**Bioscientia Medicina: Journal of Biomedicine and Translational Research**
https://bioscmed.com e-ISSN: 2598-0580 Vol. 10, No. 10 (2026); pages 2967–2979

SYSTEMATIC REVIEW AND META-ANALYSIS

# Acute Changes in Cartilage Oligomeric Matrix Protein and CTX-II Following Endurance Running in Adults: A Systematic Review and Meta-analysis

Made Anggi Paramitha[1*], Bagus Gede Krisna Astayogi[1], Komang Adinda Permitasari Bramastya[1]

[1] Faculty of Medicine, Universitas Mahasaraswati, Denpasar, Indonesia

**ARTICLE INFO**

*Article history*
**Received:** July 15, 2026
**Revised:** August 10, 2026
**Accepted:** August 29, 2026
**Available online:** September 4, 2026
**Published:** September 6, 2026

*Keywords:*
Cartilage oligomeric matrix protein
CTX-II
Endurance running
Cartilage biomarkers
Meta-analysis

**Corresponding author:*
Made Anggi Paramitha
madeanggi940@gmail.com
ORCID: 0009-0003-8169-797X

*DOI:*
https://doi.org/10.37275/bsm.v10i10.1668



**ABSTRACT**

**Background:** Acute running-related changes in cartilage oligomeric matrix protein (COMP) and the C-terminal cross-linked telopeptide of type II collagen (CTX-II) should not be interpreted automatically as structural cartilage injury. This review evaluated the magnitude and consistency of these biomarker responses after endurance running.

**Objective:** To quantify the magnitude and consistency of the immediate post-run change in serum cartilage oligomeric matrix protein (COMP), and to summarize the corresponding CTX-II responses, in adults following endurance running.

**Methods:** The protocol was registered in PROSPERO (CRD420261485170). PubMed/MEDLINE, CENTRAL, ClinicalTrials.gov, WHO ICTRP, OpenAlex, Crossref, and citation networks were searched for adult running studies reporting pre- and post-exercise COMP or CTX-II. The primary effect was the log ratio of immediate post-run to pre-run serum COMP. Random-effects models used restricted maximum likelihood with Hartung-Knapp inference.

**Results:** The search identified 221 unique reports; seven reports met the criteria for qualitative synthesis, and four compatible healthy-adult COMP estimates involving 57 participants entered the meta-analysis. The pooled post/pre ratio was 1.132 (95% CI 0.794–1.613), with $\tau^2$=0.0365, $I^2$=89.0%, and a 95% prediction interval of 0.560–2.286 at ρ=0.50. Three estimates increased and one decreased after running. CTX-II and recovery findings were too heterogeneous for pooling. Certainty was very low.

**Conclusion:** Serum COMP may change acutely after endurance running, but the direction and magnitude are inconsistent. These short-term surrogate responses do not establish cartilage damage, protection, or long-term clinical harm.

## 1. Introduction

Cartilage responds to load through fluid movement, matrix deformation, and changes in extracellular-matrix turnover. Circulating and urinary biomarkers can capture part of this biological response, although they are not specific to a single joint and their acute interpretation differs from their use as prognostic markers in osteoarthritis. Cartilage oligomeric matrix protein (COMP, thrombospondin-5) is a pentameric extracellular-matrix glycoprotein found in cartilage and other load-bearing connective tissues. C-terminal telopeptide of type II collagen (CTX-II) reflects type II collagen degradation and can be measured in urine or serum.

Running provides a repeated high-frequency loading stimulus whose magnitude depends on speed, distance, surface, footwear, and participant characteristics. Studies have reported increases, decreases, or little immediate change in circulating COMP after laboratory running and distance events[1–12]. Differences in pre-exercise rest, biospecimen, assay, sampling window, event duration, hydration, and health status complicate direct comparison. Ultra-endurance events introduce additional physiological changes, including hemoconcentration, inflammatory responses, and prolonged sampling intervals, while studies involving knee osteoarthritis or prior injury may represent distinct biological strata.

Earlier primary studies include 30-minute treadmill protocols, controlled comparisons of running with jumping or cycling, marathons, and ultramarathons[1–12]. However, a pooled estimate is only meaningful when outcome definitions, timepoints, and populations are sufficiently comparable. In particular, an immediate post-run estimand should not be combined indiscriminately with recovery samples, chronic training effects, walking-only experiments, gene-expression outcomes, or non-human models.

The protocol was registered in PROSPERO (CRD420261485170). It prespecified immediate serum COMP as the primary quantitative outcome, while CTX-II, later recovery samples, and clinical populations were analyzed in separate evidence streams. This structure was retained throughout the review so that

biologically different outcomes were not combined solely to increase the number of studies.

The distinction between an acute biochemical response and structural cartilage injury is fundamental. COMP is present in cartilage, tendon, ligament, and other connective tissues, and a temporary rise in blood concentration may reflect transport from loaded tissues, altered lymphatic or vascular clearance, or exercise-related changes in plasma volume as well as matrix turnover. CTX-II is more closely linked to type II collagen degradation, but its concentration also depends on biospecimen, renal handling, diurnal variation, and assay procedures. Neither marker should be interpreted as a direct measure of cartilage thickness, focal lesion progression, pain, or future osteoarthritis without corroborating structural and clinical evidence.

Mechanical loading is biologically necessary for cartilage homeostasis, yet loading dose is multidimensional. A 30-minute laboratory run at controlled speed differs from a competitive marathon in cumulative steps, impact distribution, substrate use, hydration, temperature, and recovery. Even within a nominally identical protocol, joint moments and contact forces vary with anthropometry, running technique, footwear, prior injury, and training status. These differences motivate a random-effects framework and a cautious definition of the target population rather than a universal claim about all running.

Sampling time is equally important. COMP may rise immediately after loading and return toward baseline during the first hour, whereas later measurements can reflect a mixture of clearance, continued release, recovery behavior, and systemic responses. Combining immediate and delayed samples would therefore obscure rather than summarize the biological pattern. The primary timepoint was accordingly defined as the sample obtained within 15 minutes of running cessation, and later samples were assigned to separate recovery windows. Intermediate checkpoints during a named event were treated as event-specific outcomes rather than as post-finish samples.

The review focuses on proportional change because absolute COMP concentrations differ substantially across assays and laboratories. A post/pre ratio is unitless and preserves the direction of change, making it more interpretable across reports that use ng/mL, U/L, or other scales. This choice does not solve every comparability problem: a ratio can still be affected by baseline selection, calibration, floor effects, and imprecise measurement. Consequently, pooling is restricted to estimates that are sufficiently aligned in population, biomarker, biospecimen, exposure, and timepoint.

The clinical motivation is not to decide whether running is categorically safe or harmful. Rather, it is to determine what the available biomarker literature can support about the immediate physiological response and where uncertainty remains. Such evidence may inform the design of mechanistic exercise studies and help clinicians explain why an acute laboratory biomarker change is not equivalent to demonstrated degeneration. Long-term recommendations must continue to rely on the broader evidence concerning symptoms, function, imaging, injury, and osteoarthritis progression.

A prior systematic review and meta-analysis evaluated running in relation to knee cartilage across imaging and biomarker outcomes, while broader reviews summarized exercise-related joint biomarker kinetics and mechanobiology[13–16]. Those syntheses established the relevance of the topic but did not resolve the narrower estimand of an immediate proportional COMP response to sustained running, did not consistently separate recovery windows from the immediate sample, and did not pool CTX-II across incompatible biospecimens. The present review therefore avoids an absolute claim of being the first review and defines novelty by a more specific estimand, outcome hierarchy, and auditable report-to-study linkage.

Novelty and aim: The novelty of this review lies in the prespecified separation of immediate and recovery biomarker kinetics, use of a unitless within-person ratio of means for serum COMP, separate handling of CTX-II by biospecimen, explicit mapping of multiple reports to underlying studies, and a provenance-linked effect dataset. Accordingly, this study aimed to quantify the immediate post-run versus pre-run change in serum COMP among adults completing endurance running, to synthesize CTX-II and later COMP responses without combining biologically incompatible streams, and to examine how population, loading dose, sampling, assay, and risk of bias influence interpretation.

