Journal of Pain Research
Dovepress

Open Access Full Text Article
REVIEW

Low Back Pain During and After Spaceflight: A Systematic Review with Meta-Analysis

Guillermo Ceniza-Bordallo[1,*], Eric Zimmermann[2,*], Marie Vigouroux[2], Kacper Niburski[3], Maryse Fortin[4], Jean Ouellet[5–7], Juan Pablo Cata[8,9], Pablo M Ingelmo[2,7,10,11]

[1]Department of Radiology, Rehabilitation and Physiotherapy, Faculty of Nursing, Physiotherapy, and Podiatry, University Complutense of Madrid, Madrid, Spain; [2]Edwards Family Interdisciplinary Center for Complex Pain, Montreal Children's Hospital, Montreal, QC, Canada; [3]Department of Anesthesia, University of British Columbia, Vancouver, BC, Canada; [4]Department of Health, Kinesiology & Applied Physiology, Concordia University, Montreal, QC, Canada; [5]Department of Paediatric Surgery & Surgery, McGill University Health Centre, Montreal, QC, Canada; [6]Department of Scoliosis & Spine Surgery, Shriner's Hospital for Children, Montreal, QC, Canada; [7]Alan Edwards Centre for Research on Pain, McGill University, Montreal, QC, Canada; [8]Department of Anesthesia and Perioperative Medicine, the University of Texas – MD Anderson Cancer Center, Houston, TX, USA; [9]Anesthesiology and Surgical Oncology Research Group, Houston, TX, USA; [10]Department of Anesthesia. McGill University, Montreal, Canada; [11]Research Institute, McGill University Health Center, Montreal, Canada

*These authors contributed equally to this work

Correspondence: Guillermo Ceniza-Bordallo, Email gceniza@ucm.es

Abstract: Space flights can produce physiological changes in the spine, leading to the development of acute and chronic pain in passengers. However, there is a lack of comprehensive literature exploring physiological spine changes and acute and chronic pain in space passengers (astronauts and animals). The first aim of this study was to identify the physiological changes experienced by passengers (humans and animals) after space flight. The second aim was to identify the incidence of low back pain during and after space flight. This systematic review was conducted following PRISMA guidelines and was pre-registered in PROSPERO (ID 451144). We included Randomized Controlled Trials or longitudinal studies in humans and animals, and the variables must be assessed either in-flight or post-flight. We conducted a literature search in major databases combining the keywords: Pain; Space; Low Back Pain; Astronauts; Spine Changes; Microgravity; Physiological Changes; Humans; Animals. Risk of bias and quality of studies were analyzed, and the level of evidence was assessed using the GRADE system. After duplicates were removed, 115 abstracts were screened by two reviewers, and finally, 11 articles were included in this review. The evidence indicates that astronauts experience muscle atrophy in the lumbar multifidus with a moderate to large effect, especially in the L4-L5 and L5-S1 segments. Space flights also decrease the range of motion with a moderate effect, along with disc herniations and disc dehydration. 77% of astronauts experience pain during spaceflight, and 47% develop acute pain after spaceflight. Chronic pain was reported by 33% of the astronauts. After space flights, astronauts suffer from lumbar muscle atrophy, reduced range of motion, disc herniations, and disc dehydration, with a high incidence of both acute and chronic pain.

Plain Language Summary: Space travel affects the spine and can cause both immediate and long-term pain. Our study aimed to understand what changes happen in the spine during and after space travel and how often these changes lead to low back pain.

Why was the study done?

We wanted to investigate the specific spinal changes and the frequency of low back pain in both astronauts and animals exposed to space travel. This is important because understanding these effects can help improve health interventions for space travelers.

What did the researchers do?

We reviewed studies that examined spine changes and pain in humans and animals during and after space flights. We used a systematic approach to find relevant research, following strict guidelines and assessing the quality of each study.

What did the researchers find?

- Astronauts often experience muscle loss in the lower back, particularly in the lumbar spine.
- Space travel reduces the spine's range of motion.
- There is a significant occurrence of disc herniations and disc dehydration.
- 77% of astronauts report pain during space flights.
- 47% experience acute pain after returning to Earth.
- 33% suffer from chronic pain post-mission.

Received: 28 August 2024
Accepted: 22 November 2024
Published: 6 December 2024

What do these results mean?

Our findings show that space travel can lead to serious spinal issues and a high risk of pain. These results highlight the need for better health strategies to protect astronauts during and after their missions. Understanding these effects is crucial for developing effective interventions and ensuring the well-being of space travelers.

Keywords: chronic low back pain, space flight, physiological changes, microgravity

Introduction

Space flight's implications for the human body are beginning to be understood. The space flights environment disrupts the homeostatic balance of many physiological systems, which are adapted to Earth.[1–3] Exposure to microgravity leads to a reduction in disc compressive loading, loss of spinal curvature, and lengthening of the vertebral column.[2,4,5]

The risk of post-space flight disc herniation in the lumbar spine is three times higher compared to those not exposed to microgravity.[6] The microgravity environment decreases physical demand on the body.[1,3,7,8] In 21 days, this results in a reduction in muscle mass of up to 40% and increased bone resorption markers in as few as 10 to 14 days.[1,3] Additionally, microgravity induces symptoms akin to aging, such as decreased cardiovascular capacity and immune dysfunction.[7,8] Preliminary data indicate that spinal stiffness persist for between 3 months to 1 year after flight.[6,9]

Astronauts experience episodes of low back pain (LBP) during space flight with over half reporting moderate to severe pain intensity.[10] This pain has been found to impact the quality of sleep, concentration, and the psychological and emotional state of astronauts, as highlighted in a previous review.[5,11] Moreover, after returning to Earth, nearly 70% of astronauts LBP within the first 3 to 10 days.[10,12] Additionally, 40% continue suffering chronic LBP up to one year after the flight[1] significantly affects the activities of daily living in post-space flight passengers, limiting their ability to perform basic tasks such as walking, bending, and lifting, which can hinder their post-flight rehabilitation and return to normal activities.[13]

Pain during space missions can hinder astronaut performance and jeopardize mission success.[14] Therefore, understanding why it occurs and how to control pain during flight has been identified as a priority for space agencies.[14]

Despite advances, research in this area faces significant challenges due to the complexity of the space environment. There is a critical need to understand better the physiological changes that occur in the low back of astronauts during and after flights, and their potential relationship with the onset of pain. The first aim of this systematic review was to identify the physiological changes in the low back experienced by passengers (humans and animals) during and after space flight. Secondly, we aimed to identify the incidence of LBP during and after space flight in astronauts.

Materials and Methods

This systematic review was conducted following The Preferred Reporting Items for Systematic Reviews and Meta-Analysis (PRISMA) guidelines.[15,16] The review protocol was pre-registered in PROSPERO database (ID 451144).

Search Strategy

Two researchers (E.Z & G.C.B) searched in MEDLine, EMBASE, Global Health, PubMed, Web Of Science, Scopus, Cochrane Library articles published between January 1st of 2010 January 31th of 2024. The last search in these databases was carried out on February 1st, 2024. The search strategy used the combined keywords: *Pain; Space; Low Back Pain; Astronauts; Spine changes; Microgravity; Physiological Changes; Humans; Animals*.

Selection Criteria

Type of Article

We include Clinical trials and longitudinal cohort studies (retrospective and prospective). We excluded narrative reviews, conference abstracts, opinion reports.

Type of Population

We include samples of humans or animals who were exposed to space flight, without restriction on the duration of exposure. The population includes astronauts, cosmonauts, participants in manned space missions, and animals used in space experiments. There were no restrictions on gender, age, or country of the human subjects. For animal studies, different species were included as long as the results are relevant and applicable to human health. We excluded studies that analyze only other environments of microgravity, such as bed rest studies.

Types of Outcomes Assessed

We included articles that evaluated physiological changes in the lower back, such as muscle atrophy, water content in spinal structures, changes in ligament density, structural alterations in the spine such as kyphosis or hernias, and other changes like disc degeneration and decreased bone mass. We also included studies that reported pain evaluation in passengers, using pain assessment methods such as pain scales, pain questionnaires, and clinical evaluations.

We excluded articles that did not distinguish between physiological changes in the lower back and those in other parts of the body.

Study Selection

Two reviewers (G.C.B and E.Z.) initially screened titles and abstracts from the search results, categorizing each as "excluded" or "potentially eligible" for our systematic review using Rayan software.[17] Throughout this process, interrater reliability was consistently monitored by having both reviewers independently assess the same ten randomly selected sets of 115 abstracts. Interrater reliability was measured using Cohen's Kappa.[18,19] Discrepancies between the two reviewers were resolved through discussion, with mediation of the senior author (P.I). Subsequently, the same two reviewers independently assessed the eligibility of 20 full-text articles randomly selected from the potentially eligible studies for the systematic review. Interrater reliability at this stage was again measured using Cohen's Kappa.[18,19] Reviewers were not blinded to authors' names, institutions, or journal titles. The remaining full-text articles were assessed by one of the reviewers (G.C.B. or E.Z). Any study that did not clearly meet the eligibility criteria was discussed and mediated by the senior author (P.I.) if needed.

