a change of -0.7807 on the logit scale (SE 0.2547; p = 0.00218), corresponding to an odds ratio of 0.458 per unit and explaining 48.6% of between-study variance. In the incidence pool the coefficient was -0.6490 (SE 0.1956; p = 0.000905), an odds ratio of 0.523 per unit, explaining 57.6%. Multivariable models were then fitted to test whether any clinical moderator survived adjustment for study size. None did. When competitive standard was added to log sample size in the prevalence pool, standard became non-significant (p = 0.515). When the observation window was added to log sample size in the incidence pool, the window became non-significant (p = 0.708). Adding ascertainment method to log sample size raised the explained variance to 55.0% while ascertainment itself remained non-significant (p = 0.151) and log sample size strengthened (p = 0.000735). In both pools the model containing log sample size alone was the best fitting on the corrected Akaike information criterion, at 35.67 and 35.34 against 49.22 and 44.12 for the two-covariate alternatives. Univariable and multivariable results are given together in Table 7.

The collapse of that signal is shown side by side with its univariable form in Figure 5 panel B. The three studies in the elite and ultra stratum are also three of the four smallest in the prevalence pool, at 40, 167 and 199 participants, so exposure tier and study size are almost perfectly confounded in this evidence base and cannot be separated by any model fitted to these ten studies.

3.8 Sex, severity and exposure time

No included study reported its low back count separately by sex, so no sex-stratified pooled estimate could be derived and sex was not entered as a moderator. Three cohorts are informative at study level. The largest study in the prevalence pool is male-only, with 4,358 participants and a 12-month prevalence of 0.69%; the female-only novice cohort in the incidence pool, with 373 participants, recorded a 12-week incidence of 2.68%; and the strongly female-predominant novice cohort of 1,696 runners recorded a 6-week incidence of 0.35%. These three cannot be compared directly, because they differ in design, window and case definition as well as in sex, and the male-only study is also the second largest in the review. Sex differences in pelvic and hip kinematics during running are established from maturation onward, so a sex effect on lumbopelvic loading is biologically plausible and simply has not been measured against low back pain in this literature.[33] The absence of a sex-stratified estimate is a substantive gap in this literature rather than an omission of this review, and it is the more conspicuous because low back pain is more common in women than in men in almost every general population in which it has been measured.

e-ISSN 2747-2051
DOI: 10.37275/amcr.v7i4.936

Table 6. Pre-specified subgroup analyses for the prevalence and incidence pools.

Pool	Moderator	Level	k	Pooled % (95% CI)	I^2 (%)	p between	R^2 (%)
Prevalence	Recall window / observation period	12-month period	6	10.13 (3.21-27.73)	99.3	0.976	0.0
Prevalence	Recall window / observation period	Point prevalence	1	13.53 (10.46-17.11)	NA	0.976	0.0
Prevalence	Recall window / observation period	Lifetime	3	9.09 (2.05-32.30)	97.4	0.976	0.0
Prevalence	Anatomical label (lumbar-only vs composite)	Composite	4	5.03 (1.47-15.81)	98.1	0.172	8.9
Prevalence	Anatomical label (lumbar-only vs composite)	Lumbar	6	15.65 (6.36-33.64)	98.8	0.172	8.9
Prevalence	Case ascertainment	Injury-based (qualifying RRI)	7	8.17 (2.60-22.86)	98.7	0.458	0.0
Prevalence	Case ascertainment	Symptom-based (any LBP)	3	16.82 (12.40-22.41)	83.9	0.458	0.0
Prevalence	Competitive standard / exposure tier	Recreational/novice	7	5.81 (2.52-12.82)	98.7	0.0143	36.1
Prevalence	Competitive standard / exposure tier	Elite/professional/ultra	3	31.78 (16.26-52.78)	92.2	0.0143	36.1
Prevalence	Publication era	Pre-2000	2	1.74 (0.48-6.14)	97.1	0.013	37.1
Prevalence	Publication era	2000 onward	8	15.17 (7.76-27.55)	98.2	0.013	37.1
Prevalence	Risk of bias	Some concerns	6	5.92 (1.87-17.18)	98.8	0.281	5.4
Prevalence	Risk of bias	High risk	2	24.76 (10.29-48.56)	87.5	0.281	5.4
Prevalence	Risk of bias	Low risk	2	18.88 (13.45-25.85)	73.0	0.281	5.4
Prevalence	Study design	Cross-sectional	7	12.24 (4.10-31.27)	99.2	0.866	0.0
Prevalence	Study design	Prospective cohort	1	4.35 (3.30-5.61)	NA	0.866	0.0
Prevalence	Study design	Retrospective descriptive	1	14.00 (9.51-19.59)	NA	0.866	0.0
Prevalence	Study design	Retrospective cohort	1	4.30 (2.92-6.09)	NA	0.866	0.0
Incidence	Recall window / observation period	Window 26 weeks or longer	4	7.61 (3.56-15.55)	69.4	0.013	37.7
Incidence	Recall window / observation period	Window under 26 weeks	7	1.46 (0.69-3.08)	94.1	0.013	37.7
Incidence	Anatomical label (lumbar-only vs composite)	Lumbar	7	2.53 (0.95-6.58)	94.6	0.87	0.0
Incidence	Anatomical label (lumbar-only vs composite)	Composite	4	2.33 (0.78-6.73)	94.3	0.87	0.0
Incidence	Case ascertainment	Injury-based (qualifying RRI)	10	2.28 (1.02-5.04)	94.5	0.651	0.0
Incidence	Case ascertainment	Symptom-based (any LBP)	1	4.50 (3.17-6.18)	NA	0.651	0.0
Incidence	Competitive standard / exposure tier	Recreational/novice	9	2.30 (0.99-5.30)	96.3	0.628	0.0
Incidence	Competitive standard / exposure tier	Elite/professional/ultra	2	4.40 (1.66-11.13)	0.0	0.628	0.0
Incidence	Publication era	Pre-2000	2	11.35 (3.73-29.69)	68.4	0.0312	29.7
Incidence	Publication era	2000 onward	9	1.79 (0.90-3.54)	93.6	0.0312	29.7
Incidence	Risk of bias	High risk	2	11.35 (3.73-29.69)	68.4	0.125	20.2
Incidence	Risk of bias	Low risk	6	1.82 (0.81-4.04)	94.2	0.125	20.2
Incidence	Risk of bias	Some concerns	3	1.72 (0.46-6.29)	78.5	0.125	20.2
Incidence	Study design	Prospective cohort	8	3.31 (1.45-7.38)	92.7	0.306	2.8
Incidence	Study design	Retrospective cohort	1	0.45 (0.09-1.32)	NA	0.306	2.8
Incidence	Study design	Cross-sectional	2	1.78 (0.49-6.29)	96.6	0.306	2.8

Where a moderator produced a level containing fewer than two studies, the level-specific estimate is reported but the omnibus between-group test is not interpretable and should be disregarded; such moderators are the recall window and the study design in the prevalence pool, and the case ascertainment method and the study design in the incidence pool. All fourteen tests are exploratory and unadjusted for multiplicity.