## 2. Methods

### *Protocol and reporting*

The review protocol was finalized before study selection and data extraction and was registered in PROSPERO (CRD420261485170). Protocol development followed PRISMA 2020 and PRISMA-S, and the review was conducted using current systematic-review guidance.[17–19] Amendments were documented with their rationale and potential analytical impact.

The protocol prespecified the review question, eligibility criteria, outcome hierarchy, primary timepoint, effect measure, synthesis model, population strata, sensitivity analyses, risk-of-bias domains, and certainty framework. Piloting of the search strategy refined syntax and platform coverage only; it did not alter the primary estimand or the rules used to select outcomes and timepoints.

PRISMA terminology was applied consistently. A record was an item returned by a database, a report was a document describing a study, and a study was the

underlying investigation that could generate several reports. Report-to-study linkage was completed before quantitative synthesis so that multiple publications from the same participants did not contribute independent effects.

### *Eligibility criteria*

Eligible participants were humans aged 18 years or older. Healthy adults, adults with knee osteoarthritis, and adults with a history of knee injury were all eligible but were not assumed to be exchangeable. The eligible exposure was at least 30 continuous minutes of running or participation in a named distance-running event. Eligible outcomes were serum or plasma COMP concentration and urinary or serum CTX-II measured before and after that exposure. Eligible designs were randomized and non-randomized crossover exercise studies, prospective repeated-measures pre-post studies, and compatible comparative observational studies. Animal, in vitro, and in silico studies were excluded, as were walking-only exposures, mixed exercise without a separable running result, COMP gene-expression outcomes without a circulating concentration, and reports without extractable biomarker data.

The age threshold was applied to the analyzed sample rather than to a broad population label. Mixed adolescent-adult samples were eligible only when adult results were separable or the entire analyzed sample met the threshold. A named distance-running event qualified even when individual duration differed from the laboratory 30-minute threshold, because the event itself defined a sustained endurance exposure. Sprint tests, intermittent sports, isolated jumps, and short walk-run intervals did not satisfy the exposure definition unless an eligible continuous-running result was reported separately.

Acute and chronic evidence streams were kept distinct. The primary question concerned the within-person response to a single running exposure. Cross-sectional comparisons of habitual runners with non-runners and longitudinal training adaptations provided context but addressed different estimands, so they were not combined with acute pre/post effects. Resting-control and low-impact comparator arms were likewise treated as secondary contrasts rather than substitutes for the primary within-person ratio.

Reports were retained whenever they identified a potentially eligible study, even when the numerical result was incomplete, so that conference abstracts, theses, and registry entries could proceed to linkage and full-text assessment. Such reports were excluded from quantitative synthesis only when the effect or a compatible variance remained unavailable. This separation kept the absence of usable numbers distinct from the absence of a biological effect.

### *Information sources and search*

The implemented search covered MEDLINE/PubMed, CENTRAL, ClinicalTrials.gov, WHO ICTRP, OpenAlex, Crossref, and backward and forward citation searching. No language, publication-status, or date restriction was applied. Search concepts combined running, jogging, marathon, ultramarathon, treadmill running, and distance events with terms for COMP, cartilage oligomeric matrix protein, thrombospondin-5, CTX-II, and the C-terminal telopeptide of type II collagen. Platform, syntax, field tags, result count, and raw export were archived for each search.

Controlled vocabulary was combined with text words on every platform that supported it, and each search string was adapted to the capabilities of its database rather than copied verbatim across platforms. Registry searches used condition and intervention fields, and bibliographic-metadata searches were restricted to records carrying both an exposure concept and a biomarker concept.

Citation searching was bidirectional. Reference lists of eligible and near-eligible reports were examined for older studies and precursor abstracts, and forward citations were used to locate later replications, secondary analyses, and dissertations. Seed reports were recorded because citation-network yield depends on which studies are chosen, and every newly identified record passed through the same de-duplication and eligibility assessment as database records.

De-duplication proceeded in stages using normalized DOI, PMID or registry identifier, title, year, and author metadata. Exact identifiers took precedence, and title-based matches were checked individually when punctuation, spelling, or transliteration differed. Because record-level de-duplication does not resolve multiple-report bias, a separate report-to-study map linked journal articles, theses, abstracts, and registry entries describing the same participants. Where linkage remained uncertain, reports were kept separate and flagged so that they could not enter a pooled analysis as independent studies.

### *Study selection and data collection*

Records were screened sequentially at title/abstract and full-text levels against the prespecified criteria. Eligibility decisions, full-text exclusion reasons, and report-to-study links were recorded in the review database. Multiple reports from one study were retained for provenance but were permitted to contribute only one independent estimate to a given meta-analysis.

A structured screening form captured the eligibility domain, reason code, evidence source, narrative rationale, and study identifier for each record. Potentially eligible reports proceeded to full-text assessment when the abstract did not provide enough information on participant age, running duration, biospecimen, sampling time, or biomarker outcome.

An inclusive screening rule was used when eligibility could not be determined from bibliographic information alone. Reports were excluded only when

the available information established an ineligible population, exposure, outcome, design, or non-human model. Duplicate and companion reports were linked rather than deleted, preserving their contribution to study identification and data verification.

Extraction distinguished study, report, participant stratum, exercise condition, biomarker, biospecimen, assay, unit, sampling time, sample size, mean, dispersion, reported ratio, change score, and variance information. Each numerical entry carried a provenance field naming the page, table, figure, supplement, repository abstract, or registry section from which it was taken. Absent values were recorded as not reported and uncertain values were flagged, and no missing result was inferred from a p-value or a narrative statement alone.

Reconciliation addressed interpretation as well as transcription, because an identical number can represent a standard deviation, a standard error, a confidence-interval limit, an adjusted estimate, a subgroup value, or a repeated observation. The analysis dataset therefore retained the originally reported scale, any transformation formula, the assumptions applied, and the reconciled analytical role of each value. Repeated sessions in the same participants and multiple timepoints were indexed explicitly to prevent unit-of-analysis errors.

Full texts were obtained from PubMed Central, publisher open-access platforms, institutional repositories, and author-posted manuscripts. A record that resolved only to a landing page was not counted as retrieved. When an abstract supplied candidate values but the corresponding full document could not be obtained, the evidence level was recorded as repository abstract and the values were used only for narrative synthesis.

### *Outcomes and effect measures*

The primary outcome was serum COMP sampled within 15 minutes of running cessation, taking the draw closest to cessation when several were available. Recovery COMP was grouped into prespecified later windows, and CTX-II was synthesized separately by biospecimen and timepoint. The primary effect measure was the natural logarithm of the post-run to pre-run ratio of means (lnROM), back-transformed to a ratio, with values above 1 indicating a higher post-run concentration. Reported ratios and confidence intervals were used directly when compatible; where paired variance was not reported, sensitivity analyses assumed within-person correlations of 0.25, 0.50, and 0.75.

The estimand hierarchy was specified before pooling. First priority was the immediate within-person serum COMP ratio in healthy adults. Second-priority streams were immediate COMP in prespecified clinical strata, recovery COMP within defined windows, and CTX-II separated by serum and urine. Where a report provided several eligible immediate samples, the draw closest to exercise cessation was selected; where several reports described one cohort, the most complete compatible estimate was used and all contributing reports remained linked in the provenance map. Fixing this hierarchy in advance prevented outcome and timepoint multiplicity from being resolved after inspection of statistical significance.

The logarithmic transformation renders proportional increases and decreases symmetric for modeling. For a report providing pre- and post-run means, the point estimate was ln(mean post/mean pre), and its variance depended on the standard errors of both means and on their within-person correlation. Where a report gave a ratio with a compatible confidence interval, those values were used preferentially because they preserved the paired analysis. Where only separate standard deviations were available, the correlation sensitivity analyses quantified how the unreported pairing affected precision.