Data Extraction

Two reviewers (G.C.B. and E.Z.) independently extracted data from all the articles included. Then the information extracted were checked by a second reviewer (P.I). The following data were extracted into a excel document: 1) Article characteristics: a) Country of the article; b) Type of study; c) Sample characteristics d) Flight duration, 2) Studies characteristics and spaceflight assessment, 3) Low back physiological changes, 4) Acute and chronic pain reports.

Methodological Quality Assessment

Risk of Bias

The risk of bias in each included study was assessed according to systematic review guidelines.[15,16] Risk of bias was evaluated using each specific tool by two independent reviewers (E.Z. and G.C.B). Disagreements or discrepancies were resolved by the senior author (P.I). Cohen's Kappa[18,19] was calculated for both evaluations, considering the number of identical ratings between the two reviewers and the number of different ratings between the two reviewers in each domain of risk of bias and study quality.

For animal studies, the widely employed scale Systematic Review Centre for Laboratory Animal Experimentation (SYRCLE) risk of bias tool was used.[20] The SYRCLES[20] assesses risk of bias in ten items, taking into account allocation sequence, distribution of participants' baseline characteristics, randomized allocation and outcomes evaluation, blinded evaluation, and problems in study design. Each item is evaluated with three options: "yes", "no", and "unclear".

For human studies, the Quality of Prognostics Studies Tool (QUIPS Tool),[21] which assesses risk of bias in prognostic observational studies, was used. The outcomes were: 1) Spine impairments 2) LBP and the prognostics factors all covariables described before. The QUIPS Tool[21] assesses the risk of bias in observational studies using six points: 1) *Study participation*, 2) *Study attrition* 3) *Prognostic factor measurement* 4) *Outcome measurement* 5) *Study*

Confounding 6) *Statistical Analysis and Reporting*. For each point, it evaluates with three possibilities, "low", "moderate" or "high" risk of bias introduced into potential prognostic factor and outcome.

Quality of Studies

Study quality was evaluated using each specific tool by two independent reviewers (E.Z. and G.C.B). Disagreements or discrepancies were resolved by the senior author (P.I). Cohen's Kappa[18,19] was calculated in both evaluations, considering the number of equal ratings between the two reviewers and the number of different ratings between the two reviewers in each domain of quality of studies.

For animal studies, the Collaborative Approach to Meta-Analysis and Review of Animal Data from Experimental Studies Tool (CAMARADES tool),[22] a scale used for these types of articles was used. The CAMARADES tool assesses the quality of studies according to ten items. Among these items, we excluded items 3, 6, and 7 because they did not apply to our included studies. Each item is sorted into three categories: "yes", "no", and "unclear".

For human studies, the National Institute of Health (NIH) Quality Assessment Tool[23] is widely used to assess study quality as it allows for different study designs to be assessed.[23,24] The NIH Quality Assessment Tool[23] assesses quality of studies taking into account problems in study design, and focusing on internal validity. Each item is evaluated with three possibilities: "yes" "no" or "other" (CD, *cannot determine*; NA, *not applicable*; NR, *not reported*)".

Data Synthesis

The quality of evidence was assessed using the Grading of Recommendations, Assessment, Development and Evaluations (GRADE) approach.[25,26] For each domain, the following were analyzed: (1) *phase of the research*; (2) *limitations of the study* (3) *inconsistency of the results*, (4) *indirectness (not generalizable)*, (5) *imprecision (insufficient data)* and (6) *publication bias*, (7) *effect size* and (8) *dose effect*. We used the guidelines provided by Huguet et al,[26] the evidence was classified as *high* (++++), *moderate* (+++), *low* (++), and *very low* (+).

The assessment of all GRADE[25,26] factors was pilot-tested by two reviewers (G.C.B. and E.Z.) on 30% of the outcomes included in this review. Since the level of agreement was deemed to be adequate, one reviewer (E.Z.) assessed the evidence for the remaining outcomes, except for "Study limitations". Any uncertainties were discussed by another reviewer (G.C.B). The "Study limitations" factor was independently assessed by two reviewers (G.C.B. and E.Z. or P.I.) using the QUIPS.[21] Interrater reliability for the QUIPS Tool[21] study limitations rating was evaluated using Cohen's Kappa.[18,19]

Quantitative Analysis

A meta-analysis was conducted to analyze the incidence of pain during and after spaceflight using statistical software *Jamovi version 2.3*.[27] There is not a minimum number of articles required to perform a meta-analysis; therefore, we included at least two articles to examine associations.[28,29] Sample sizes and the number of astronauts experiencing pain were selected and extracted from studies included in the analysis. Due to expected heterogeneity among studies, a random-effects model was applied.[30,31] This approach is appropriate for handling variability among studies and provides more accurate estimates when the approximate normal within-study likelihoods are replaced with the appropriate exact likelihoods, leading to a generalized linear mixed model.[32] This is especially useful in the context of sparse and heterogeneous data, ensuring that differences among studies are adequately addressed.[32]

The presence of heterogeneity among studies was assessed by Cochran's Q statistic.[33] Additionally, the I^2 index was calculated as a measure of the proportion of total variability across studies due to heterogeneity rather than chance.[34] I^2 values of 75% or higher indicate a high degree of heterogeneity, reflecting significant variability among the results of included studies.[34] Because we expected a low number of studies, we also incorporated Tau and tau squared.[35] Tau and tau squared were calculated as estimates of heterogeneity among studies, providing a measure of variability among studies beyond what is expected by chance.[35,36] Higher values of tau and tau squared indicate greater heterogeneity among studies, suggesting that observed differences among studies are not solely due to chance but to real variations in the effects of the included studies.[35,36]

https://doi.org/10.2147/JPR.S491060
DovePress

Publication bias was evaluated using funnel plots and specific statistical tests. The Egger test was used to directly assess funnel plot asymmetry.[37] A significant regression slope suggests potential asymmetry in the funnel plot, indicating publication bias.[37]

To evaluate the effect of microgravity exposure in space on the lumbar spine, we analyzed the observed changes in variables before and after spaceflight. The meta-analysis was conducted using the standardized mean difference (SMD). SMDs of 0.2, 0.5, and 0.8 are considered small, medium, and large, respectively.[38] A random-effects model was fitted to the data, and the amount of heterogeneity.[30,31] There is not a minimum number of articles required to perform a meta-analysis; therefore, we included at least two articles to examine associations.[28,29] Tau-squared was estimated using the restricted maximum-likelihood method.[39] In addition, Cochran's Q test for heterogeneity and the I^2 statistic were reported.[33,34] If any heterogeneity was detected, a confidence interval for the true outcomes was also provided. Studentized residuals and Cook's distances were used to examine whether any studies were outliers or influential within the model's context.[30] Studies with studentized residuals larger than the 97.5th percentile of a standard normal distribution, adjusted for Bonferroni correction, were considered potential outliers.[30] Studies with Cook's distances larger than the median plus six times the interquartile range of Cook's distances were considered influential.[40] Rank correlation tests and regression tests, using the standard error of the observed outcomes as a predictor, were employed to check for funnel plot asymmetry.[37]

Results

Results of the Search

Two researchers independently searched the databases and identified 115 records after removing duplicates (Figure 1). After analyzing the articles by "Title" and "Abstract", 20 records were selected. After applying the inclusion and exclusion criteria, in the "Full text" analysis, finally, 11 articles[1,13,14,41–48] were included in this review. Interrater reliability for screening the titles and the abstracts retrieved was substantial (Kappa = 0.85). Interrater reliability for evaluating the eligibility of the 20 randomly selected full-text articles was appropriate (Kappa = 0.70).

Characteristics of the Articles Included

Of the 11 studies included, 6 (54%)[1,13,14,41–43] were prospective observational cohort studies and 4[45–48] also had an intervention (post-flight rehabilitation) and one[44] include an ultrasound protocol. The studies were published between 2014 and 2023. They were predominantly conducted in the United States (64% n= 7 studies),[1,13,41–44,48] two in Australia,[45,46] and two in Germany[14,43] (Table 1).

Characteristics of the Participants

Of the 11 studies included in the review, 10 analyzed human samples (n=93)[1,13,14,41,43–48] and one study analyzed mouse samples (n=24).[42] The human samples[1,13,14,41,43–48] had a simple sizes ranging from 6 to 20 participants with a mean age ranging from 38 to 55 years. Two studies analyze same samples.[1,43] The animal study[42] included a 24 mice divides into two groups, spaceflight and control with the same environmental characteristics except for the exposure to microgravity in the spaceflight group (Table 1).

Characteristics of Spaceflights

All studies included (n=11)[1,13,14,41–48] assessed participants before and after the spaceflight, while four studies[13,14,44,46] also evaluated participants during the flight.