Table 7. Univariable and multivariable meta-regression, with model comparison. Intercept terms are omitted for readability.

Model	Pool	Covariate	Beta (logit)	SE	Odds ratio per unit	p	R^2 (%)	AICc
Univariable	Prevalence	Publication year	0.0846	0.0304	1.088	0.00542	43.6	-
Univariable	Prevalence	Log sample size	-0.7807	0.2547	0.458	0.00218	48.6	-
Univariable	Incidence	Publication year	-0.0575	0.0288	0.944	0.0459	25.9	-
Univariable	Incidence	Log sample size	-0.6490	0.1956	0.523	0.000905	57.6	-
Multivariable: logn + exposure	Prevalence	Log sample size	-0.5924	0.3918	0.553	0.131	44.8	49.22
Multivariable: logn + exposure	Prevalence	Recreational or novice standard	-0.7442	1.1432	0.475	0.515	44.8	49.22
Multivariable: logn + year	Prevalence	Log sample size	-0.5458	0.258	0.579	0.0344	60.8	46.93
Multivariable: logn + year	Prevalence	Publication year	0.0537	0.0293	1.055	0.0672	60.8	46.93
Multivariable: logn + ascertainment	Prevalence	Log sample size	-0.8107	0.2401	0.445	0.000735	55.0	47.84
Multivariable: logn + ascertainment	Prevalence	Symptom-based ascertainment	0.997	0.6936	2.710	0.150578	55.0	47.84
Multivariable: logn + window	Incidence	Log sample size	-0.5337	0.3584	0.586	0.136	50.0	44.12
Multivariable: logn + window	Incidence	Observation window under 26 weeks	-0.4016	1.073	0.669	0.708	50.0	44.12
Multivariable: logn + year	Incidence	Log sample size	-0.5808	0.2891	0.559	0.0445	50.6	43.86
Multivariable: logn + year	Incidence	Publication year	-0.0105	0.0339	0.990	0.7563	50.6	43.86
Single-covariate fit: logn only	Prevalence	Log sample size	-0.7807	0.2547	0.458	0.00218	48.6	35.67
Single-covariate fit: exposure only	Prevalence	Recreational or novice standard	-2.0236	0.826	0.132	0.0143	36.1	37.47
Single-covariate fit: year only	Prevalence	Publication year	0.0846	0.0304	1.088	0.00542	43.6	36.59
Single-covariate fit: logn only	Incidence	Log sample size	-0.649	0.1956	0.523	0.000905	57.6	35.34
Single-covariate fit: window only	Incidence	Observation window under 26 weeks	-1.6969	0.6833	0.183	0.013	37.7	37.36
Single-covariate fit: year only	Incidence	Publication year	-0.0575	0.0288	0.944	0.0459	25.9	38.81

The odds ratio per unit is the exponentiated coefficient and is a restatement of it, not a separate estimate. AICc is the corrected Akaike information criterion; lower values indicate better fit and are reported only for the models fitted for comparison. The single-covariate model containing log sample size was the best fitting in both pools.

Severity was not synthesisable either. No included study reported pain intensity, disability or care-seeking in a form that could be pooled, and none reported recurrence. The closest available proxy is the time-loss threshold embedded in each study's case definition, which functions as a crude severity scale: the qualifying threshold ranged from any complaint reported on a standardised health-problem questionnaire, through pain restricting running for at least one day, to an injury causing an absence of two weeks or more. Twenty-one of the twenty-five studies used a time-loss or clinical injury definition of some kind, which means the pooled proportions are dominated by episodes that were at least disruptive enough to alter training, and are not estimates of trivial or self-limiting back discomfort. Four studies counted any self-reported episode and are the only ones that capture the milder end of the spectrum.

Exposure time is the third unsynthesisable dimension and it bears directly on the incidence estimate. Observation windows ranged from a single race to two years, a difference of roughly two orders of magnitude in exposure, and only one included study reported injury rates per unit of running exposure in a form that could be converted. A raw incidence proportion pooled across such a range is not a rate, and the window-stratified estimates of 7.61% for windows of 26 weeks or longer and 1.46% for shorter windows are for that reason at least as informative as the single pooled value of 2.44%, which should be read as a summary of a heterogeneous set rather than as an estimate of a well-defined quantity. Conversion to person-time was not possible with the data reported by the included studies.

3.9 Sensitivity analyses

Six pre-specified sensitivity analyses were run in each primary pool. Restricting the prevalence pool to lumbar-specific anatomical labels raised the estimate to 15.65%, and restricting to studies published from 2000 onward raised it to 15.17%; both restrictions remove the very large early event surveys and therefore move the estimate in the direction predicted by the small-study gradient. Excluding studies at high risk of bias lowered it to 7.91%, and excluding samples of fewer than 100 participants lowered it to 8.43%. In the incidence pool every restriction moved the estimate downward, from 2.44% to between 1.75% and 1.79%, except the lumbar-only restriction which left it essentially unchanged at 2.53%. Results are given in Table 8.

e-ISSN 2747-2051
DOI: 10.37275/amcr.v7i4.936

Table 8. Pre-specified sensitivity analyses.

Pool	Analysis	k	Total n	Cases	Pooled % (95% CI)	I^2 (%)	τ^2
Prevalence	a. Primary analysis (all studies)	10	10817	928	10.11 (4.48-21.24)	99.0	1.9580
Prevalence	b. Lumbar-only anatomical labels	6	4772	786	15.65 (6.36-33.64)	98.8	1.5410
Prevalence	c. Excluding derived / re-extracted numerators	9	10120	898	11.09 (4.61-24.37)	99.1	2.0740
Prevalence	d. Published 2000 onward	8	5171	842	15.17 (7.76-27.55)	98.2	1.1475
Prevalence	e. Excluding high risk of bias	8	10577	883	7.91 (3.14-18.53)	99.1	1.9496
Prevalence	f. Excluding samples n<100	9	10777	911	8.43 (3.69-18.10)	99.0	1.7692
Incidence	a. Primary analysis (all studies)	11	9682	162	2.44 (1.17-5.04)	94.9	1.4233
Incidence	b. Lumbar-only anatomical labels	7	6684	97	2.53 (0.95-6.58)	94.6	1.5664
Incidence	c. Excluding derived / re-extracted numerators	10	9651	161	2.44 (1.12-5.25)	95.6	1.5179
Incidence	d. Published 2000 onward	9	9577	149	1.79 (0.90-3.54)	93.6	0.9857
Incidence	e. Excluding high risk of bias	9	9577	149	1.79 (0.90-3.54)	93.6	0.9857
Incidence	f. Excluding samples n<100	8	9546	148	1.75 (0.84-3.60)	94.6	1.0474

Reading across the restriction rows of Table 8, the leave-one-out analysis showed that no single study drives either primary result. Prevalence ranged from 8.20% to 13.28% across the ten iterations and incidence from 1.93% to 3.00% across the eleven, and every iteration lay within the confidence interval of the corresponding primary analysis. Removing the single-event adolescent elite study changed the pooled incidence not at all, at 2.44% before and after, which is reassuring given that it contributes one case from 31 participants. Full leave-one-out results are given in Table 9. The pooled estimates are therefore robust to the removal of any one study, but not to the small-study gradient documented above, which is a property of the evidence base as a whole rather than of any single report. No leave-one-out analysis was computed for the secondary pool, and this review has not recomputed one; the study-level values in the characteristics table allow the reader to gauge the influence of the two clinic series and of the study carried there as an approximation.