A single study could contribute more than one eligible participant stratum, but correlated strata were not treated as independent. Healthy controls and participants with prior knee injury are clinically distinct, so the primary healthy-adult analysis used the healthy stratum only. Where several exercise conditions were studied in the same participants, the prespecified eligible running condition was selected rather than counting each session as a separate study, which would have exaggerated precision.

Percentage changes reported without a measure of uncertainty were retained for narrative synthesis and direction-of-effect summaries. They were not assigned an assumed standard error and were not entered into inverse-variance pooling. Where a report did not state whether error bars represented a standard deviation, a standard error, or a confidence interval, the value was not reconstructed for the primary model.

### *Synthesis and certainty*

Clinically and methodologically compatible estimates were synthesized by inverse-variance random-effects meta-analysis with restricted maximum likelihood, following current review guidance[19]. Hartung-Knapp intervals were used because at least three studies were pooled[19]. Statistical heterogeneity was summarized with $\tau^2$, $I^2$, Cochran's Q, and a prediction interval. Healthy adults and clinical populations were analyzed separately. Small-study-effect tests were prespecified to require at least ten compatible studies and were therefore not performed. Risk of bias was assessed with ROBINS-I for non-randomized intervention evidence, with RoB 2 reserved for randomized trials, and with attention to the target effect at result level[20]. GRADE was applied by outcome[19].

For studies reporting pre- and post-run means, the study-specific effect and paired sampling variance were calculated as follows; directly reported normalized ratios and confidence intervals were

converted to the log scale and preferred when they preserved the paired analysis.

$y_i = \ln(M_{post,i} / M_{pre,i})$

$Var(y_i) = SD_{post,i}^2 / (n_i M_{post,i}^2) + SD_{pre,i}^2 / (n_i M_{pre,i}^2)$

$- 2\rho SD_{post,i} SD_{pre,i} / (n_i M_{post,i} M_{pre,i})$.

The choice of random effects reflects the scientific expectation that true acute responses differ across settings, not merely a statistical test for heterogeneity. REML estimates between-study variance, while Hartung-Knapp inference acknowledges uncertainty in that estimate when few studies are available. With only four estimates, both the pooled confidence interval and prediction interval deserve greater weight than a binary significance test. The prediction interval addresses the range in which a future comparable study's true effect might lie under the model.

Prespecified sensitivity analyses varied the assumed paired correlation, compared alternative random-effects inference where appropriate, and assessed influential observations by leave-one-out estimation. These analyses were interpretive rather than a search for a preferred significant result, and instability across plausible assumptions is reported as such rather than resolved by selecting one model after inspecting the output.

Risk of bias was assessed at the result level because outcomes and timepoints within one study could have different vulnerabilities. The domains addressed uncontrolled time-varying confounding, participant selection, exposure classification, deviations from the running protocol, missing samples, assay and sampling procedures, and selective result reporting. Judgements were supported by study-specific evidence and were considered in the certainty assessment.

Certainty was assessed by outcome using GRADE. Non-randomized repeated-measures evidence began at low certainty and could be rated down for risk of bias, inconsistency, indirectness, imprecision, or suspected publication bias. Rating up would have required defensible evidence such as a large effect or a dose-response gradient. The resulting certainty statement concerns confidence in the estimated acute biomarker response, not the long-term safety or harm of running.

Funnel-plot asymmetry tests were not performed, because such tests are underpowered with four studies and their assumptions are unlikely to hold under marked clinical heterogeneity. Meta-regression was likewise not performed, since the number of independent estimates was inadequate. Potential modifiers including event duration, assay, age, sex, knee health, and sampling delay were therefore described narratively and reserved for subgroup analysis should sufficient data become available.

## 3. Results

### *Study selection*

The PubMed search returned 106 records, ClinicalTrials.gov returned 12 registry records, forward citation searching contributed 56 additional unique reports, CENTRAL contributed 36, and supplementary OpenAlex and Crossref searches contributed 11. After record-level de-duplication, the evidence library contained 221 unique reports. The source-specific contributions and de-duplication trail were retained for reproducibility.

Of the 221 screened reports, 178 were excluded at title and abstract and 2 were removed as duplicate reports of an already-identified study, leaving 41 reports sought for retrieval. Twenty-one could not be obtained in full and were assessable only from an abstract or bibliographic record; 20 were assessed at full text, of which 13 were excluded (11 ineligible exposure, 1 non-human, 1 ineligible outcome). Seven reports therefore met the eligibility criteria for qualitative synthesis, and four provided compatible immediate serum-COMP estimates for quantitative synthesis. The remaining included reports contributed to the narrative assessment of recovery kinetics, longer-distance events, alternative biospecimens, or outcomes lacking a compatible variance estimate. The complete flow of records is shown in Supplementary Figure S2.

Supplementary searches identified dissertations, conference reports, earlier abstracts, and companion publications that enriched the evidence map. These sources were linked to their corresponding study whenever participant characteristics, exposure protocol, outcome timing, and bibliographic information indicated a shared cohort. Reports without sufficient numerical information contributed only to the descriptive synthesis.

Record-level de-duplication and study-level linkage were treated as distinct procedures. The Woodland thesis was linked to the later Osteoarthritis and Cartilage report,[21] the Denning dissertation to its Gait & Posture publication,[22] the Cattano repository report to the Journal of Athletic Training article,[3] and an earlier marathon abstract to the later report by Neidhart et al.[23] Each linked set was treated as one study for quantitative purposes.

An eligibility reason and an evidence level were recorded for every screened report. Irrelevant hits arose most often because the abbreviation COMP occurred within an unrelated word or label, because running was mentioned only as a performance test, or because the record described animal treadmill experiments, adolescent sport samples, or walking and cycling exposures. Records advanced when they combined a sustained running exposure with a circulating COMP or CTX-II outcome, and were carried forward to full-text assessment whenever duration, adult status, biospecimen, or report identity could not be established from bibliographic metadata.

### *Study characteristics*

The seven included reports covered continuous 30-minute treadmill running, a 5-km run in older adults, marathons, and a 308-km ultramarathon. Sample sizes

ranged from 14 to 46 in the relevant running streams. Most participants were healthy or recreationally trained adults; one report separated participants with previous knee injury from matched controls. COMP was measured in serum or plasma using different ELISA platforms and reporting units. CTX-II was reported less consistently and was synthesized separately. Population, exposure, specimen, timepoint, and analytical role are detailed in Table 1.

**Table 1. Characteristics and analytical role of the included reports.**

| Report/population | Exposure/specimen | Outcome, timepoint, and analytical role |
|---|---|---|
| Shin 2012<br>Healthy runners; n not pooled | 308-km event<br>Serum | COMP serial checkpoints; narrative because variance incompatible |
| Vuolteenaho 2014<br>Trained men; n=46 | Marathon<br>Plasma | Immediate COMP; narrative because post mean unavailable |
| Cattano 2017<br>Healthy/prior injury; n=11 each | 30-min treadmill<br>Serum | COMP and CTX-II <=10 min; healthy COMP in meta-analysis |
| Dreiner 2020<br>Healthy men; n=15 | 30-min treadmill<br>Serum | Immediate COMP; meta-analysis |
| Hernández-Hermoso 2021<br>Recreational men; n=17 | Marathon<br>Serum | COMP immediate/48 h; immediate COMP in meta-analysis |
| Dreiner 2022<br>Healthy men; n=14 | 30-min treadmill<br>Serum | COMP immediate/30/60 min; direct ratio in meta-analysis |
| Hay 2024<br>Adults >=50 y; n=20 runners | 5-km run<br>Serum | Immediate COMP; narrative because variance unavailable |

Additional repository evidence broadened the descriptive map but did not change the four-estimate meta-analysis. The Woodland thesis described 12 adults completing control, sham, and experimental-pain 30-minute running sessions.[21] Across sessions, serum COMP increased from 116.02 to 132.19 ng/mL immediately after running before declining toward baseline at 60 minutes. Because the dispersion measure was not unambiguously defined and the repeated sessions involved the same participants, this evidence was retained narratively rather than assigned an assumed independent variance.