The human spaceflights lasted between 14 days and 6 months. The post-spaceflight assessments varied significantly among the human studies. Three studies[1,43,44] conducted only one post-flight assessment. These single evaluations also differed, being conducted immediately after the flight[44] or within the first three months,[1] or at six months.[43] Five studies[13,14,41,45,46] conducted multiple post-flight assessments, with only one study including a long-term follow-up at 12 months.[41] Four studies additionally performed rehabilitation interventions after spaceflight exposure (n=4 studies),[45–48] with only one study[48] including a long-term follow-up at 1, 2, and 4 years after the flight (Tables 1 and 2).

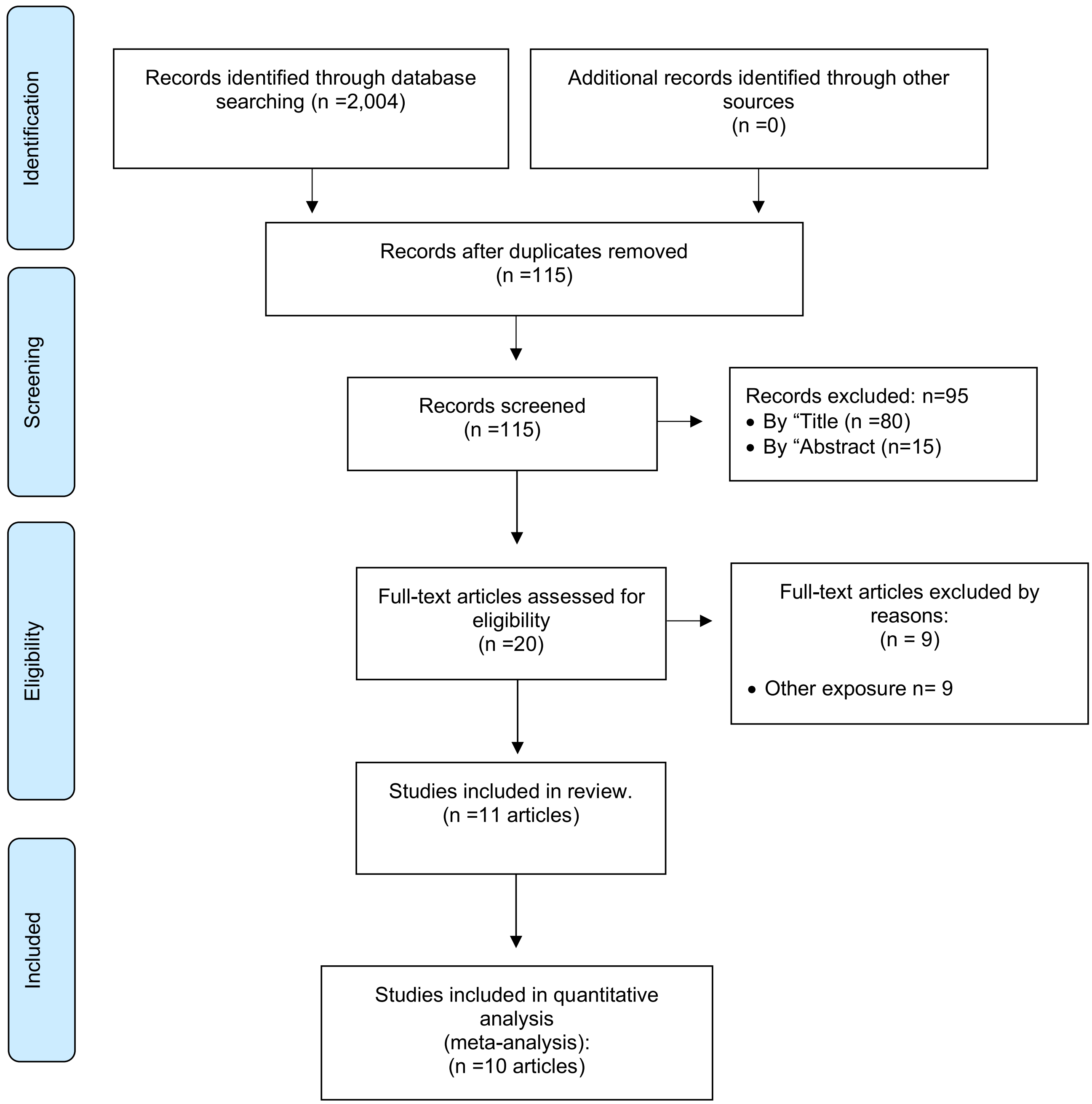

Figure 1 PRISMA Flow Diagram.

The animal spaceflights study lasted 15 days,[42] with post-spaceflight assessments conducted 48 hours after landing. In this study, sixteen C57BL/C mice (space flight group, n=8; ground-based control group, n=8) were sacrificed immediately after the spaceflight. The assessments focused on the biomechanical properties of the lumbar and caudal discs, specifically measuring physiological disc height and conducting compressive creep tests to evaluate parameters such as endplate permeability, nuclear swelling pressure strain dependence, and annular viscoelasticity.

Physiological Spine Changes in Low Back After Space Flights

The physiological impact of spaceflight on the human body leads to significant changes in spinal health. Astronauts experience notable alterations in spinal structures after spaceflight, affecting both muscle characteristics and lumbar

Table 1 Characteristics of Studies and Participants Included

Study	Country	Study type	Type of sample	Study size (% female)	Mean age (SD)/ Range	Space flight duration
Bailey et al 2014[42]	U.S.	Prospertive cohort study	Mice	16 (100%)	16 weeks	15 days
Pool-Goudzwaard et al 2015[14]	Germany, Netherlands	Prospertive cohort study	Humans	60 (83%)	47 years (6 years)	15 days
Chang et al 2016[1]	U.S.	Prospertive cohort study	Humans	6 (16%)	46–55 years	117–213 days
Hides et al 2016[45]	Australia	Prospertive cohort Study + Intervention	Humans	1*	38 years	14 days
Bailey et al 2018[43]	Germany	Prospertive cohort study	Humans	6 (16%)	46–55 years	6 months
Garcia et al 2018[44]	U.S.	Prospertive cohort Study + US protocol development	Humans	7 (14%)	46.1 years (6.4 years)	150 days
Burkhart et al, 2019[48]	U.S.	Prospective cohort study + Intervention	Humans	17 (NA)	45 years (4 years)	5.9 months (SD=1 month)
Hides et al 2021[46]	Australia	Case report + Intervention	Humans	5 (20%)	NA	6 months
Bailey et al, 2022[41]	U.S.	Prospertive cohort study	Humans	12 (16%)	51.3 years (5.6 years)	6 months (168 days, 138 to 289 days)
Coulombe et al 2023[47]	U.S.	Prospertive cohort study + Intervention	Humans	17 (NA)	45 years (4 years)	5.8 months (two astronauts were in space for 4 months, three astronauts for 5 months, and seven astronauts for 6 months, and five astronauts for 7 months).
Sauer et al, 2023[13]	U.S.	Prospertive cohort Study	Humans	2 (NA)	NA	17 days

Note: *Male.
Abbreviations: US, Ultrasound; NA, Not Available.

Table 2 Objectives by Studies and Spaceflight Assessment

Study	Objective	Outcomes evaluation	Statistical Analysis
Bailey et al, 2014[42]	To assess microgravity's relative effects on lumbar and caudal intervertebral discs, they quantized changes in physiological disc height and biomechanical properties of tissues from mice that had returned from 15-day NASA shuttle mission, as compared to those of ground-based controls. Since tail discs have a significantly greater range of motion as compared to lumbar discs, they hypothesized that microgravity would have a more pronounced detrimental effect among the caudal discs.	The personnel sacrificed mice three hours after landing and control mice were sacrificed 48 h later.	Creep parameters (initial modulus, compressive strain, and from the fluid transport model) and physiological disc height of each tested disc were analyzed using a non-parametric statistical test (Wilcoxon–Mann–Whitney), comparing space eight and control groups for each parameter. Statistical significance was based on $p < 0.05$
Pool-Goudzwaard et al, 2015[14]	To describe prospectively the development and course of low back pain in microgravity in full detail regarding onset, localization, severity, and relieving countermeasures undertaken by astronauts per day in short-term flight. They compare the development and course of low back pain between astronauts with a history of low back pain prior to flight versus healthy astronauts (no low back pain). They also aim to compare data from space- flight to data on low back pain in two bed rest studies.	The questionnaire was completed by astronauts 10 days prior to the flight, each day during the 12- to 15-days flight, as well as 10 days postflight. At 3 to 6 months postflight, a debriefing was scheduled to determine the effectiveness of the countermeasures for the astronauts who experienced low back pain.	The occurrence, localization, intensity, continuity, and duration of LBP has been described for all subjects and tested for significant difference between astronauts experiencing LBP with no history of LBP on Earth and those who experienced LBP in their life prior to flight using Wilcoxon ranking for the dichotomous variable and ordinal variable and the Student's *t*-test for the continuous variable (NRS, duration). Descriptive data on provoking movements or periods and successful countermeasures undertaken by astronauts will be described
Chang et al, 2016[1]	To understand the factors involved in lumbar spine strength and back pain in crewmembers during a long mission and after increased goads of landing and re-adaptation to Earth.	Supine lumbar spine magnetic resonance imaging scans were conducted pre-flight, immediate post-flight and at least 30 days post-flight recovery after a mission. Pre-flight imaging was performed on average 214 days prior to launch. The immediate post-flight imaging was performed within 1–2 days. The recovery period images were performed an average of 46 days (range 33–67 days) after landing.	One-way, repeated measures ANOVA to establish significance, defined as $p<0.05$, followed by post-hoc testing with the Newman-Keuls multiple comparison test 33 with alpha = 0.05, using GraphPad Prism. Change in the average disc height was calculated at post-flight (Post-Preflight), recovery (Recovery-Postflight), and overall change from pre-flight to recovery (Recovery-Preflight)