Table 9. Leave-one-out analysis for both primary pools.

Pool	Study omitted	k	Pooled % (95% CI)	I^2 (%)
Prevalence	Marti B 1988	9	13.28 (6.89-24.08)	98.4
Prevalence	Walter SD 1989	9	11.08 (4.60-24.37)	99.0
Prevalence	Woolf SK 2002	9	9.79 (3.95-22.26)	99.0
Prevalence	Chang WL 2012	9	11.41 (4.83-24.64)	99.0
Prevalence	Ellapen TJ 2013	9	9.76 (3.94-22.18)	99.1
Prevalence	Malliaropoulos N 2015	9	8.43 (3.69-18.10)	99.0
Prevalence	Teixeira RN 2016	9	9.75 (3.94-22.16)	99.1
Prevalence	Vitez L 2017	9	11.09 (4.61-24.37)	99.1
Prevalence	Koech RC 2021	9	8.20 (3.68-17.25)	98.8
Prevalence	Maselli F 2021	9	9.17 (3.76-20.68)	98.5
Incidence	Bach DK 1985	10	1.93 (1.01-3.67)	92.8
Incidence	Lysholm J 1987	10	2.32 (1.05-5.05)	95.6
Incidence	Buist I 2010	10	2.26 (1.01-4.98)	94.6
Incidence	Rasmussen CH 2013	10	2.88 (1.39-5.87)	94.7
Incidence	Kluitenberg B 2015	10	3.00 (1.51-5.89)	93.9
Incidence	van der Worp MP 2016	10	2.43 (1.07-5.39)	95.5
Incidence	von Rosen P 2017	10	2.44 (1.12-5.25)	95.6
Incidence	Messier SP 2018	10	2.22 (1.00-4.84)	94.9
Incidence	Besomi M 2019	10	2.81 (1.31-5.92)	94.0
Incidence	Dallinga J 2019	10	2.52 (1.12-5.59)	95.4
Incidence	Wu B 2021	10	2.28 (1.02-5.04)	94.5

3.10 Certainty of the evidence

Certainty was rated separately for each pool. Both primary outcomes were downgraded once for risk of bias, twice for inconsistency and once each for indirectness and publication bias, and were rated very low. The inconsistency and indirectness downgrades are not fully independent, because variation in case definition contributes to both; the ratings are reported as pre-specified, but a reader who judged that variation to have been counted once rather than twice would place both primary outcomes one level higher, at low certainty. The secondary outcome was downgraded once each for risk of bias, inconsistency and imprecision, and was rated low; publication bias could not be assessed because fewer than ten studies contribute. The assessment is summarised in Table 10.

Table 10. Certainty of the evidence, rated with GRADE adapted for prevalence data.

Outcome	Starting certainty	Risk of bias	Inconsistency	Indirectness	Imprecision	Publication bias	Certainty
Prevalence among runners (k = 10)	High	Serious (-1): 2 low risk, 6 some concerns, 2 high risk	Very serious (-2): I^2 99.0%, prediction interval 0.4 to 76.9%	Serious (-1): case definitions range from any self-reported pain to injuries causing more than two weeks off	Not serious	Serious (-1): Peters $p < 0.0001$	Very low
Incidence among runners (k = 11)	High	Serious (-1): 6 low risk, 3 some concerns, 2 high risk	Very serious (-2): I^2 94.9%, prediction interval 0.1 to 29.9%	Serious (-1): observation windows range from a single race to two years	Not serious	Serious (-1): Peters $p < 0.0001$	Very low
Low back share of injuries (k = 7)	High	Serious (-1): two clinic series at high risk	Serious (-1): I^2 79.7%	Not serious	Serious (-1): fewer than ten studies, wide interval	Not assessable	Low

4. Discussion

This review provides the first pooled estimates of low back pain in runners, and its most transportable result is not the one the question was originally framed around. Across seven studies and 3,503 running-related injuries, the low back accounted for 4.88% of injuries, with markedly lower heterogeneity than either runner-denominator pool and a prediction interval of 1.40% to 15.63%. That consistency should not be mistaken for generalisability: more than half of that denominator comes from a single 1981 clinic series rated at high risk of bias, and no influence analysis was computed for this pool. About one running injury in twenty is a low back injury, and that proportion is remarkably consistent from a Canadian clinic series of the early 1980s to a South African trail-running surveillance cohort four decades later.[3,8] Pooled prevalence among runners was 10.11% and pooled incidence 2.44%, but both are dominated by study size, and in cohorts of at least 500 runners the corresponding values were 4.07% and 1.32%. Both figures are lower than the prevalence of low back pain in the general adult population, which is the most immediately useful thing this review has to say to a clinician, although the comparison cannot be made formally because no included study carried a non-running comparator.[1]

4.1 Why the published literature looks irreconcilable

The most consequential finding of this review is methodological rather than clinical. The apparent contradiction in the published literature, where estimates run from 0.3% to above 40%, is not principally a clinical phenomenon. It is driven by three things, in descending order of importance: the denominator, the size of the study and the case definition.

The denominator problem is structural and, once seen, is difficult to unsee. Four of the twenty-five studies express the low back count against the number of injuries rather than the number of runners, and two of those four were carried into the earlier synthesis of this literature as though they reported runner-denominator prevalence.[2,3,31] A study reporting that 3.7% of running injuries are low back injuries is not reporting that 3.7% of runners have low back pain, and the two quantities differ by roughly the injury rate of the cohort. Any future synthesis in this field should assign studies to a pool by the denominator the source actually uses, and should report the injury-site share as a distinct outcome rather than folding it into a prevalence estimate.