The Denning dissertation described 18 adults, equally divided by sex, who completed 4000 treadmill steps at slow walking, medium jogging, and fast running speeds.[22] Immediate serum COMP increased by 5%, 18%, and 28%, respectively, suggesting a speed- or load-related pattern. The report was not pooled because compatible variance was unavailable and the duration of the running condition could not be aligned confidently with the prespecified 30-minute exposure threshold.

The included evidence base was dominated by small repeated-measures studies. Laboratory protocols afforded better control of speed and sampling but often enrolled homogeneous groups of healthy men and imposed a fixed session order. Field events provided ecologically valid endurance exposure but introduced uncontrolled variation in pace, environmental conditions, food and fluid intake, and sample timing. The evidence map therefore comprised complementary designs rather than interchangeable replications.

Outcome reporting was uneven. Dreiner et al. (2022) provided normalized ratios with confidence intervals, whereas several reports gave pre- and post-run means without a paired correlation or a standard deviation of change[6]. Hay et al. reported a 29% immediate increase after a 5-km run but no compatible variance[7]. Shin et al. reported progressively larger percentage changes at ultramarathon checkpoints without a variance suitable for the immediate model[1]. Vuolteenaho et al. described no mean post-marathon change despite marked individual variability, and the exact post-run value was not reported numerically[2,24].

### *Outcome-specific evidence map*

COMP and CTX-II were analyzed as distinct outcome streams because they are not analytically interchangeable. COMP was the only biomarker with four compatible immediate estimates. CTX-II remained a narrative stream because the evidence mixed serum and urinary measurements, used different collection windows, and provided too few compatible independent estimates. Recovery COMP was also separated from the immediate estimand because clearance and post-exercise behavior change the biological question. The outcome hierarchy is detailed in Table 2.

**Table 2. Prespecified outcome map and synthesis approach for each evidence stream.**

| Outcome stream | Scope/effect | Current synthesis |
|---|---|---|
| Immediate serum COMP | Healthy adults; <=15 min<br>ln post/pre ratio | 4 estimates; random-effects meta-analysis |
| Immediate COMP, clinical strata | Knee injury or osteoarthritis; <=15 min<br>ln post/pre ratio | Narrative synthesis; no compatible quantitative pool |
| Recovery COMP | >15–30 min, >30–120 min, >24–48 h<br>Window-specific ratio | Narrative; never combined with immediate sample |
| Serum CTX-II | Immediate and recovery separated<br>Ratio or change | Narrative; sparse and highly dispersed |
| Urinary CTX-II | Time-normalized collection windows<br>Creatinine-adjusted ratio | Narrative; not pooled with serum CTX-II |

For CTX-II, Cattano et al. reported serum values for prior-injury and matched-control strata, but both were small (n=11 each) and the standard deviations were large relative to the means[3]. Marathon reports described type-II collagen synthesis and catabolism markers using different biological constructs and timepoints[2,5]. Contemporary CTX-II reviews also emphasize biospecimen, diurnal, renal, and post-injury context[25–26]. Consequently, absence of a pooled CTX-II estimate is a prespecified comparability decision, not evidence that CTX-II is unimportant or unchanged.

### *Quantitative synthesis of immediate serum COMP*

Effect inputs, variance sources, and report-level provenance are detailed in Table 3.

**Table 3. Effect-level provenance for the immediate serum-COMP meta-analysis.**

| Estimate | Effect input and ratio | Variance and provenance |
|---|---|---|
| Cattano 2017, healthy controls<br>n=11 | Pre: 830.32 (381.30) pg/mL<br>Post: 660.93 (382.33) pg/mL<br>Ratio: 0.796 | Means/SD; rho sensitivity<br>Table 3 |
| Dreiner 2020, running<br>n=15 | Pre: 9.10 (2.98) U/L<br>Post: 11.50 (3.16) U/L<br>Ratio: 1.264 | SD reconstructed from 95% CI; rho sensitivity<br>Reported table/CI |
| Hernández-Hermoso 2021<br>n=17 | Pre: 222.50 (56.86) ng/mL<br>Post: 233.43 (54.95) ng/mL<br>Ratio: 1.049 | Means/SD; rho sensitivity<br>Table 2 |
| Dreiner 2022, running<br>n=14 | Pre: 456.2; dispersion as CI<br>Post: 617.8; dispersion as CI<br>Ratio: 1.368 | Reported normalized ratio, 95% CI 1.252–1.483<br>Supplementary Tables S1-S2 |

Four healthy-adult estimates had compatible immediate serum-COMP information: the healthy-control subgroup from Cattano et al.,[3] the running condition from Dreiner et al. (2020),[4] the marathon cohort from Hernández-Hermoso et al.,[5] and the running condition from Dreiner et al. (2022).[6] At an assumed paired correlation of 0.50, the REML/Hartung-Knapp pooled post/pre ratio was 1.132 (95% CI 0.794–1.613), equivalent to a 13.2% average increase. Heterogeneity was substantial ($\tau^2$=0.0365; $I^2$=89.0%; Q=21.902, p<0.001), and the 95% prediction interval was 0.560–2.286.

The contributing estimates were directionally inconsistent: one subgroup decreased after running, whereas the other three increased. Differences in assay, baseline concentration, participant selection, event duration, pre-exercise standardization, and variance construction plausibly contributed to the heterogeneity. The 5-km, marathon, and ultramarathon reports without compatible variance were not forced into the model. Study-specific ratios and the pooled estimate are shown in Figure 1.

Correlation sensitivity yielded pooled ratios of 1.149 (95% CI 0.819–1.613; $I^2$=82.8%) at ρ=0.25, 1.132 (95% CI 0.794–1.613; $I^2$=89.0%) at ρ=0.50, and 1.114 (95% CI 0.770–1.610; $I^2$=94.2%) at ρ=0.75. Leave-one-out analysis at ρ=0.50 produced ratios from 1.050 when Dreiner et al. (2022) was omitted to 1.223 when the Cattano healthy-control stratum was omitted; every confidence interval included 1.0, and residual $I^2$ ranged from 80.6% to 92.8%. The complete correlation sensitivity results are given in Table 4 and displayed in Supplementary Figure S1. The positive point estimate was therefore not produced by a single study, but its magnitude and precision were unstable and heterogeneity remained substantial.

**Table 4. Correlation sensitivity of the random-effects serum-COMP meta-analysis.**

| Assumed ρ | Pooled ratio (95% CI) | $\tau^2$ / $I^2$ | 95% prediction interval |
|---|---|---|---|
| 0.25 | 1.149 (0.819–1.613) | 0.0294 / 82.8% | 0.604–2.184 |
| 0.50 | 1.132 (0.794–1.613) | 0.0365 / 89.0% | 0.560–2.286 |
| 0.75 | 1.114 (0.770–1.610) | 0.0457 / 94.2% | 0.514–2.414 |