Hides et al, 2016[45]	To assess the size of the lumbar multifidus and antero-lateral abdominal muscles using ultrasound imaging pre- and post-space flight, to determine the effects of microgravity (and exercise in microgravity) on these muscles. To assess the response of the multifidus and antero-lateral abdominal muscles to post-space flight rehabilitation	Pre-flight and at Day 1, Day 8 and Day 14, respectively, following return to Earth after a period of six months in microgravity on the International Space Station	NA
Bailey et al, 2018[43]	To identify back pain and injury mechanisms, they conducted a longitudinal study of six NASA astronaut crewmembers in whom lumbar spine anatomy and biomechanics were quantified before and after six months of microgravity exposure on the International Space Station.	Two time-points: before launch ("pre") and one day following six months spaceflight on the International Space Station ("post").	All pre-flight variables were tested for normal distribution using a Shapiro–Wilk test. Statistical analyses included paired t-tests to compare changes in pre- and post-flight variables among subjects, and simple regression analysis to test for relationships between pre- to post-flight changes for separate variables. Significance was defined as $p<0.05$. Statistical analyses were performed with Stata
Garcia et al, 2018[44]	To develop a simplified and reproducible US protocol with a set of common semantics for operational spinal US examinations during long-duration orbital and exploration-class spaceflight.	Preflight and post- flight spinal US images, three in-flight sessions were scheduled for each participant on flight day 30 (615), flight day 90 (615), and flight day 150 (615)	A Student *t* test was performed to identify any differences between the preflight and post- flight anthropometric measurements. A x2 analysis was performed to identify any differences between the occurrences of abnormalities found by US and MRI. A generalized linear model repeated-measures analysis of variance with a Bonferroni post hoc analysis was per- formed to identify differences over the course of the US measurements from before flight to in-flight to after flight. For all statistical analyses, significance was determined as $P < 0.05$.

(*Continued*)

Table 2 (Continued).

Study	Objective	Outcomes evaluation	Statistical Analysis
Burkhart et al, 2019[48]	To determine the effect of long-duration spaceflight and multiyear recovery on Earth on lumbar paraspinal musculature. In addition, they aimed to determine the association between in-flight exercise and trunk musculature, as this information can inform future exercise countermeasure development, both on Earth and in microgravity	All crewmembers underwent pre- and postflight scanning, with preflight scans acquired 30 to 60 days prior to flight and postflight scans acquired 7 to 10 days after landing. Fifteen individuals underwent scanning approximately 1 year after landing, and 8 underwent an additional scan 2 to 4 years after landing.	They used a paired *t* test to examine changes in pre- versus immediate postflight measurements of muscle CSA and attenuation. They computed the rate of change in the muscle CSA and attenuation by dividing the percent difference between post and preflight values by mission duration (in months). The immediate postflight measurements were compared by paired t tests to those made 1 year after landing. In subjects with additional follow-up measurements, another paired *t* test was employed to test whether muscle morphology or spinal measures remain different from preflight during extended recovery time in normal gravity. They used Pearson correlation coefficients to determine the association between muscular declines (in CSA and attenuation) and in-flight exercise.
Hides et al, 2021[46]	To examine the changes in muscle size and function of the lumbar multifidus and anterolateral abdominal muscles across three time periods: (1) Time in-flight on the ISS (Preflight to Return day to determine changes associated with exposure to microgravity), (2) Reconditioning time (day 1 to day 15, with an additional midpoint measure at day 8 to determine changes associated with performing daily reconditioning exercises), and (3) Total time (Preflight to day 15 to determine if muscle values measured after reconditioning returned to preflight values)	Three time periods: (1) Time in-flight (Preflight to Return day to determine changes associated with exposure to microgravity), (2) Reconditioning time (Day 1 to day 15, with an additional midpoint measure at day 8 to determine changes associated with performing daily reconditioning exercises), and (3) 15 days after flight	Multilevel linear models that incorporated random intercepts for participant, astronaut mission number (two astronauts had flown twice) and side for CSA measures were used to estimate the change in muscle CSA of the lumbar MF muscles at four lumbar vertebral levels (L2, L3, L4, and L5), and thickness and contraction of the anterolateral abdominal muscles and the multifidus muscles. These measures were analyzed to summarize the change in size over three time periods: (1) time in flight; (2) reconditioning time (3) and total time (preflight to R+15) Restricted maximum likelihood and an autoregressive correlation structure was used, and all analyses were adjusted for the baseline measure at the beginning of the specific time period The estimates were the beta-coefficients for the time in- flight and total time models: A linear combination of the daily change for 15 days (15 x B) was used to estimate the total change over the reconditioning time.

Bailey et al, 2022[41]	To examine the relationship between prolonged exposure to microgravity and the elevated incidence of postflight disc herniation, they conducted a longitudinal study to track the spinal health of twelve NASA astronauts before and after approximately 6 months in space	Three separate time points in relation to spending 6 months in space: (1) within a year before launch ("preflight"), (2) within a week after return to Earth ("postflight"), and (3) between 1 and 2 months after return to Earth ("recovery"). Astronauts chose to volunteer for the study after hearing a short briefing about the research and the recovery period (30–60 days) was determined by NASA and the availability of the partici- pating astronauts	Assessed for potential differences (based *t* test for age, frequency for sex) in relation to incidence for disc herniation The analysis for this study involved a repeated-measures, mixed methods linear regression model was used to assess changes in continuous quantitative variables from 3T MRI over three time points (preflight, postflight, and recovery). The three time points were treated as factors to flexibly model longitudinal trends before and after the post- flight measurement. Interactions between spinal level and timepoint were modelled as fixed effects and repeated measurements within subjects and spinal levels were modeled as nested random intercepts. Pairwise comparisons were used to compare between timepoint data. To explore longitudinal effects from disc herniation within spinal levels, secondary repeated-measures, mixed methods linear regression analyses were conducted for models with a significant effect from disc herniation across timepoints. These models were used for specific spinal levels and included main effects and interaction terms for disc herniation and timepoint modelled as fixed effects with repeated measurements within subjects modeled as random intercepts. One sided paired t tests were used to assess changes in spinal segment kinematics because this measure was only collected at pre- and postflight timepoints. Unpaired t tests were used to assess baseline differences in variables between subjects who experienced postflight disc herniation versus those that did no

(*Continued*)

https://doi.org/10.2147/JPR.S491060
DovePress

Table 2 (Continued).

Study	Objective	Outcomes evaluation	Statistical Analysis
Coulombe et al, 2023[47]	To use computed tomography scans of the lumbar spine from long-duration astronauts to determine the effect of space- flight and 12 months of reloading on the spatial heterogeneity of vertebral trabecular bone density	Before flight (n = 17), immediately after flight (n = 17), and 12 months readaptation to gravitational loading on Earth (follow-up n = 15)	Differences between time points were evaluated via paired t-tests. They used a one-sample *t*-test to determine whether the monthly rate of change of these variables was significantly different from 0. Next, they evaluated the associations among global muscle density, and CSA via general linear regression models. They computed the monthly rate of change for each of the three regions for spaceflight and readaptation as described earlier.
Sauer et al, 2023[13]	To evaluate the pain experience and potential sensory changes in astronauts during a short travel commercial space flight.	Four time points: T0, within two weeks before the flight (Pre-Flight); T1, during the flight (In-Flight) except for quantitative sensory testing; T2, within two weeks after the flight (Post-Landing); and T3, three months after the flight (3-month follow-up).	NA

Notes: NA= Not Applicable because Sauer et al 2023, only include 2 participants, and Hides et al 2015 only include 1 participants.
Abbreviations: US, Ultrasound; LPB, Low Back Pain; CSA, Cross-sectional area; MRI, Magnetic Resonance Imaging; NRS, Numerical Rating Scale.

mobility. The studies included in our analysis highlight physiological changes across four domains: 1) spinal pathologies, 2) disc dimensions and lumbar lordosis, 3) muscle atrophy, and 4) range of motion (ROM) (Table 3).