The size effect is larger than any clinical moderator examined. Log sample size explained 48.6% of between-study variance in the prevalence pool and 57.6% in the incidence pool, and Peters regression was significant at $p < 0.0001$ in both. Modelled at 2,500 runners, prevalence falls from 10.11% to 3.18% and incidence from 2.44% to 0.88%; restricting to cohorts of at least 500 runners gives 4.07% and 1.32%. These procedures are diagnostics rather than corrections, and the true population value is not established by them; what they do establish is that the naive pooled estimates sit toward the upper end of the plausible range and that the lower part of the pooled confidence interval is the more defensible region for clinical expectation. Two mechanisms plausibly contribute. Small runner cohorts are usually recruited in a clinic, a specialist race or a national squad, and are therefore enriched for symptomatic athletes; and small studies reporting an unremarkable low back finding are less likely to have that finding highlighted in the abstract, and so less likely to be found by a search.

Case definition matters, but less than might be expected. Symptom-based ascertainment produced a higher pooled prevalence than injury-based ascertainment, 16.82% against 8.17%, which is the expected direction, since a definition requiring time loss will always count fewer cases than one asking whether the runner has had any back pain. The difference did not reach significance on the omnibus test, and ascertainment remained non-significant when added to log sample size. That should not be read as evidence that definition is unimportant. With ten studies the subgroup analysis is underpowered, definition is confounded with study size in this literature because the largest surveys tend to use the stricter injury definitions, and all fourteen moderator tests reported here are exploratory and unadjusted for multiplicity.

4.2 What cannot be concluded

The univariable finding that elite, professional and ultra-distance cohorts report roughly five times the prevalence of

recreational cohorts is arresting, and it is the kind of result that would ordinarily be reported as a headline. It should not be. Once log sample size is entered into the model, competitive standard is no longer significant (p = 0.515), and the model containing study size alone fits better on the corrected Akaike information criterion. The three studies in the elite and ultra stratum are also three of the four smallest in the prevalence pool, with 40, 167 and 199 participants, so exposure tier and study size are almost perfectly confounded in this evidence base. Whether elite and ultra-distance runners genuinely carry a higher burden of low back pain remains an open question that this literature cannot answer, and answering it will require a large elite cohort rather than another small one.

That said, the highest single estimate in the review deserves more than a statistical dismissal. Forty-seven per cent of 167 professional endurance runners in the Rift Valley of Kenya reported low back pain in the preceding year, in the only professional African cohort in this literature and in a training environment of hills, valleys and unpaved surfaces at altitude that is unrepresented elsewhere in the evidence base.[12] That environment is not incidental: downhill and uphill running alter trunk posture and paraspinal muscle activity substantially relative to level running, so a terrain-dominated training load is a mechanistically plausible reason for a higher lumbar burden.[34] The correct reading is not that the figure is unreliable because the study is small, but that it is the most important unreplicated observation in the field and that the training environments producing the world's fastest distance runners have essentially never been studied for spinal morbidity. It is also, notably, an injury-based figure: the cohort's case definition required more than 48 hours out of running, so 47% is not the loose symptom-based measure that a figure of that size might suggest, and it is therefore more striking rather than less when set beside the symptom-based estimates elsewhere in the pool. The same applies, in a smaller way, to the ultra-trail cohort in which seventeen of forty runners reported low back pain over a lifetime.[11] Both are priorities for replication in adequately powered cohorts.

Nor can this review support any claim that running is protective against low back pain. That claim has been made on the basis that reported prevalence in runners appears low relative to population figures, but no included study recruited a non-running comparator, and the comparison across data sources is confounded by age, by sex distribution, by the healthy-exerciser effect and by case definition. What can be said is narrower and more useful: among people who already run, low back pain is an uncommon reason for interrupting training, and the low back is not among the anatomical regions where running injuries concentrate.[8,35,36]

Finally, no pooled risk-factor estimate could be produced. The factors reported across the pool are heterogeneously defined and sparsely replicated: hip flexion restriction measured by the Thomas test, hamstring and lumbar flexibility, leg-length discrepancy, body mass index, warm-up adequacy, fatigue and running gait posture each appear in only one or two studies.[3,10,14,20,25] Two of the largest and most methodologically careful surveys in the pool found no association at all between low back pain and training volume, equipment or lifestyle.[9,30] The determinants that are reproducible in the general adult population, such as trunk muscle strength and postural control, have not been measured in a runner cohort against this outcome at all.[37,38] The risk-factor literature for this outcome is at the stage of hypothesis generation, not of synthesis.

4.3 What the pooled proportions do and do not describe

A pooled proportion of runners with low back pain says nothing about how severe that pain was, and this is a real limitation for a clinical reader. None of the included studies reported pain intensity, disability or care-seeking in a poolable form, and none reported recurrence. The one thing that can be said is that twenty-one of the twenty-five studies required an episode to meet a time-loss or clinical injury threshold, so the pooled proportions are dominated by episodes disruptive enough to alter training rather than by trivial stiffness. Readers should therefore interpret the figures as estimates of clinically consequential low back pain in runners, not of any back discomfort whatever.

The pooled proportions also say nothing about what kind of low back pain is being counted. Almost every included study recorded a single anatomical category without specifying pathology, and the granularity that does exist is instructive: the oldest clinic survey in the pool separated its 68 low back cases into non-specific low back pain, sciatica, spondylolysis and spondylolisthesis, and that distinction is lost the moment the categories are summed.[3] The pooled estimates in this review are therefore dominated by non-specific low back pain, and they should not be read as reassurance in a runner whose pain does not behave mechanically. Even within non-specific chronic low back pain, imaging and symptom patterns identify subgroups that a single anatomical label cannot separate.[39] Sacral and pelvic bone stress injury presents as low back or buttock pain in exactly this population and has a distinct risk profile in female runners; early-stage spondylolysis in a younger runner produces low back pain that behaves differently from the mechanical pattern; radicular pain changes management entirely; and none of the three can be identified from an anatomical injury tally.[40,41] The low prevalence of low back pain in runners does not lower the threshold for investigating a runner whose pain fails to settle with load modification, and this review offers no evidence bearing on when to image or when to refer, which were outside its scope. That caution belongs in the same breath as the low figure rather than elsewhere in the paper: a base rate of roughly four in a hundred describes how often the complaint arises, not how carefully any individual presentation should be assessed.