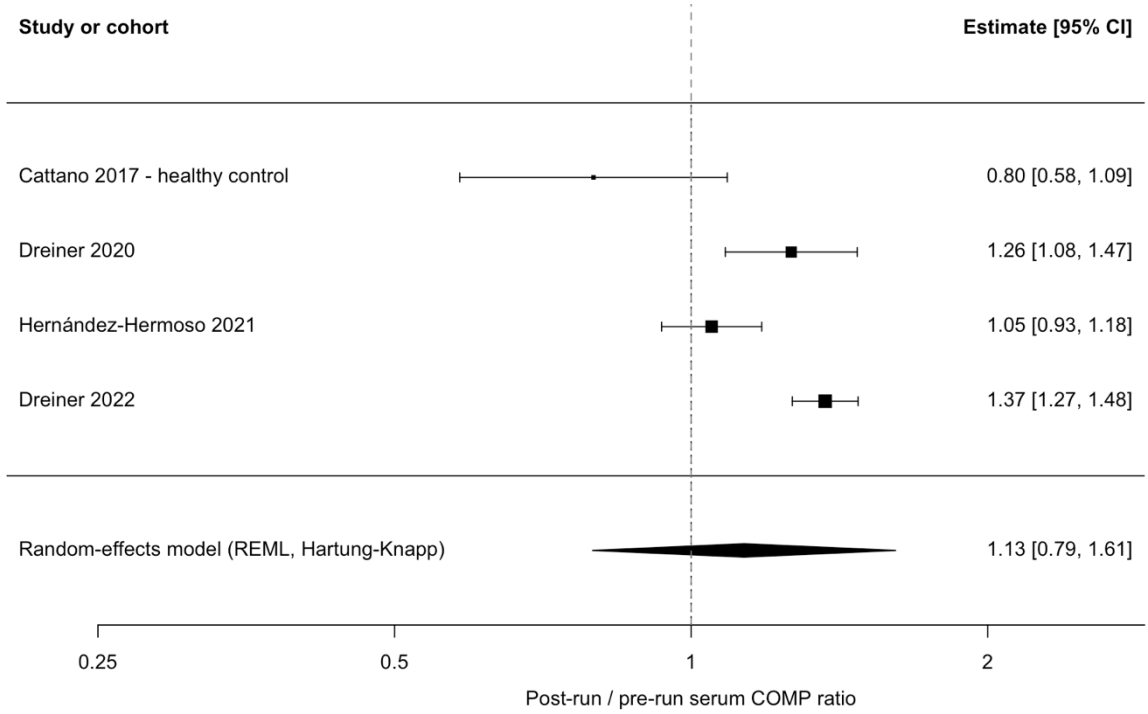


**Figure 1.** Random-effects forest plot for the immediate post-run versus pre-run serum-COMP ratio. Ratios above 1.0 indicate higher post-run concentrations. The model used REML with Hartung-Knapp inference and an assumed within-person correlation of 0.50.

The healthy-control subgroup of Cattano et al. decreased from 830.32 to 660.93 pg/mL, whereas Dreiner et al. (2020) increased from 9.1 to 11.5 U/L, Hernández-Hermoso et al. increased from 222.50 to 233.43 ng/mL, and Dreiner et al. (2022) reported a normalized ratio of 1.368[3–6]. These values illustrate why a unitless ratio is useful but also why a pooled mean alone is insufficient. The observed range includes a substantial decrease and increases of different magnitude, so clinical interpretation must foreground heterogeneity.

### *Risk of bias and certainty of evidence*

The four pooled studies were judged to have serious overall risk of bias. The most frequent concern was confounding in uncontrolled pre/post biomarker contrasts: acute plasma-volume shifts, circadian variation, hydration, food intake, and time effects could alter observed concentrations independently of cartilage matrix release. Field studies also faced environmental and behavioral variation, while fixed exercise order and incomplete paired-variance reporting affected several laboratory studies.

Study-specific concerns differed. In Cattano et al., large dispersion, possible outliers, assay-unit verification, and a blood-draw window of up to 10 minutes raised outcome-measurement concerns[3]. Dreiner et al. (2020) scheduled the running session last, creating potential period and carryover bias[4]. Hernández-Hermoso et al. analyzed a restricted final sample and measured COMP in singlet without plasma-volume correction[5]. Dreiner et al. (2022) also used a fixed session order and had one participant with insufficient serum, although laboratory analysis was blinded and normalized ratios were well reported[6].

Certainty was rated very low for the immediate serum-COMP ratio in healthy adults. The evidence began at low certainty because it consisted of non-randomized repeated-measures studies and was downgraded for serious risk of bias, very serious inconsistency, and serious imprecision. Indirectness was not downgraded for the narrowly defined biomarker response, although COMP remains an indirect surrogate for clinical cartilage damage. Publication bias could not be assessed meaningfully because only four studies contributed to the meta-analysis.

## 4. Discussion

### *Principal findings*

The included evidence suggests that serum COMP can respond acutely to endurance running, but it does not support a single precise magnitude or direction. The pooled estimate was above 1.0, yet its confidence interval included both a reduction and a substantial increase, and the prediction interval was very wide. This pattern is compatible with variation in loading protocols, biospecimen handling, assays, participant characteristics, and sampling schedules. A transient concentration increase may reflect matrix transport, acute turnover, vascular or lymphatic clearance, or plasma-volume change and should not automatically be interpreted as cartilage injury.

Distance events cannot be assumed equivalent to standardized laboratory running. Marathon and ultramarathon studies differ in exercise duration, hydration, nutritional intake, environmental exposure,

and sampling logistics, and a sample drawn at 100 or 200 km does not address the same estimand as an immediate post-finish sample. Older adults, people with knee osteoarthritis, and individuals after knee injury may also differ in baseline COMP concentration and mechanosensitivity. Population and timepoint strata were therefore kept separate throughout the synthesis.

The broadest defensible interpretation is that circulating COMP is mechanically responsive in some running contexts. Evidence for a consistent direction is weaker. Three pooled estimates increased, one decreased, one open 5-km report described an increase, one marathon report described no mean change, and ultramarathon reports suggested time-varying increases. This mixture is compatible with true effect modification and measurement differences; it is not well represented by a single claim that running always raises or lowers COMP.

An acute rise should not be described as proof of degradation. Mechanical compression can increase the movement of matrix molecules or fragments into synovial fluid and the circulation without indicating irreversible tissue loss. Conversely, a fall in serum concentration does not prove protection because distribution volume, clearance, and tissue transport can also change. The outcome is best understood as a systemic biomarker response to loading whose structural and clinical meaning remains conditional.

The very-low certainty of evidence limits clinical conclusions. The point estimate was compatible with a moderate average increase, but the confidence interval included a decrease and the prediction interval encompassed effects in both directions. No validated threshold links a short-term COMP ratio to pain, injury, imaging progression, or osteoarthritis incidence. These findings should therefore not be used either to discourage otherwise appropriate running or to claim a cartilage-protective effect.

The findings refine rather than overturn previous syntheses. Dong et al. found no simple evidence that running causes uniform adverse knee-cartilage change across imaging and biomarker outcomes.[13] Reviews of exercise biomarkers similarly describe rapid, load-sensitive responses whose direction depends on sampling, tissue transport, and the distinction between acute exercise and chronic adaptation.[14–16,25,27–28] The present serum-COMP synthesis adds an explicit proportional estimand and prediction interval, but the wide interval and very-low certainty prevent stronger clinical claims.[16,29–33]

Mechanistic studies help explain the observed inconsistency. Running and cycling can produce different serum COMP and cartilage-thickness responses[15], while intra-articular COMP may decrease even when circulating concentrations change in another direction[34]. Walking, drop landing, and controlled ambulatory loading also evoke acute biomarker responses[14,35]. Studies of COMP fragmentation and loaded versus unloaded exercise further show that assay target and impact exposure affect the measured signal[15,24,36]. These observations caution against treating serum concentration as a direct tissue balance and support the decision to keep loading modes and timepoints distinct[14–15,34,37–38].

Population and loading dose may modify the response. Studies after anterior cruciate ligament injury and in older adults indicate that lower-limb loading, delayed COMP kinetics, cartilage composition, and age can alter load-response relationships[39–41]. Extreme-distance events produce serial changes across thousands of kilometers, showing that cumulative dose and recovery cannot be represented by a single immediate laboratory sample[32]. Evidence from endurance athletes, marathon cohorts, and habitual-activity analyses further supports heterogeneity by training status and systemic context[16,32–33,36,39–43].

### *Methodological implications*

Incomplete paired-data reporting is a central methodological limitation. Several reports provide pre- and post-run means but omit the within-person correlation, the standard deviation of change, and a compatible confidence interval. Treating such repeated measures as independent loses efficiency and can misstate variance, so the present analysis prioritized directly reported paired ratios and used correlation sensitivity analyses elsewhere. Reports lacking this information cannot enter an inverse-variance model without reconstruction, and any reconstruction should be documented explicitly.