Spinal pathologies were evaluated in three studies.[41,43,44] Specifically, eight spinal pathologies were assessed: lumbar disc herniation, disc desiccation, disc degeneration, osteophytes, intervertebral disc bulge, lumbar endplate irregularities, lumbar facet arthropathy, and adjacent high-intensity zones. Spinal structures were examined before and after the flight using ultrasound[44] and Magnetic Resonance Imaging (MRI).[41,43] One study[44] found one lumbar hernia, three new osteophytes, two new intervertebral disc bulges, and nine new cases of disc desiccation post-flight.

The disc size was evaluated in two studies (one in humans and one in mice)[1,42], lumbar lordosis in one study,[43] and water content in two studies.[41,43] Significant changes in caudal intervertebral disc size before and after the flight were found in mice, but not in the lumbar region.[42] In humans, a decrease in the intervertebral disc size was observed in the L1-L2, L3-L4, L4-L5, and L5-S1 segments.[1] All studies reported changes[41,43] in the water content in the intervertebral disc, but none were statistically significant. Specifically, a decrease in water content was found in the L1-L2, L2-L3, L3-L4, and L1-S1 segments, while an increase was observed in the L4-L5 and L5-S1 segments after spaceflight.

We carried out a meta-analysis to explore the differences in water content of intervertebral discs before and after spaceflights. A total of k=10 samples were included in the analysis. The estimated SMD based on the random-effects model was −0.052 (95% CI: −0.347 to 0.242) with the majority of estimates showing an decrease in water content after spaceflight (60%). Therefore, the average outcome did not differ significantly from zero ($z = -0.350$, $p = 0.725$) (Figure 2). These results suggest that exposure to microgravity during spaceflight does not lead to significant changes in the water content of intervertebral discs. There was no significant heterogeneity in the samples (Q p=0.739; Tau=0.0; I^2=0%). An examination of the studentized residuals revealed that none of the studies had a value larger than ±2.8070, hence there was no indication of outliers in the context of this model. According to the Cook's distances, none of the studies could be considered to be overly influential. Neither the rank correlation nor the regression test indicated any funnel plot asymmetry ($p = 0.600$ and $p = 0.842$, respectively) (Figure 3).

Muscle atrophy was the most studied outcome, with six studies[1,41,43,45,46,48] investigating this aspect. The muscles examined included the multifidus, erector spinae, transversus abdominis, internal oblique, external oblique, psoas, quadratus lumborum, and lumbar paraspinals. Both the cross-sectional area (CSA) and the functional cross-sectional area (FCSA) were calculated. Findings of muscle atrophy were observed in all muscles both before and after the flight.

We carried out a meta-analysis to explore the differences in the CSA of the lumbar multifidus before and after spaceflights. A total of k=8 samples were included in the analysis. The estimated average SMD based on the random-effects model was 0.743 (95% CI: 0.239 to 1.248; $p = 0.003$) indicating a moderate to large effect size[38] (Figure 4). The majority of the estimates were positive (88%), suggesting a consistent trend of increased CSA in the lumbar multifidus post-spaceflight. These findings imply that exposure to microgravity during spaceflights leads to significant hypertrophy of the lumbar multifidus muscle. The Q-test for heterogeneity was not significant, but some heterogeneity may still be present in the true outcomes ($Q = 13.180$, $p = 0.067$, $tau^2 = 0.246$, $I^2 = 47.6\%$). A 95% confidence interval for the true outcomes is given by −0.351 to 1.839. Hence, although the average outcome is estimated to be positive, in some studies the true outcome may in fact be negative. An examination of the studentized residuals revealed that none of the studies had a value larger than ± 2.734 and hence there was no indication of outliers in the context of this model. According to the Cook's distances, none of the studies could be considered to be overly influential. Neither the rank correlation nor the regression test indicated any funnel plot asymmetry ($p = 0.719$ and $p = 0.453$, respectively) (Figure 5).

Additionally, ROM in the lumbar spine was evaluated in two studies.[41,43] Flexion-extension and lateral movements were assessed actively and passively by vertebral segments. In active flexion-extension, reductions in ROM were found at the L2-L3, L3-L4, L4-L5, and L5-S1 levels in both studies.[41,43] One study[41] found reductions at L1-L2, while another[43] found an increase. In passive flexion-extension, the L1-L2 and L5-S1 segments showed reduced ROM, while the rest showed increases, though none were statistically significant.[43] Lastly, active lateral ROM was evaluated in only one study,[41] which found reductions in ROM across all segments after the flight.

Additionally, we carried out a meta-analysis to explore the differences in ROM before and after spaceflights. A total of k=10 samples were included in the analysis. The estimated average SMD based on the random-effects model was 0.490 (95% CI: 0.069 to 0.910, $p = 0.022$), indicating a moderate effect size[38] (Figure 6). The majority of the estimates

Table 3 Spine Physiological Changes in Low Back Before and After Space Flights

1. **Spinal Pathologies**				
Study	**Outcome**	**Before flight n spinal pathologies/ n total of segments**	**After flight n spinal pathologies/ n total of segments**	**Incidence/ number of differences**
Bailey et al, 2018[43]	Lumbar disc herniation (number of patients)	0/6	1/6	16,6%
Garcia et al, 2018[44]	Lumbar disk herniation (number of clinical findings by US)	0	0	0
Garcia et al, 2018[44]	Disk desiccation (number of clinical findings by US)	3	11	+9
Garcia et al, 2018[44]	Disk degeneration (number of clinical findings by US)	0	0	0
Garcia et al, 2018[44]	Osteophytes (number of clinical findings by US)	5	8	+3
Garcia et al, 2018[44]	Intervertebral disk bulge (anterior) (number of clinical findings by US)	4	6	+2
Bailey et al, 2022[41]	Lumbar endplate irregularities (number of clinical findings by MRI)	22/60	22/60	0
Bailey et al, 2022[41]	Lumbar facet arthropathy (number of clinical findings by MRI)	10/60	10/60	0
Bailey et al, 2022[41]	Adjacent high intensity zones (number of clinical findings by MRI)	18/60	18/60	0
2. **Disc Size and Lumbar Lordosis**				
2.1 Disc size				

https://doi.org/10.2147/JPR.S491060
DovePress

Study	Outcome	Spaceflight Mean (SD)	No Spaceflight Mean (SD)	Mean differences	P-value	Effect Size
Bailey et al, 2014 (mouse)[42]	Caudal disc height (mm)	0,28 (0,02)	0,41 (0,02)	−0,32	0,034	r=0,80
	Lumbar disc height (mm)	0,52 (0,01)	0,53 (0,01)	+0,01	0,39	r=0,24
	Caudal compressive strain	0,38 (0,22)	0,11 (0,03)	0,27	0,034	r= −0,80
	Lumbar compressive strain	0,15 (0,001)	0,15 (0,01)	0	0,89	r= −0,04
	Caudal strain-dependent swelling pressure	2,48 (0,95)	7,58 (1.09)	+5,1	0,034	r=0,80
	Lumbar strain-dependent swelling pressure	8,16 (0,64)	8,19 (0,69)	+0,03	0,89	r= −0,04
Chang et al, 2016[1]	L1L2 disc height change (mm)	–	–	−0,1 (1,2)	–	–
Chang et al, 2016[1]	L2L3 disc height change (mm)	–	–	0 (0,4)	–	–
Chang et al, 2016[1]	L3L4 disc height change (mm)	–	–	−0,8 (1,5)	–	–
Chang et al, 2016[1]	L4L5 disc height change (mm)	–	–	−0,3 (0,5)	–	–
Chang et al, 2016[1]	L5S1 disc height change (mm)	–	–	0,1 (1,0)	–	–
2.2 Water content						
Study	**Outcome**	**Before flight Mean (SD)**	**After flight Mean (SD)**	**Mean differences (Percentage of change)**	**P value**	**Effect size**
Bailey et al, 2018[43]	L1L2 water content (mean T2 intensity)	109.7 (49.9)	107.9 (55.7)	1.83 (−1.6%)	0.48	–
Bailey et al, 2022[41]	L1L2 water content (mean T2 intensity)	101.4 (17.8)	98.7 (10.7)	−2,7	-	–

(*Continued*)

Table 3 (Continued).