4.4 Comparison with adjacent literatures

The picture that emerges here is consistent with the broader running-injury literature, in which the knee, the shin and the foot dominate the anatomical distribution and the lumbar spine sits well down the list.[35,42-44] Large prospective cohorts continue to find that the modifiable drivers of running injury are training-session characteristics, previous injury, body mass index, sleep quality, comorbidity and psychosocial state rather than anatomical site-specific vulnerability, and none of them identifies the lumbar spine as a region of concentrated risk.[36,45-54] It is also consistent with the more recent evidence that running is compatible with, and may be therapeutic for, chronic low back pain: a randomised trial of a progressive running programme in adults with non-specific chronic low back pain reported it to be acceptable and efficacious, with secondary analyses showing improvement in mental health symptoms and pain catastrophising, and qualitative work documenting that the fears runners and patients hold about running with a painful back are not borne out.[55-57] Imaging studies point the same way: a narrative synthesis of thirteen studies found that any negative change in intervertebral disc morphology after running is short-lived, while cross-sectional comparisons suggest a mild positive

long-term effect, and recent magnetic resonance work relates disc health to lifetime physical loading rather than to running per se.[58,59]

Comparison with other sports is instructive. Pooled estimates in professional soccer, in university athletes, in adolescent field hockey players, in college rifle athletes and in elite cross-country skiers are all higher than the large-cohort figures reported here, and in several of those sports the lumbar spine is a leading rather than a marginal site of injury.[15,16,60-62] Sports that load the lumbar spine in repeated flexion, extension or rotation, or that hold it in a sustained non-neutral posture, appear to carry a different lumbar risk profile from running, in which the spine is loaded axially and rhythmically at a cadence that itself modifies spinal load.[63-66] Triathletes, whose training includes running alongside two other disciplines, again report higher musculoskeletal injury prevalence overall.[67] A quantitative comparison across sports was outside the scope of this review, but it is the obvious next synthesis, and this review supplies the running reference value it would need.

4.5 Clinical implications

Three practical points follow, and they are easiest to hold in natural frequencies. Of a hundred recreational runners seen over a year, roughly four would report low back pain on the large-cohort estimate, and roughly ten on the naive pooled estimate; the true figure is likely to sit in the lower part of that range. Of every twenty running injuries presenting to a sports clinic, about one will be at the low back, against roughly six at the knee and four at the shin or lower leg in the same cohorts. Of a hundred novice runners followed through a structured beginner programme of six to twelve weeks, fewer than three would be expected to develop new low back pain.

First, therefore, when a runner presents with low back pain, running is unlikely to be the explanation by default. The clinician should look for the ordinary causes of low back pain in an active adult before attributing the pain to running mileage. This matches what runners themselves report: in the largest survey in this pool, 77% of those who had experienced low back pain believed it was not caused by running.[9] Second, blanket advice to stop running is not supported by the evidence assembled here or by the trial evidence in chronic low back pain, and large cohort data show that weight-bearing activity volume is inversely rather than positively associated with chronic low back pain and that common physical activities carry small transient rather than sustained risk in people who already have it.[55,57,68,69] Third, the wide prediction intervals mean that any single published figure should be quoted with its uncertainty. A clinician who reads that one runner in ten has low back pain, and a clinician who reads that fewer than one in a hundred does, are both reading real studies from this literature.

The heterogeneity finding also has a direct implication for how running-injury surveillance should be reported. Studies using standardised health-problem questionnaires with defined anatomical regions and explicit denominators produced estimates that reconcile with each other; studies using bespoke instruments and composite anatomical labels did not.[6-8,28] Wider adoption of a standard anatomical taxonomy, explicit reporting of both denominators, and reporting of exposure in person-time would substantially improve the synthesisability of this literature at no additional cost to investigators.

4.6 Research agenda

Four priorities follow directly from what this review could not do. The first is a diagnostic accuracy study of the lumbar clinical examination in runners. No included study reported any lumbar examination test against a reference standard in a runner cohort, which means that the physical examination performed daily on runners with low back pain has never been evaluated in runners; a recent meta-analysis of lumbar multifidus morphology in athletes with chronic non-specific low back pain shows that examination-based constructs can be synthesised in athletic populations when the primary studies exist, so the gap is one of primary data rather than of method.[70] The second is a large cohort of elite and ultra-distance runners, adequately powered and prospectively designed, to separate competitive standard from study size. The third is sex-stratified reporting: not one of the twenty-five studies reported its low back count by sex, and until that changes no sex-specific estimate is possible in a population where sex is among the strongest known determinants of low back pain generally. The fourth is severity and exposure reporting, so that a future synthesis can express incidence as a rate per unit of running exposure and can distinguish an episode of two days from an episode of two months.

4.7 Strengths and limitations

The principal strength of this review is that every extracted value was recomputed from its own numerator and denominator, and every value that failed to reconcile was returned to the source document. That process changed the structure of the analysis: it identified the denominator taxonomy, corrected a mislabelled study, reclassified one study from prevalence to incidence and re-extracted another whose denominator was wrong by a factor of six. A second strength is that bibliographic records were rebuilt field by field from live database records rather than transcribed. A third is that the analysis used a model appropriate to the data, since the double-arcsine transformation would have been unstable across a 140-fold range of denominators. A fourth is that the search was subsequently re-executed with broadened concept blocks and across all sources indexed by Europe PMC, which returned no additional eligible study and provides some assurance that the small size of this evidence base reflects the literature rather than the search.

Several limitations qualify the findings. Certainty was very low for both primary outcomes, principally because of inconsistency and small-study effects. Heterogeneity remained extreme after every subgroup and sensitivity analysis, and the prediction intervals are so wide that the pooled point estimates should never be quoted alone. The evidence base is small and its denominators are concentrated. Ten and eleven studies contribute to the primary pools, but the 1988 Bern event survey supplies 40.3% of the prevalence denominator, is male-only and reports the lowest estimate in that pool, so the pooled prevalence is anchored to a substantial degree by a single-sex cohort in a review that can otherwise say nothing about sex; the 2019 Chilean race survey supplies 45.2% of the incidence denominator and is one of only two cross-sectional designs among eleven contributors; and the 1981 clinic series supplies 51.9% of the secondary denominator while being rated at high risk of bias on both sampling domains. The last of these bears directly on how far the injury-site share can be carried, and is the reason that estimate is described below as the most consistent rather than the most generalisable. Six studies used composite anatomical labels that also capture the buttock, sacrum, pelvis or hip, which inflates the count relative to a lumbar-specific definition, although the lumbar-only sensitivity analysis moved the estimate upward rather than downward because it also

removed the large early surveys. One study is carried in the secondary pool as an acknowledged approximation because its runner-denominator numerator could not be recovered from an image-only source table. Screening and extraction were performed independently and reconciled, but no formal inter-rater agreement statistic was computed. Scopus, SPORTDiscus and CINAHL could not be searched because the necessary institutional licences were not available to the review team, and reference harvesting from the prior synthesis therefore supplies part of this review's reach; grey literature was not searched beyond the preprint and thesis records returned by Europe PMC, and no attempt was made to contact authors for unpublished counts. The review was not prospectively registered. A further limitation lies inside the secondary pool itself. Four of its seven studies divide the low back count by injuries and two divide it by injured runners, and those denominators differ by the mean number of injuries each injured runner sustains; 2,508 of the 3,503 units in that pool, 71.6%, are injury events, and 995, 28.4%, are injured runners. This is a milder form of the same denominator problem the review was built to expose, it was not separable without splitting a seven-study pool into two of four and two, and it means the injury-site share should be read as the share of recorded injury events rather than as a per-injured-runner risk. Coverage is also uneven: sixteen countries are represented but none in Oceania, and only two in East Asia, despite substantial running-injury surveillance literatures in both regions. Finally, three studies contribute to both a primary and the secondary pool under different denominators; they are counted once in the review total and are flagged in the characteristics table, but the three pool totals are not additive and comparisons between pools are made across partly overlapping evidence.