Assay diversity is another source of heterogeneity. COMP values were reported in different units and measured with different ELISA kits. The ratio-of-means effect measure avoids direct unit conversion but does not remove differences in calibration, analytical sensitivity, pre-analytic handling, or biological specificity. CTX-II should remain separate from COMP and urinary CTX-II should not be pooled with serum CTX-II without a defensible model[27,44–45].

Pre-exercise standardization deserves more attention than it commonly receives. COMP can respond to ordinary ambulation and posture, so the duration of seated or supine rest before baseline sampling may influence the observed ratio. Recent exercise, training load, time of day, fasting status, and hydration can further alter baseline and post-run concentrations. Analytical assay precision cannot correct a poorly standardized biological baseline[30,35].

Plasma-volume adjustment is methodologically important but should not be applied opaquely. Endurance exercise can concentrate or dilute blood, and an apparent biomarker change may partly reflect fluid redistribution. Experimental estimates confirm that running, cycling, and adrenergic stimulation can produce material plasma-volume changes[46]. Some field studies attempted correction, while others reported raw concentrations. The review therefore preserves

the authors' reported result and any transparent corrected sensitivity rather than imposing an undocumented universal adjustment.

The report-to-study map is more than a bibliographic housekeeping tool. A dissertation may contain complete methods and several timepoints, while the journal article provides peer-reviewed summary results. A conference abstract may precede the full paper by years. Counting these documents independently would inflate study number and precision, while discarding the earlier report could lose unique data. The correct unit is the underlying cohort, with report-specific provenance for each extracted result.

Small samples amplify the influence of individual observations and distributional assumptions. Several COMP datasets are highly dispersed, and a mean ratio may be sensitive to outliers or skewness. When authors analyze data nonparametrically but report means and standard deviations, the synthesis must acknowledge the mismatch. Median-based evidence may require separate narrative presentation unless individual data or a defensible transformation is available.

A narrow primary estimand improves coherence but reduces the number of poolable studies. That tradeoff is appropriate: combining incompatible timepoints or biospecimens would create a larger but less meaningful meta-analysis. Narrative synthesis is not a failure when variance is missing or biological comparability is poor. Direction, magnitude, timing, and methodological context can be reported transparently without manufacturing precision.

### *Strengths and limitations*

Strengths of the review include prospective registration, a prespecified primary timepoint and effect measure, explicit separation of immediate and recovery outcomes, report-to-study linkage, effect-level provenance, and a reproducible random-effects analysis. The study-selection rules also prevented walking-only exposures, non-concentration outcomes, and incompatible biospecimens from entering the primary synthesis.

The evidence base has important limitations. Database coverage was narrower than originally planned, and some potentially relevant reports did not provide sufficient full-text or variance information for quantitative use. The meta-analysis was based on four small studies involving 57 participants, largely healthy adults and men. These features limit precision, representativeness, and the ability to examine modifiers such as age, sex, previous knee injury, running distance, and assay platform.

Clinical and methodological heterogeneity further limits generalization. Laboratory treadmill running and field marathons differ in cumulative load, hydration, environment, and recovery; serum and plasma COMP are not fully interchangeable; and paired variance was incompletely reported. Consequently, the pooled average should be read as a summary of the included conditions rather than a universal response to all endurance running.

Interpretation is also constrained by the surrogate nature of the outcomes. COMP is found in several load-bearing connective tissues, and CTX-II varies with biospecimen, renal handling, diurnal timing, and assay procedures. None of the pooled studies established that an acute biochemical change predicts pain, focal cartilage loss, imaging progression, or future osteoarthritis. Mechanistic and clinical conclusions must therefore remain separate.

The meta-analysis included only four estimates and 57 participants. Hartung-Knapp inference and a prediction interval appropriately widened uncertainty, but statistical methods cannot compensate for sparse evidence. Estimation of between-study variance is unstable with few studies, subgroup analysis would be underpowered, and formal publication-bias testing was not appropriate. The pooled value should therefore be interpreted together with the study-specific effects and prediction interval.

The review also inherits the limitations of its source reports. Selective recruitment, male-dominant samples, incomplete reporting of paired variance, fixed session order, heterogeneous assay platforms, and inconsistent plasma-volume handling all constrain inference. A small number of candidate values were available only from repository abstracts rather than full reports; these were labeled accordingly and kept out of the primary analysis wherever the residual uncertainty prevented valid variance estimation.

The title's inclusion of CTX-II is justified by the protocol but requires restraint in the conclusions. CTX-II is biologically closer to type-II collagen degradation than COMP, yet serum and urinary CTX-II are not interchangeable and are influenced by renal handling, timing, injury state, and assay design. Recent evidence in anterior cruciate ligament reconstruction and knee osteoarthritis underscores that clinical context matters[26,47]. The current evidence map therefore treats CTX-II as an important but unpooled secondary stream rather than granting it equal quantitative weight to COMP[26,28,47].

### *Research implications*

Future primary studies should standardize pre-exercise rest, hydration assessment, running dose, blood-draw timing, assay details, and paired statistical reporting. Reporting raw or individual-level pre/post values would permit reliable within-person synthesis. Studies should distinguish acute transport or turnover signals from structural damage and incorporate imaging or longer-term clinical outcomes where possible[28,44,48–52].

A useful core protocol would document activity during the preceding 24–48 hours, seated or supine rest before baseline sampling, time of day, fasting and hydration, menstrual and hormonal context when

relevant, running speed and distance, footwear, surface, environmental conditions, and exact blood or urine collection times. Hematocrit or plasma-volume indicators would permit sensitivity analysis. Samples should be randomized across assay plates, analyzed in duplicate where feasible, and reported with coefficients of variation and detection limits.

Future reports should provide participant-level or at least paired summary data: pre and post means, standard deviations, mean change and its standard deviation, within-person correlation, and the number contributing to each timepoint. Confidence intervals for ratios or changes are especially helpful. When several exercise conditions are studied in the same participants, the correlation between conditions should be available so that multilevel or multivariate synthesis can preserve the crossover design.

Research should also connect biochemical kinetics with outcomes that matter clinically. Serial COMP and CTX-II measurements could be paired with quantitative cartilage imaging, symptoms, recovery, and longitudinal follow-up. Such designs may distinguish transient transport from persistent matrix catabolism and identify whether baseline joint health, age, sex, prior injury, training history, or running biomechanics modify the response. Until then, biomarker findings should remain mechanistic rather than prescriptive[29,31,53].

For evidence synthesis, prospective harmonization and data sharing would be more valuable than simply increasing the number of small heterogeneous studies. A common data dictionary and pre-agreed sampling windows could enable individual-participant meta-analysis, explore nonlinear dose-response, and reduce dependence on reconstructed variance. Registering analysis plans before examining pooled effects would further limit selective choices among biomarkers, timepoints, and loading conditions.

## 5. Conclusion

Available adult evidence indicates that serum COMP may change acutely after endurance running, but the magnitude and direction vary substantially across studies. The pooled ratio was above 1.0, while both its confidence interval and prediction interval included effects in opposing directions. CTX-II and recovery findings were too heterogeneous for defensible pooling. These short-term surrogate biomarker responses do not establish cartilage degradation, protection, or long-term clinical harm.

## Declarations

### *Registration*

PROSPERO CRD420261485170 (https://www.crd.york.ac.uk/PROSPERO/view/CRD420261485170).

### *Funding*

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

### *Conflicts of interest*

The authors declare no conflicts of interest.

### *Author contributions*

MAP, BGKA, and KAPB contributed equally to Conceptualization, Methodology, Investigation, Data Curation, Formal Analysis, Validation, Writing - Original Draft, Writing - Review & Editing, and Project Administration. MAP is guarantor.