Bailey et al, 2018[43]	L2L3 water content (mean T2 intensity)	93.2 (38.1)	84.9 (36.2)	−8.3 (−8,9%)	0.3	–
Bailey et al, 2022[41]	L2L3 water content (mean T2 intensity)	81.2 (8.2)	79.0 (7.7)	−2.1	-	–
Bailey et al, 2018[43]	L3L4 water content (mean T2 intensity)	71.6 (37.1)	66.8 (29.1)	−4,8 (−6.7%)	0.19	–
Bailey et al, 2022[41]	L3L4 water content (mean T2 intensity)	73.0 (7.6)	70.5 (8.3)	−2,5	-	–
Bailey et al, 2018[43]	L4L5 water content (mean T2 intensity)	78.0 (33.5)	78.5 (22.8)	0.49 (0.6%)	0.53	–
Bailey et al, 2022[41]	L4L5 water content (mean T2 intensity)	65.7 (6.9)	68.8 (6.2)	+3,1	-	–
Bailey et al, 2018[43]	L5S1 water content (mean T2 intensity)	46.7 (15.6)	51.7 (16.2)	5.03 (10.7%)	0.87	–
Bailey et al, 2022[41]	L5S1 water content (mean T2 intensity)	48.8 (4.7)	52.4 (4.3)	+3,6	–	–
Bailey et al, 2018[43]	L1S1 water content (mean T2 intensity)	399,1 (138,4)	389,7 (127,0)	−9,41 (−2,4%)	0,43	–
2.3 Lumbar lordosis						
Study	**Outcome**	**Before flight Mean (SD)**	**After flight Mean (SD)**	**Mean differences (Percentage of change)**	**P value**	**Effect size**
Bailey et al, 2018[43]	Lumbar lordosis (grades)	41,9 (12,9)	37,2 (11,0)	−4,73 (−11,1%)	0,009	–
3. **Muscle Atrophy**						
Study	**Outcome**	**Before flight Mean (SD)**	**After flight Mean (SD)**	**Mean differences (Percentage of change)**	**P value**	**Effect size**
Multifidus CSA						
Hides et al, 2016[45]	Multifidus CSA L2 (cm^2)	–	–	7	–	–

Hides et al, 2016[45]	Multifidus CSA L3 (cm²)	–	–	7	–	-
Hides et al, 2016[45]	Multifidus CSA L4 (cm²)	–	–	1	–	-
Hides et al, 2016[45]	Multifidus CSA L5 (cm²)	9,86	6,99	−2,87	–	–
Hides et al, 2021[46]	Multifidus CSA L1L2 (cm²)	3,26 (0,72)	3,47 (0,96)	0,21	–	–
Bailey et al, 2022[41]	Multifidus CSA L1L2 (cm²)	2.28 (0.20)	1.97 (0.16)	−0,31	–	–
Hides et al, 2021[46]	Multifidus CSA L2L3 (cm²)	5,22 (1,02)	4,56 (1,07)	−0,66	–	–
Bailey et al, 2022[41]	Multifidus CSA L2L3 (cm²)	3.30 (0.31)	3.27 (0.31)	−0,03	–	–
Hides et al, 2021[46]	Multifidus CSA L3L4 (cm²)	8 (1,12)	7,51 (0,87)	−0,49	–	–
Bailey et al, 2022[41]	Multifidus CSA L3L4 (cm²)	5,04 (0,56)	4,66 (0,39)	−0,38	–	–
Hides et al, 2021[46]	Multifidus CSA L4L5 (cm²)	10,07 (1,41)	9,03 (1,86)	−1,04	–	–
Bailey et al, 2022[41]	Multifidus CSA L4L5 (cm²)	7,32 (0,52)	6,51 (0,37)	−0,81	–	–
Bailey et al, 2022[41]	Multifidus CSA L5S1 (cm²)	7,52 (0,61)	6,78 (0,52)	−0,74	–	–
Bailey et al, 2018[43]	Multifidus CSA (mm²)	1235,7 (252,2)	1158,1 (231,4)	−77,7 (−6,2%)	0,16	–
Burkhart et al, 2019[48]	Multifidus CSA (mm^2)	395.6 (64)	370.2 (61)	−6.1% (5.1%)	–	–
Bailey et al, 2018[43]	Multifidus FCSA (mm²)	1002,5 (319,9)	847,3 (253,1)	−155,2 (−14,2%)	0,06	–

(*Continued*)

https://doi.org/10.2147/JPR.S491060

Table 3 (Continued).

Erector spinae CSA						
Bailey et al, 2018[43]	Erector spinae CSA (mm²)	5010,7 (815,2)	4817,9 (1026,1)	−192,9 (−3,9%)	0,28	–
Bailey et al, 2018[43]	Erector spinae FCSA (mm²)	3903,7 (457,6)	3486,5 (1186,2)	−417,2 (−11,5%)	0,18	–
Burkhart et al, 2019[48]	Erector spinae CSA (mm^2)	2010.1 (209)	1915.7 (240)	−4.6%(7.0%)	–	–
Lumbar CSA						
Chang et al, 2016[1]	Lumbar CSA (mm²)	10,122 (1905)	9769 (2239)	−353	< 0.05	–
Chang et al, 2016[1]	Lumbar paraspinal FCSA (mm²)	8737 (1758)	7049 (1822)	−1688	< 0.05	–
Transversus, Internal oblique, External oblique, Psoas, Quadratus Lumborum						
Hides et al, 2021[46]	Transversus abdominis thickness (cm)	0,44 (0,1)	0,29 (0,12)	−0,15	–	–
Hides et al, 2021[46]	Internal oblique thickness (cm)	1,17 (0,18)	0,99 (0,27)	−0,18	–	–
Hides et al, 2021[46]	External oblique thickness (cm)	0,98 (0,13)	0,89 (0,16)	−0.09	–	–
Burkhart et al, 2019[48]	Psoas CSA (mm^2)	902.6 (234)	853.9 (205)	−4.6% (9.6%)	–	–
Burkhart et al, 2019[48]	Quadratus Lumborum CSA (mm^2)	486.8 (96)	441.3 (82)	−8.4% (10.8%)		
4. **Range of Motion**						
Study	**Outcome**	**Before flight Mean (SD)**	**After flight Mean (SD)**	**Mean Differences (Mean percentage differences)**	**P value**	**Effect size**
Active Flexion-Extension						
Bailey et al, 2018[43]	Active Flexion-Extension ROM – L1L2 (Grades)	6,8 (4,3)	7,1 (4,8)	0,35 (7,7%)	0.73	–

Bailey et al, 2022[41]	Active Flexion-Extension ROM – L1L2 (Grades)	9.1 (0.8)	8.8 (1.3)	−0,3	–	–
Bailey et al, 2018[43]	Active Flexion-Extension ROM - L2L3 (Grades)	8,3 (4,3)	6,9 (4,9)	−1.42 (−22,1%)	0,049	–
Bailey et al, 2022[41]	Active Flexion-Extension ROM – L2L3 (Grades)	10,8 (1,1)	9.3 (1.2)	−1,5	–	–
Bailey et al, 2018[43]	Active Flexion-Extension ROM - L3L4 (Grades)	8,8 (4,9)	7,6 (4,8)	−1,27 (−17,3%)	0,016	–
Bailey et al, 2022[41]	Active Flexion-Extension ROM – L3L4 (Grades)	11,2 (1,1)	9,1 (1,2)	−2,1	–	–
Bailey et al, 2018[43]	Active Flexion-Extension ROM - L4L5 (Grades)	8,9 (3,1)	6,3 (2,5)	−2,65 (−30,3%)	0,004	–
Bailey et al, 2022[41]	Active Flexion-Extension ROM – L4L5 (Grades)	10,4 (0,9)	9,9 (1,0)	−0,5	–	–
Bailey et al, 2018[43]	Active Flexion-Extension ROM – L5S1 (Grades)	6,4 (1,4)	7,0 (3,4)	0,59 (5,3%)	0,69	–
Bailey et al, 2022[41]	Active Flexion-Extension ROM – L5S1 (Grades)	10,4 (1,7)	10,7 (1,3)	+0,3	–	–
Passive Flexion-Extension						
Bailey et al, 2018[43]	Passive Flexion-Extension ROM – L1L2 (Grades)	8,5 (5,5)	8.3 (8,3)	−0,25 (−17,1%)	0,46	–
Bailey et al, 2018[43]	Passive Flexion-Extension ROM – L2L3 (Grades)	3,9 (1,1)	4,2 (2,1)	0,27 (10,5%)	0,68	–
Bailey et al, 2018[43]	Passive Flexion-Extension ROM – L3L4 (Grades)	7,4 (3,8)	7,7 (3,1)	0,21 (17,7%)	0,56	–
Bailey et al, 2018[43]	Passive Flexion-Extension ROM – L4L5 (Grades)	9,0 (2,5)	10,8 (2,2)	1,76 (35,7%)	0,79	–
Bailey et al, 2018[43]	Passive Flexion-Extension ROM – L5S1 (Grades)	11,8 (6,0)	7,2 (4,5)	−4,5 (−40%)	0,031	–

(*Continued*)

Table 3 (Continued).