5. Conclusion

Low back pain is an uncommon site of running-related morbidity. The most consistent quantity in this literature is the low back share of running-related injuries, 4.88%, or about one running injury in twenty. Pooled across ten studies and 10,817 runners, prevalence was 10.11%, and pooled across eleven studies and 9,682 runners, the incidence of new low back pain was 2.44%; but both figures carry prediction intervals too wide to be quoted alone, and both are dominated by a strong and highly significant gradient with study size, the corresponding values in cohorts of at least 500 runners being 4.07% and 1.32%. The apparent contradiction between published estimates is largely explained by which denominator a study used and how large it was, rather than by any characteristic of the runners themselves, and the univariable finding that elite and ultra-distance cohorts report more low back pain does not survive adjustment for study size. Clinically, a runner presenting with low back pain should not have that pain attributed to running by default, blanket advice to stop running is not supported, and the low prevalence does not lower the threshold for investigating pain that fails to settle. Methodologically, future running-injury surveillance should report both denominators explicitly, use a standard anatomical taxonomy, report exposure in person-time and stratify by sex; and a definitive estimate for elite and ultra-distance runners will require a large purpose-designed cohort rather than another small one.

6. Declarations

Statement on the Use of Artificial Intelligence

None declared.

Author Contribution Statement

BGKA: Conceptualization, Methodology, Investigation, Data curation, Formal analysis, Visualization, Writing - original draft, Writing - review and editing. INDMWJ: Conceptualization, Methodology, Investigation, Data curation, Validation, Writing - review and editing. IWASN: Conceptualization, Methodology, Validation, Supervision, Writing - review and editing. All authors approved the final manuscript and accept accountability for the work.

Funding Statement

This research received no specific grant from any funding agency in the public, commercial or not-for-profit sectors.

Conflict of Interest

The authors declare that they have no competing interests.

Data Availability Statement

All data analysed in this review are contained within the published reports of the 25 included studies. The extracted study-level dataset, the search strategies with their yields, and the analysis outputs are available from the corresponding author on reasonable request.

Ethical approval and consent to participate

Not applicable. This study analysed aggregate data from previously published research and did not involve human participants, human material or human data collected by the authors.

Consent for publication

Not applicable.

Acknowledgements

The authors thank the investigators of the included studies, several of whom made full anatomical breakdowns available in open-access supplementary tables without which the denominator adjudication would not have been possible, and the two reviewers whose reports prompted the extended search, the denominator flagging in the characteristics table and the addition of the definition, severity and exposure-time analyses.

Author Approval

All authors have reviewed and approved the final version of the manuscript. The authors agree to be accountable for the work.

7. References

1. GBD 2021 Low Back Pain Collaborators. Global, regional, and national burden of low back pain, 1990-2020, its attributable risk factors, and projections to 2050: a systematic analysis of the Global Burden of Disease Study 2021. Lancet Rheumatol. 2023;5(6):e316-e329. doi: 10.1016/S2665-9913(23)00098-X.
2. Maselli F, Storari L, Barbari V, et al. Prevalence and incidence of low back pain among runners: a systematic review. BMC Musculoskelet Disord. 2020;21(1):343. doi: 10.1186/s12891-020-03357-4.
3. Clement DB, Taunton JE, Smart GW, et al. A survey of overuse running injuries. Phys Sportsmed. 1981;9(5):47-58. doi: 10.1080/00913847.1981.11711077.
4. Marti B, Vader JP, Minder CE, et al. On the epidemiology of running injuries. The 1984 Bern Grand-Prix study. Am J Sports Med. 1988;16(3):285-94. doi: 10.1177/036354658801600316.
5. Walter SD, Hart LE, McIntosh JM, et al. The Ontario cohort study of running-related injuries. Arch Intern Med. 1989;149(11):2561-4. doi: 10.1001/archinte.1989.00390110113025.
6. Kluitenberg B, van Middelkoop M, Smits DW, et al. The NLstart2run study: incidence and risk factors of running-related injuries in novice runners. Scand J Med Sci Sports. 2015;25(5):e515-23. doi: 10.1111/sms.12346.
7. Dallinga J, Van Rijn R, Stubbe J, et al. Injury incidence and risk factors: a cohort study of 706 8-km or 16-km recreational runners. BMJ Open Sport Exerc Med. 2019;5(1):e000489. doi: 10.1136/bmjsem-2018-000489.