### *Ethics*

Ethics committee approval was not required because this study synthesizes published and publicly registered data.

### *Data and code availability*

Search strategies, bibliographic exports, de-duplication records, effect-level extraction data, statistical code, and generated results are retained by the authors and can be made available with the publication materials or in a persistent repository.

### *AI and automated tools disclosure*

The authors did not use generative artificial intelligence to generate the scientific content, the data extraction or the statistical analysis reported in this manuscript. The authors verified the manuscript and take full responsibility for the accuracy and integrity of its content.

## 6. References

1. Shin KA, Kim AC, Kim YJ, et al. Effect of Ultra-marathon (308 km) Race on Bone Metabolism and Cartilage Damage Biomarkers. Ann Rehabil Med. 2012;36(1):80-7. doi:10.5535/arm.2012.36.1.80.
2. Vuolteenaho K, Leppänen T, Kekkonen R, et al. Running a marathon induces changes in adipokine levels and in markers of cartilage degradation--novel role for resistin. PLoS One. 2014;9(10):e110481. doi:10.1371/journal.pone.0110481.
3. Cattano NM, Driban JB, Barbe MF, et al. Biochemical Response to a Moderate Running Bout in Participants With or Without a History of Acute Knee Injury. J Athl Train. 2017;52(6):567-574. doi:10.4085/1062-6050-51.5.09.
4. Dreiner M, Willwacher S, Kramer A, et al. Short-term Response of Serum Cartilage Oligomeric Matrix Protein to Different Types of Impact Loading Under Normal and Artificial Gravity. Front Physiol. 2020;11:1032. doi:10.3389/fphys.2020.01032.
5. Hernández-Hermoso JA, Nescolarde L, Roca E, et al. Marathon Running Increases Synthesis and Decreases Catabolism of Joint Cartilage Type II Collagen Accompanied by High-Energy Demands and an Inflamatory Reaction. Front Physiol. 2021;12:722718. doi:10.3389/fphys.2021.722718.
6. Dreiner M, Munk T, Zaucke F, et al. Relationship between different serum cartilage biomarkers in the acute response to running and jumping in healthy male individuals. Sci Rep. 2022;12(1):6434. doi:10.1038/s41598-022-10310-z.
7. Hay AM, Rhoades MJ, Bangerter S, et al. Serum Cartilage Oligomeric Matrix Protein Concentration Increases More After Running Than Swimming for Older People. Sports Health. 2024;16(4):534-541. doi:10.1177/19417381231195309.
8. Kersting UG, Stubendorff JJ, Schmidt MC, et al. Changes in knee cartilage volume and serum COMP concentration after running exercise. Osteoarthritis Cartilage. 2005;13(10):925-34. doi:10.1016/j.joca.2005.06.005.
9. Kim HJ, Lee YH, Kim CK. Changes in serum cartilage oligomeric matrix protein (COMP), plasma CPK and plasma hs-CRP in