Active Lateral ROM						
Bailey et al, 2022[41]	Active Lateral ROM – L1L2 (Grades)	9,1 (0,6)	8,2 (0,9)	−0,9	–	–
Bailey et al, 2022[41]	Active Lateral ROM – L2L3 (Grades)	8,9 (0,9)	8,2 (0,7)	−0,7	–	–
Bailey et al, 2022[41]	Active Lateral ROM – L3L4 (Grades)	10.1 (0.9)	8.1 (0.9)	−2,0	–	–
Bailey et al, 2022[41]	Active Lateral ROM – L4L5 (Grades)	8.9 (1.1)	8.3 (1.6)	−0,5	–	–
Bailey et al, 2022[41]	Active Lateral ROM – L5S1 (Grades)	4.6 (0.8)	3.7 (0,4)	−0,9	–	–

Note: "-"Not Analyzed.
Abbreviations: NR, Not reported; US, Ultrasound; CSA, Cross-sectional area; FCSA, Functional Cross-sectional area.

https://doi.org/10.2147/JPR.S491060

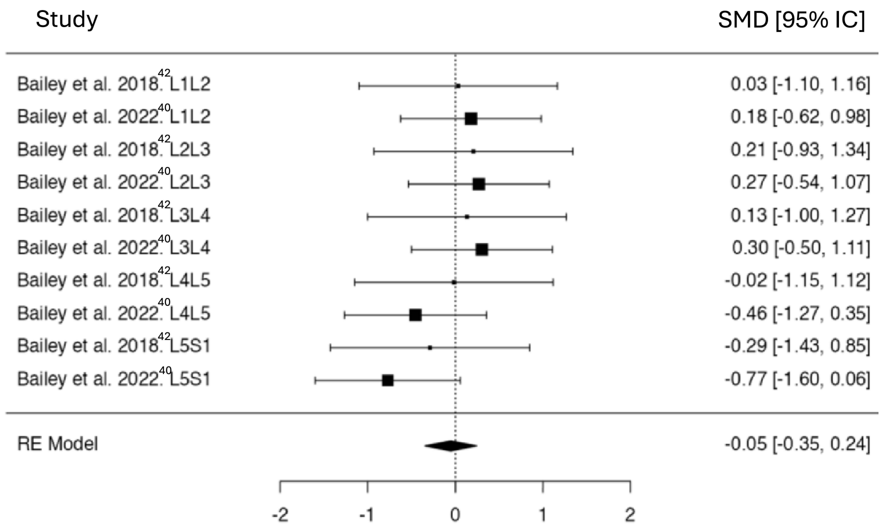

Figure 2 Forest Plot of Standardized Mean Differences in Intervertebral Disc Water Content Before and After Spaceflight.

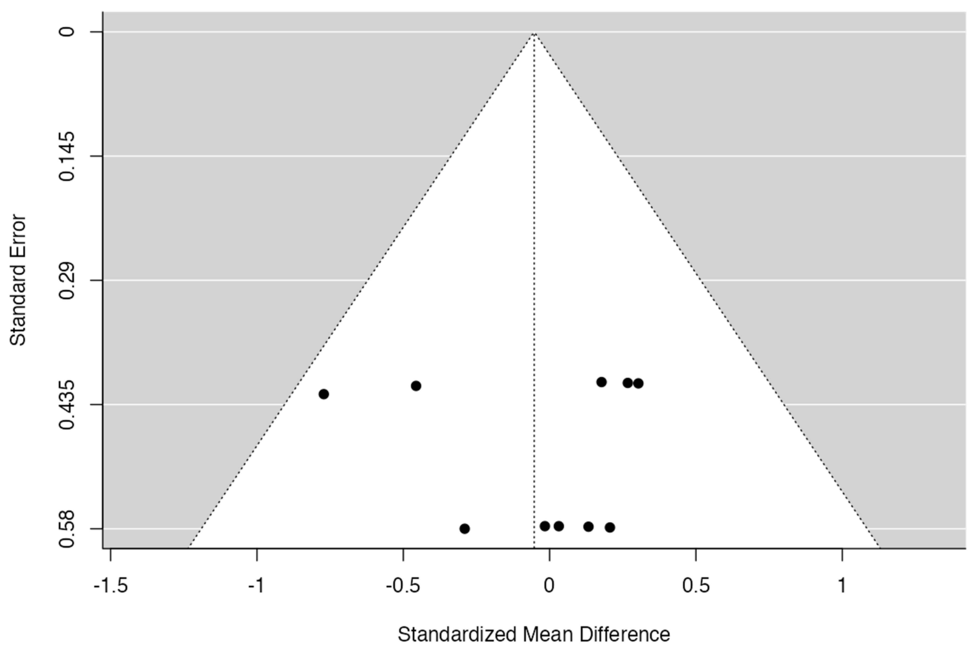

Figure 3 Funnel Plot Assessing Publication Bias in Studies of Intervertebral Disc Water Content Before and After Spaceflight.

were positive (70%), suggesting a consistent trend of increased ROM post-spaceflight. These findings imply that exposure to microgravity during spaceflights leads to a significant improvement in the ROM.The Q-test for heterogeneity was not significant, but some heterogeneity may still be present in the true outcomes ($Q = 16.502$, $p = 0.05$, $tau^2 = 0.20$, $I^2 = 46.20\%$). A 95% confidence interval for the true outcomes is given by −0.498 to 1.478. Hence, although the average outcome is estimated to be positive, in some studies the true outcome may in fact be negative. An examination of the studentized residuals revealed that none of the studies had a value larger than ± 2.807 and hence there was no indication of outliers in the context of this model. According to the Cook's distances, none of the studies could be considered to be overly influential. Neither the rank correlation nor the regression test indicated any funnel plot asymmetry ($p = 0.380$ and $p = 0.731$, respectively) (Figure 7).

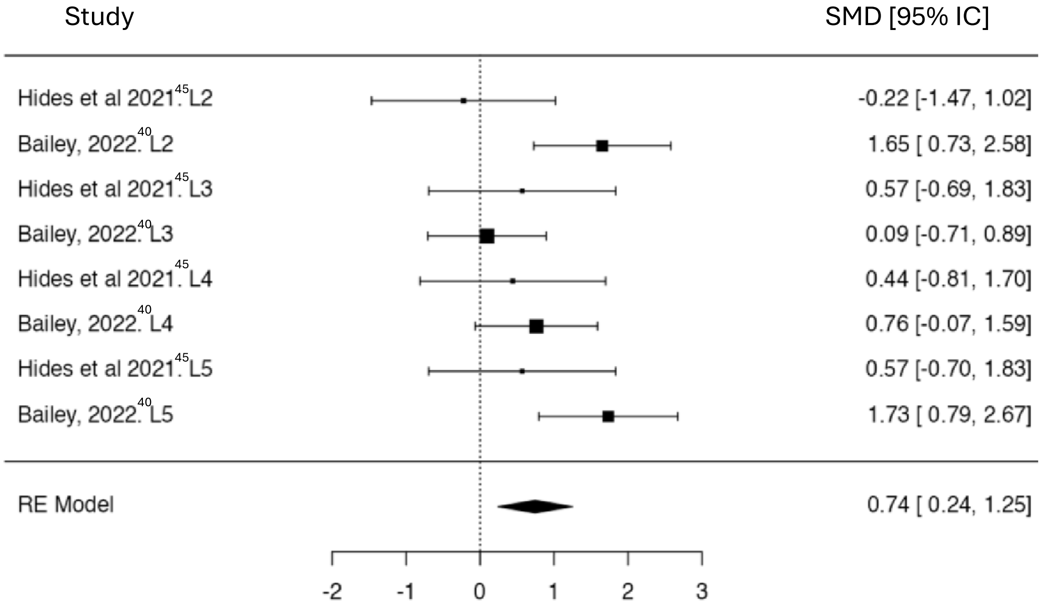

Figure 4 Forest Plot of Standardized Mean Differences in Multifidus Muscle Size Before and After Spaceflight.

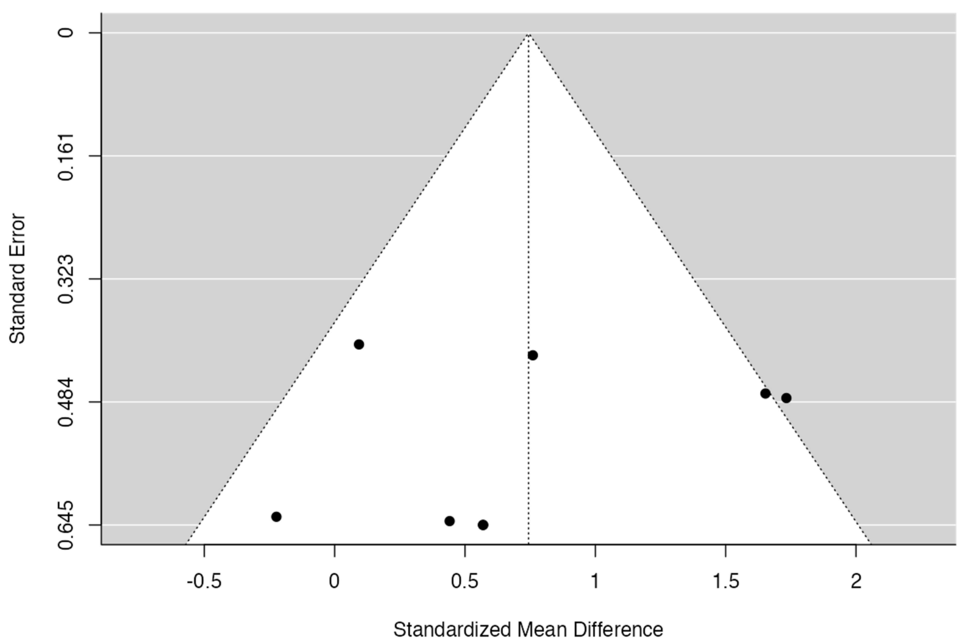

Figure 5 Funnel Plot Assessing Publication Bias in Studies of Multifidus Muscle Size Before and After Spaceflight.