8. Viljoen CT, Janse van Rensburg DC, Verhagen E, et al. Epidemiology, clinical characteristics, and risk factors for running-related injuries among South African trail runners. Int J Environ Res Public Health. 2021;18(23):12620. doi: 10.3390/ijerph182312620.
9. Maselli F, Esculier JF, Storari L, et al. Low back pain among Italian runners: a cross-sectional survey. Phys Ther Sport. 2021;48:136-145. doi: 10.1016/j.ptsp.2020.12.023.
10. Wu B, Chen CC, Wang J, et al. Incidence and risk factors of low back pain in marathon runners. Pain Res Manag. 2021;2021:6660304. doi: 10.1155/2021/6660304.
11. Malliaropoulos N, Mertyri D, Tsaklis P. Prevalence of injury in ultra trail running. Hum Mov. 2015;16(2):55-59. doi: 10.1515/humo-2015-0026.
12. Koech RC, Olivier B, Tawa N. A prevalence of running-related injuries among professional endurance runners in the Rift Valley, Kenya. S Afr J Sports Med. 2021;33(1):v33i1a10690. doi: 10.17159/2078-516X/2021/v33i1a10690.
13. Buist I, Bredeweg SW, Bessem B, et al. Incidence and risk factors of running-related injuries during preparation for a 4-mile recreational running event. Br J Sports Med. 2010;44(8):598-604. doi: 10.1136/bjsm.2007.044677.
14. Ellapen TJ, Satyendra S, Morris J, et al. Common running musculoskeletal injuries among recreational half-marathon runners in KwaZulu-Natal. S Afr J Sports Med. 2013;25(2):39-43. doi: 10.7196/sajsm.360.
15. Diz JBM, Dutra MTP, Feijo IC, et al. Low back pain estimates in professional soccer: a systematic review and meta-analysis. Acta Ortop Bras. 2023;31(6):e266012. doi: 10.1590/1413-785220233105e266012.
16. Ansari S, Sharma S. Prevalence and risk factors of chronic low back pain in university athletes: a cross-sectional study. Phys Sportsmed. 2023;51(4):361-370. doi: 10.1080/00913847.2022.2108351.
17. Morimoto M, Okada R, Sugiura K, et al. Low back pain and lumbar degeneration in Japanese professional baseball players. Orthop J Sports Med. 2022;10(10):23259671221125513. doi: 10.1177/23259671221125513.
18. Page MJ, McKenzie JE, Bossuyt PM, et al. The PRISMA 2020 statement: an updated guideline for reporting systematic reviews. BMJ. 2021;372:n71. doi: 10.1136/bmj.n71.
19. Rasmussen CH, Nielsen RO, Juul MS, et al. Weekly running volume and risk of running-related injuries among marathon runners. Int J Sports Phys Ther. 2013;8(2):111-20. PMID: 23593549.
20. Woolf SK, Barfield WR, Nietert PJ, et al. The Cooper River Bridge Run Study of low back pain in runners and walkers. J South Orthop Assoc. 2002;11(3):136-43. PMID: 12539937.
21. Higgins JPT, Morgan RL, Rooney AA, et al. A tool to assess risk of bias in non-randomized follow-up studies of exposure effects (ROBINS-E). Environ Int. 2024;186:108602. doi: 10.1016/j.envint.2024.108602.
22. Vitez L, Zupet P, Zadnik V, et al. Running injuries in the participants of Ljubljana Marathon. Zdr Varst. 2017;56(4):196-202. doi: 10.1515/sjph-2017-0027.
23. Chang WL, Shih YF, Chen WY. Running injuries and associated factors in participants of ING Taipei Marathon. Phys Ther Sport. 2012;13(3):170-4. doi: 10.1016/j.ptsp.2011.08.001.
24. Teixeira RN, Lunardi A, da Silva RA, et al. Prevalence of musculoskeletal pain in marathon runners who compete at the elite level. Int J Sports Phys Ther. 2016;11(1):126-31. PMID: 26900507.
25. Bach DK, Green DS, Jensen GM, et al. A comparison of muscular tightness in runners and nonrunners and the relation of muscular tightness to low back pain in runners. J Orthop Sports Phys Ther. 1985;6(6):315-23. doi: 10.2519/jospt.1985.6.6.315.
26. Lysholm J, Wiklander J. Injuries in runners. Am J Sports Med. 1987;15(2):168-71. doi: 10.1177/036354658701500213.
27. van der Worp MP, de Wijer A, van Cingel R, et al. The 5- or 10-km Marikenloop Run: a prospective study of the etiology of running-related injuries in women. J Orthop Sports Phys Ther. 2016;46(6):462-70. doi: 10.2519/jospt.2016.6402.
28. von Rosen P, Flostrom F, Frohm A, et al. Injury patterns in adolescent elite endurance athletes participating in running, orienteering, and cross-country skiing. Int J Sports Phys Ther. 2017;12(5):822-832. doi: 10.26603/ijspt20170822.
29. Messier SP, Martin DF, Mihalko SL, et al. A 2-year prospective cohort study of overuse running injuries: the Runners and Injury Longitudinal Study (TRAILS). Am J Sports Med. 2018;46(9):2211-2221. doi: 10.1177/0363546518773755.
30. Besomi M, Leppe J, Mauri-Stecca MV, et al. Training volume and previous injury as associated factors for running-related injuries by race distance: a cross-sectional study. J Hum Sport Exerc. 2019;14(3). doi: 10.14198/jhse.2019.143.06.
31. Taunton JE, Ryan MB, Clement DB, et al. A prospective study of running injuries: the Vancouver Sun Run In Training clinics. Br J Sports Med. 2003;37(3):239-44. doi: 10.1136/bjsm.37.3.239.
32. Benca E, Listabarth S, Flock FKJ, et al. Analysis of running-related injuries: the Vienna study. J Clin Med. 2020;9(2):438. doi: 10.3390/jcm9020438.
33. Taylor-Haas JA, Long JT, Garcia MC, et al. The influence of maturation and sex on pelvis and hip kinematics in youth distance runners. J Sci Med Sport. 2022;25(3):272-278. doi: 10.1016/j.jsams.2021.09.193.
34. Engeler N, Lichtenstein E, Faude O, et al. How downhill and uphill running interfere posture and muscle activity: a descriptive laboratory study. Int J Sports Phys Ther. 2025;20(8):1186-1197. doi: 10.26603/001c.142485.
35. Rizzo F, Vallio CS, Hespanhol L. HealthyTrailsBR - the prevalence of running-related injuries and cramps, and the description of personal and running characteristics in Brazilian trail runners: a cross-sectional study. Braz J Phys Ther. 2024;28(5):101117. doi: 10.1016/j.bjpt.2024.101117.
36. Tamer I, Kocak UZ. Incidence and running characteristics associated with running related injuries in runners: a cross-sectional study. Res Sports Med. 2026;34(3):312-321. doi: 10.1080/15438627.2025.2568935.
37. Wang P, Lu X, Wen M, et al. Association between muscle strength and low back pain among middle-aged and older adults: a cross-sectional study. BMC Public Health. 2025;25(1):1869. doi: 10.1186/s12889-025-23050-2.
38. Madsen AL, Gyntelberg F, Marott JL, et al. Cohort study of postural sway and low back pain: the Copenhagen City Heart Study. Eur Spine J. 2023;32(12):4390-4396. doi: 10.1007/s00586-023-07925-9.
39. Ziegeler K, Gensler LS, Link TM, et al. Association between MRI findings and inflammatory symptoms in non-specific chronic low back pain. Eur Spine J. 2025;34(12):5530-5538. doi: 10.1007/s00586-025-09492-7.
40. Tenforde AS, Ackerman KE, Bouxsein ML, et al. Factors associated with high-risk and low-risk bone stress injury in female runners: implications for risk factor stratification and management. Orthop J Sports Med. 2024;12(5):23259671241246227. doi: 10.1177/23259671241246227.
41. Sugiura S, Aoki Y, Oyama T, et al. Low back pain characteristics in adolescent patients with early-stage spondylolysis: a prospective study. Eur Spine J. 2024;33(10):3709-3714. doi: 10.1007/s00586-024-08478-1.
42. Goodrum M, Viljoen C, Kaulback K. Incidence, severity, and risk factors for injuries in female trail runners - a retrospective cross-sectional study. Phys Ther Sport. 2025;71:1-7. doi: 10.1016/j.ptsp.2024.11.004.
43. Moreira PF, Veras PM, Oliveira TMD, et al. Incidence and biomechanical risk factors for running-related injuries: a prospective cohort study. J Clin Orthop Trauma. 2024;57:102562. doi: 10.1016/j.jcot.2024.102562.
44. Burke A, Dillon S, O'Connor S, et al. Aetiological factors of running-related injuries: a 12 month prospective Running Injury Surveillance Centre (RISC) study. Sports Med Open. 2023;9(1):46. doi: 10.1186/s40798-023-00589-1.
45. Schuster Brandt Frandsen J, Hulme A, Parner ET, et al. How much running is too much? Identifying high-risk running sessions in a 5200-person cohort study. Br J Sports Med. 2025;59(17):1203-1210. doi: 10.1136/bjsports-2024-109380.
46. Lindman I, Abrahamson J, Nielsen RO. Runners with a high body mass index and previous running-related problems is a high-risk population for sustaining a new running-related injury: a 18-month cohort study. Eur J Sport Sci. 2025;25(1):e12206. doi: 10.1002/ejsc.12206.
47. Goldberg M, Le Mat Y, Metra M, et al. Poor sleep quality is associated with an increased risk of running-related injuries: a prospective study of 339 runners over six months. Scand J Med Sci Sports. 2025;35(11):e70164. doi: 10.1111/sms.70164.
48. van Vreden C, Schwellnus M, Ramagole D, et al. History of multiple allergies and gradual onset running-related injuries in distance runners: SAFER XXXV. Clin J Sport Med. 2025;35(1):67-74. doi: 10.1097/JSM.0000000000001245.
49. van Hoek JPKD, Cloosterman KLA, de Vos RJ, et al. Sex-differences in characteristics and factors associated with new running-related injuries among recreational runners. Scand J Med Sci Sports. 2024;34(11):e14758. doi: 10.1111/sms.14758.
50. Markussen Bergum J, Lindman I, Abrahamson J, et al. Runners with osteoarthritis sustain more injuries than healthy runners: an observational cohort study. Phys Ther Sport. 2025;76:44-49. doi: 10.1016/j.ptsp.2025.09.008.
51. Li X, Chen Y, Zhao X, et al. Risk factors for running-related injuries among Chinese marathon runners: a cross-sectional study. Prev Med. 2026;204:108510. doi: 10.1016/j.ypmed.2026.108510.