relation to running distance in a marathon (42.195 km) and an ultra-marathon (200 km) race. Eur J Appl Physiol. 2009;105(5):765-70. doi:10.1007/s00421-008-0961-x.
10. Niehoff A, Kersting UG, Helling S, et al. Different mechanical loading protocols influence serum cartilage oligomeric matrix protein levels in young healthy humans. Eur J Appl Physiol. 2010;110(3):651-7. doi:10.1007/s00421-010-1529-0.
11. Niehoff A, Müller M, Brüggemann L, et al. Deformational behaviour of knee cartilage and changes in serum cartilage oligomeric matrix protein (COMP) after running and drop landing. Osteoarthritis Cartilage. 2011;19(8):1003-10. doi:10.1016/j.joca.2011.04.012.
12. Firner S, Willwacher S, de Marées M, et al. Effect of increased mechanical knee joint loading during running on the serum concentration of cartilage oligomeric matrix protein (COMP). J Orthop Res. 2018;36(7):1937-1946. doi:10.1002/jor.23859.
13. Dong X, Li C, Liu J, et al. The effect of running on knee joint cartilage: A systematic review and meta-analysis. Phys Ther Sport. 2021;47:147-155. doi:10.1016/j.ptsp.2020.11.030.
14. Herger S, Vach W, Liphardt AM, et al. Experimental-analytical approach to assessing mechanosensitive cartilage blood marker kinetics in healthy adults: dose-response relationship and interrelationship of nine candidate markers. F1000Res. 2021;10:490. doi:10.12688/f1000research.52159.2.
15. Bjerre-Bastos JJ, Nielsen HB, Andersen JR, et al. A biomarker perspective on the acute effect of exercise with and without impact on joint tissue turnover: an exploratory randomized cross-over study. Eur J Appl Physiol. 2021;121(10):2799-2809. doi:10.1007/s00421-021-04751-z.
16. Zhang Y, Huang Y, Zhang L, et al. Effects of long-term running on the structure and biochemical composition of knee cartilage in males: a cross-sectional study. Quant Imaging Med Surg. 2024;14(8):6036-6047. doi:10.21037/qims-23-1563.
17. 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.
18. Rethlefsen ML, Kirtley S, Waffenschmidt S, et al. PRISMA-S: an extension to the PRISMA Statement for Reporting Literature Searches in Systematic Reviews. Syst Rev. 2021;10(1):39. doi:10.1186/s13643-020-01542-z.
19. Higgins JPT, Thomas J, Chandler J, et al. Cochrane Handbook for Systematic Reviews of Interventions. Cochrane. 2024;Version 6.5. Available from: https://training.cochrane.org/handbook/current.
20. Sterne JA, Hernán MA, Reeves BC, et al. ROBINS-I: a tool for assessing risk of bias in non-randomised studies of interventions. BMJ. 2016;355:i4919. doi:10.1136/bmj.i4919.
21. Denning WM, Woodland S, Winward JG, et al. The influence of experimental anterior knee pain during running on electromyography and articular cartilage metabolism. Osteoarthritis Cartilage. 2014;22(8):1111-9. doi:10.1016/j.joca.2014.05.006.
22. Denning WM, Becker Pardo M, Winward JG, et al. Ambulation speed and corresponding mechanics are associated with changes in serum cartilage oligomeric matrix protein. Gait Posture. 2016;44:131-6. doi:10.1016/j.gaitpost.2015.11.007.
23. Neidhart M, Müller-Ladner U, Frey W, et al. Increased serum levels of non-collagenous matrix proteins (cartilage oligomeric matrix protein and melanoma inhibitory activity) in marathon runners. Osteoarthritis Cartilage. 2000;8(3):222-9. doi:10.1053/joca.1999.0293.
24. de Vasconcelos BM, Peeler JD, Scribbans T, et al. A preliminary study on the effect of loaded and unloaded exercise on N-propeptide of type II collagen and serum cartilage oligomeric matrix protein activity of articular cartilage in healthy young adults. Appl Physiol Nutr Metab. 2023;48(12):954-961. doi:10.1139/apnm-2023-0124.
25. Mündermann A, Nüesch C, Herger S, et al. Load-induced blood marker kinetics in patients with medial knee compartment osteoarthritis are associated with accumulated load and patient reported outcome measures. F1000Res. 2023;12:299. doi:10.12688/f1000research.131702.2.
26. Selistre LFA, Gonçalves GH, Vasilceac FA, et al. The relationship between urinary C-Telopeptide fragments of type II collagen, knee joint load, pain, and physical function in individuals with medial knee osteoarthritis. Braz J Phys Ther. 2021;25(1):62-69. doi:10.1016/j.bjpt.2020.02.002.
27. Udomsinprasert W, Mookkhan N, Tabtimnark T, et al. Cartilage oligomeric matrix protein as a potential biomarker for knee osteoarthritis. Bone Joint Res. 2024;13(6):261-271. doi:10.1302/2046-3758.136.BJR-2023-0180.R1.
28. Okada S, Taniguchi M, Yagi M, et al. Characteristics of Acute Cartilage Response After Mechanical Loading in Patients with Early-Mild Knee Osteoarthritis. Ann Biomed Eng. 2024;52(5):1326-1334. doi:10.1007/s10439-024-03456-6.
29. Erhart-Hledik JC, Chehab EF, Asay JL, et al. Longitudinal changes in tibial and femoral cartilage thickness are associated with baseline ambulatory kinetics and cartilage oligomeric matrix protein (COMP) measures in an asymptomatic aging population. Osteoarthritis Cartilage. 2021;29(5):687-696. doi:10.1016/j.joca.2021.02.006.
30. Liphardt AM, Godonou ET, Dreiner M, et al. Immobilization by 21 days of bed rest results in type II collagen degradation in healthy individuals. Osteoarthritis Cartilage. 2024;32(2):177-186. doi:10.1016/j.joca.2023.11.007.
31. Ivanochko NK, Gatti AA, Stratford PW, et al. Interactions of cumulative load with biomarkers of cartilage turnover predict knee cartilage change over 2 years: data from the osteoarthritis initiative. Clin Rheumatol. 2024;43(7):2317-2327. doi:10.1007/s10067-024-07014-2.
32. Della Rosa T, Gaulin B, Schwach M, et al. Evaluation of the impact of ultra-trail running on knee cartilage using magnetic resonance imaging t2 mapping. J Sports Med Phys Fitness. 2024;64(12):1321-1328. doi:10.23736/S0022-4707.24.15966-X.
33. Jandacka D, Casula V, Hamill J, et al. Regular Running Is Related to the Knee Joint Cartilage Structure in Healthy Adults. Med Sci Sports Exerc. 2024;56(6):1026-1035. doi:10.1249/MSS.0000000000003386.
34. Hutcherson CW, Mao M, Thakur B, et al. Low-Grade Inflammatory Mediators and Metalloproteinases Yield Synchronous and Delayed Responses to Mechanical Joint Loading. Cartilage. 2024;15(4):417-427. doi:10.1177/19476035231193089.
35. Harkey MS, Blackburn JT, Hackney AC, et al. Sex-Specific Associations between Cartilage Structure and Metabolism at Rest and Acutely Following Walking and Drop-Landing. Cartilage. 2021;13(1_suppl):1772S-1781S. doi:10.1177/1947603520959386.
36. Armitano-Lago C, Evans-Pickett A, Davis-Wilson H, et al. Modifying loading during gait leads to biochemical changes in serum cartilage oligomeric matrix protein concentrations in a subgroup of individuals with anterior cruciate ligament reconstruction. Clin Rheumatol. 2024;43(4):1363-1373. doi:10.1007/s10067-024-06898-4.
37. Bjerre-Bastos JJ, Nielsen HB, Andersen JR, et al. Does moderate intensity impact exercise and non-impact exercise induce acute changes in collagen biochemical markers related to osteoarthritis? - An exploratory randomized cross-over trial. Osteoarthritis Cartilage. 2021;29(7):986-994. doi:10.1016/j.joca.2021.02.569.
38. Boxer B, Zhang Z, Folland JP, et al. Acute effect of impact and resistance exercise on Wnt signaling modulators, bone and cartilage metabolism. J Bone Miner Res. 2026;41(3):282-292. doi:10.1093/jbmr/zjaf128.
39. Erhart-Hledik JC, Titchenal MR, Migliore E, et al. Cartilage oligomeric matrix protein responses to a mechanical stimulus associate with ambulatory loading in individuals with anterior cruciate ligament reconstruction. J Orthop Res. 2022;40(4):791-798. doi:10.1002/jor.25121.
40. Lisee C, Evans-Pickett A, Davis-Wilson H, et al. Delayed cartilage oligomeric matrix protein response to loading is associated with femoral cartilage composition post-ACLR. Eur J Appl Physiol. 2023;123(11):2525-2535. doi:10.1007/s00421-023-05253-w.
41. Herger S, Nüesch C, Liphardt AM, et al. Effect of older age and/or ACL injury on the dose-response relationship between ambulatory load magnitude and immediate load-induced change in serum cartilage oligomeric matrix protein. J Sport Health Sci. 2025;14:100993. doi:10.1016/j.jshs.2024.100993.
42. Davis-Wilson HC, Thoma LM, Johnston CD, et al. Fewer daily steps are associated with greater cartilage oligomeric matrix protein response to loading post-ACL reconstruction. J Orthop Res. 2022;40(10):2248-2257. doi:10.1002/jor.25268.
43. Evans-Pickett A, Davis-Wilson HC, Johnston CD, et al. Immediate Effects of Walking With a Knee Brace After Anterior Cruciate Ligament Reconstruction: A Biomechanical, Biochemical, and Structural Approach. J Athl Train. 2023;58(6):542-553. doi:10.4085/1062-6050-0700.20.
44. Papaneophytou C, Alabajos-Cea A, Viosca-Herrero E, et al. Associations between serum biomarkers of cartilage metabolism and serum hyaluronic acid, with risk factors, pain categories, and disease severity in knee osteoarthritis: a pilot study. BMC Musculoskelet Disord. 2022;23(1):195. doi:10.1186/s12891-022-05133-y.
45. Jørgensen AEM, Schjerling P, DellaValle B, et al. Acute loading has minor influence on human articular cartilage gene expression and glycosaminoglycan composition in late-stage

knee osteoarthritis: a randomised controlled trial. Osteoarthritis Cartilage. 2023;31(7):884-893. doi:10.1016/j.joca.2023.01.317.
46. Bjerre-Bastos JJ, Sejersen C, Bihlet AR, et al. An Estimate of Plasma Volume Changes Following Moderate-High Intensity Running and Cycling Exercise and Adrenaline Infusion. Front Physiol. 2022;13:948087. doi:10.3389/fphys.2022.948087.
47. Chow TT, Uddin N, McManus C, et al. Biological, biomechanical, and pain sensitivity effects of walk-run in people with self-reported knee osteoarthritis. Front Sports Act Living. 2026;8:1792156. doi:10.3389/fspor.2026.1792156.
48. Thudium CS, Engstrøm A, Bay-Jensen AC, et al. Cartilage tissue turnover increases with high- compared to low-intensity resistance training in patients with knee OA. Arthritis Res Ther. 2023;25(1):22. doi:10.1186/s13075-023-03000-2.
49. Bjerre-Bastos JJ, Sejersen C, Nielsen HB, et al. The Impact of Weight-bearing Exercise, Non-Weight-bearing Exercise, and Cardiovascular Stress on Biochemical Markers of Cartilage Turnover in Patients With Mild to Moderate Knee Osteoarthritis - A Sequential, Cross-Over, Clinical Study. Cartilage. 2025;16(2):159-168. doi:10.1177/19476035241258170.
50. Song L, Meng Y. Effect of aerobic exercise of different intensity on articular cartilage metabolism in patients with knee osteoarthritis: A randomized controlled trial. Medicine (Baltimore). 2026;105(22):e48913. doi:10.1097/MD.0000000000048913.
51. Oğuz R, Belviranlı M, Okudan N. Effects of Exercise Training Alone and in Combination With Kinesio Taping on Pain, Functionality, and Biomarkers Related to the Cartilage Metabolism in Knee Osteoarthritis. Cartilage. 2021;13(1_suppl):1791S-1800S. doi:10.1177/19476035211007895.
52. Vassão PG, de Souza ACF, da Silveira Campos RM, et al. Effects of photobiomodulation and a physical exercise program on the expression of inflammatory and cartilage degradation biomarkers and functional capacity in women with knee osteoarthritis: a randomized blinded study. Adv Rheumatol. 2021;61(1):62. doi:10.1186/s42358-021-00220-5.
53. Erhart-Hledik JC, Mahtani GB, Asay JL, et al. Changes in knee adduction moment wearing a variable-stiffness shoe correlate with changes in pain and mechanically stimulated cartilage oligomeric matrix levels. J Orthop Res. 2021;39(3):619-627. doi:10.1002/jor.24770.