52. Garcia M, Martins TB, Schmidt de Mesquita R, et al. Epidemiology of injuries in road runners: associations with individual characteristics, training, habits and emotional aspects. Res Sports Med. 2026. [Epub ahead of print]. doi: 10.1080/15438627.2026.2641588.
53. Masilana V, Kunene S. Running-related injuries and risks among runners in township running clubs in Soweto, Johannesburg. S Afr J Sports Med. 2026;38(1):v38i1a20618. doi: 10.17159/2078-516X/2026/v38i1a20618.
54. Nair A, Jansen Van Rensburg A, Viljoen CT, et al. Pre-existing injuries and illnesses in trail running: sex-based epidemiological findings from a 2022 forest marathon. Phys Sportsmed. 2026. [Epub ahead of print]. doi: 10.1080/00913847.2026.2680966.
55. Neason C, Samanna CL, Tagliaferri SD, et al. Running is acceptable and efficacious in adults with non-specific chronic low back pain: the ASTEROID randomised controlled trial. Br J Sports Med. 2025;59(2):99-108. doi: 10.1136/bjsports-2024-108245.
56. Wittchen PJ, Vincent GE, Samanna CL, et al. Running improves mental health symptoms and pain catastrophising in adults with chronic low back pain: a secondary analysis of the ASTEROID randomised controlled trial. J Sci Med Sport. 2026;29(1):79-84. doi: 10.1016/j.jsams.2025.08.001.
57. Neason C, Samanna CL, Tagliaferri SD, et al. Fears and beliefs about running in adults with chronic low back pain: a mixed methods study from the ASTEROID randomised controlled trial. Sports Med Open. 2025;11(1):66. doi: 10.1186/s40798-025-00861-6.
58. Shu D, Dai S, Wang J, et al. Impact of running exercise on intervertebral disc: a systematic review. Sports Health. 2024;16(6):958-970. doi: 10.1177/19417381231221125.
59. Samanna CL, Neason C, Tagliaferri SD, et al. Lifetime physical loading and magnetic resonance-derived intervertebral disc health in adults with chronic low back pain: a cross-sectional study. JOR Spine. 2026;9(2):e70186. doi: 10.1002/jsp2.70186.
60. De Wit D, Pillay JD. Low back pain in South African adolescent field hockey players: implications for future practice. J Clin Med. 2025;14(10):3309. doi: 10.3390/jcm14103309.
61. Urbach B, Mansfield C, Harr M, et al. High prevalence of low back pain in college rifle athletes. Int J Sports Phys Ther. 2025;20(1):87-96. doi: 10.26603/001c.127385.
62. Aavikko A, Pesonen J, Ristolainen L, et al. Prevalence of low back pain and disc degeneration in elite cross-country skiers: a comparative study with non-athlete controls. Scand J Med Sci Sports. 2025;35(12):e70187. doi: 10.1111/sms.70187.
63. Nahhas Rodacki CL, Monteiro CA, Paulo AC, et al. Cadence matters: influence of cadence on spinal load during running. Gait Posture. 2024;113:528-533. doi: 10.1016/j.gaitpost.2024.07.298.
64. Drozda D, Thompson Z, Vincent KR, et al. Gait signatures of endurance runners with low back pain: a case controlled cross sectional study. Gait Posture. 2024;113:184-190. doi: 10.1016/j.gaitpost.2024.06.015.
65. Carraro A, Gilic B, Bertolo R, et al. Lower back pain in young climbers: a retrospective cross-sectional study. Front Sports Act Living. 2023;5:1328811. doi: 10.3389/fspor.2023.1328811.
66. Wall J, Cook DL, Meehan WP, et al. Adolescent athlete low back pain diagnoses, characteristics, and management: A retrospective chart review. J Sci Med Sport. 2024;27(9):618-623. doi: 10.1016/j.jsams.2024.05.004.
67. Anselmo-e-Silva CI, Santos-de-Araujo AD, Ferreira-Silva H, et al. Prevalence and predictive factors of musculoskeletal injuries in triathletes: a cross-sectional study. BMC Sports Sci Med Rehabil. 2025;18(1):20. doi: 10.1186/s13102-025-01451-5.
68. Haddadj R, Nordstoga AL, Nilsen TIL, et al. Volume and intensity of walking and risk of chronic low back pain. JAMA Netw Open. 2025;8(6):e2515592. doi: 10.1001/jamanetworkopen.2025.15592.
69. Suri P, Timmons AKI, Korpak AM, et al. Transient and long-term risks of common physical activities in people with low back pain. JAMA Netw Open. 2025;8(12):e2547915. doi: 10.1001/jamanetworkopen.2025.47915.
70. Wong WK, Fu ACL, Tsang SMH. Morphological and compositional features of lumbar multifidus in athletes with chronic non-specific low back pain: a systematic review and meta-analysis. Int J Sports Phys Ther. 2026;21(6):162340. doi: 10.26603/001c.162340.