Pain During and After Space Flight

Four studies assessed pain,[13,14,43,44] three during the flight,[13,14,44] three upon landing,[13,14,44] and one also evaluated pain at 12 months post-flight.[43]

During the flight, a high percentage of astronauts developed LBP (70%-85%).[14,44] Additionally, in the study by Pool-Goudzwaard et al,[14] astronauts reported a mean pain intensity of 5 out of 10 points. None of the astronauts who were pain-free during the flight had a history of LBP on Earth.[14] Out of the 12 astronauts without a history of LBP before the

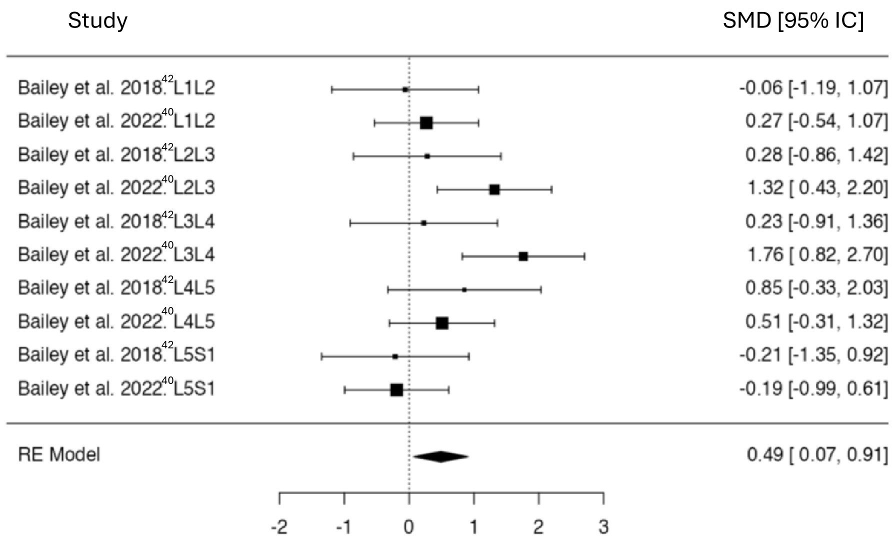

Figure 6 Forest Plot of Standardized Mean Differences in Lumbar Range of Motion Before and After Spaceflight.

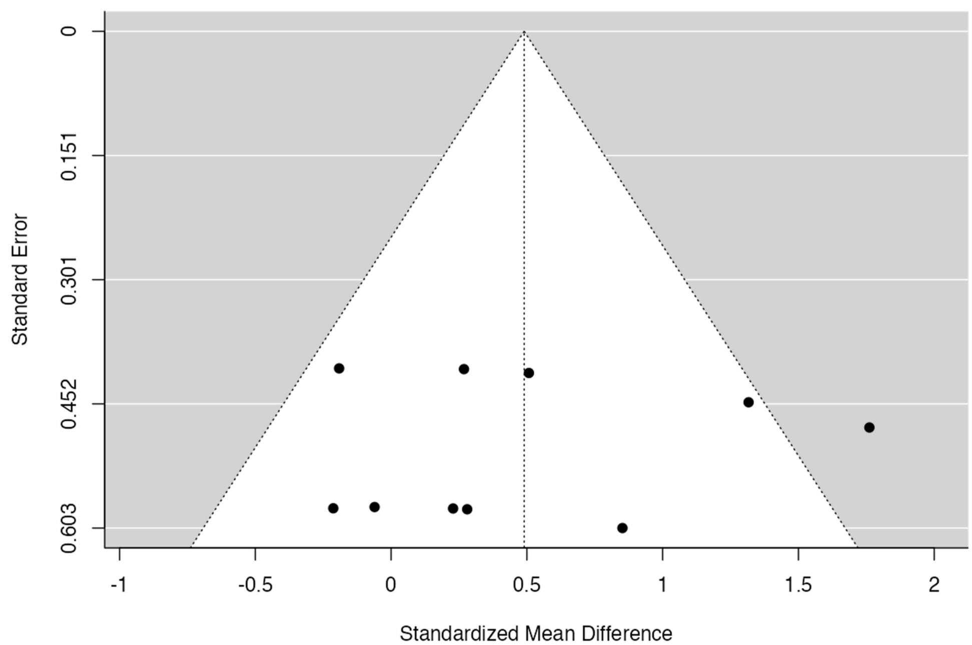

Figure 7 Funnel Plot Assessing Publication Bias in Studies on Lumbar Range of Motion Before and After Spaceflight.

flight, 4 experienced it during the flight.[14] There was a significant difference between the proportion of astronauts with and without previous back pain, and in the duration of LBP episodes (P < 0.01).[14] The most commonly reported regions of pain included the iliac crest at the posterior iliac spines on both sides, a broad central lower lumbar region, a small area at the height of the iliac crest, and at L5.[14] The main activities that triggered LBP were unknown, with 45% reporting pain after sleeping[14] (Table 4).

To assess the incidence of pain during spaceflight, a random-effects model with a total sample size of k = 3 was used. The estimated effect of pain during spaceflight was 0.768 (SE = 0.075, Z = 10.1, $p < 0.001$, 95% CI = 0.620–0.917), indicating that 77% of astronauts experience pain during spaceflight (Figure 8). There was no significant heterogeneity observed among the included studies (Q value = 0.987, p-value of 0.611; tau = 0.000, $tau^2 = 0$, SE = 0.0199, $I^2 = 0\%$, H^2

Table 4 Low Back Pain Incidence During and After Space Flights

Study	Outcome	Total of participants	Participants in pain	Incidence
Pool-Goudzwaard et al, 2015[13]	LBP in-flight	20	14	70%
Garcia et al, 2018[44]	LBP in-flight	7	6	85%
Sauer et al, 2023[48]	LBP in-flight	2	2	100%
Pool-Goudzwaard et al, 2015[13]	LBP (acute*) after flight	20	2	10%
Garcia et al, 2018[44]	LBP (acute*) after flight	7	4	57%
Sauer et al, 2023[48]	LBP (acute*) after flight	2	2	100%
Bailey et al, 2018[43]	Chronic LBP post-flight (12 months after flight)	6	2	33%

Note: *=Acute pain was between 1 day and 17 days after landed.

= 1.00).[33–36] Furthermore, the Egger test found that the regression slope was not significant (Z = 0.651, p = 0.515), suggesting no publication bias in the analyzed studies[37] (Figure 9).

After the flight, astronauts also experienced pain ranging from 10%[14] to 100%,[13] and at 12 months, 33% still had pain.[43] One study[13] assessed the pain of two astronauts during and after a 17-day spaceflight. Both astronauts reported musculoskeletal pain, managed with anti-inflammatories and stretching techniques during the flight.[13] Pain levels returned to baseline three months after landing.[13] Pain questionnaires revealed intense pain experiences during and immediately after the flight.[13]

To assess the incidence of acute pain following spaceflight, a random-effects model was employed with a total sample size of k = 3. The estimated effect of pain during spaceflight was 0.47 (SE = 0.242, Z = 1.95, $p < 0.052$, 95% CI = −0.003–0.943), indicating that 47% of astronauts experience pain during spaceflight (Figure 10). Significant high heterogeneity was observed among the included studies (Q = 14.814, p-value < 0.001, $\tau = 0.385$; $\tau^2 = 0.1483$ SE = 0.1832; $I^2 = 86.5\%$; $H^2 = 7.407$).[33–36] Furthermore, the Egger test found a significant regression slope (Z = 3.819, $p < 0.001$), suggesting publication bias in the analyzed studies[37] (Figure 11).

Qualitative interviews allowed the astronauts to describe their pain experiences during the flight.[13] When exposed to microgravity, Astronaut 1 described experiencing lower back pain and headaches. Although the lower back pain decreased after two to three days on the International Space Station (ISS), they continued to experience "a higher number of headaches than usual". They also mentioned feelings of "nausea, disorientation, and general discomfort", which gradually improved in two to three days. Notably, they expressed relief at the absence of shoulder pain during the

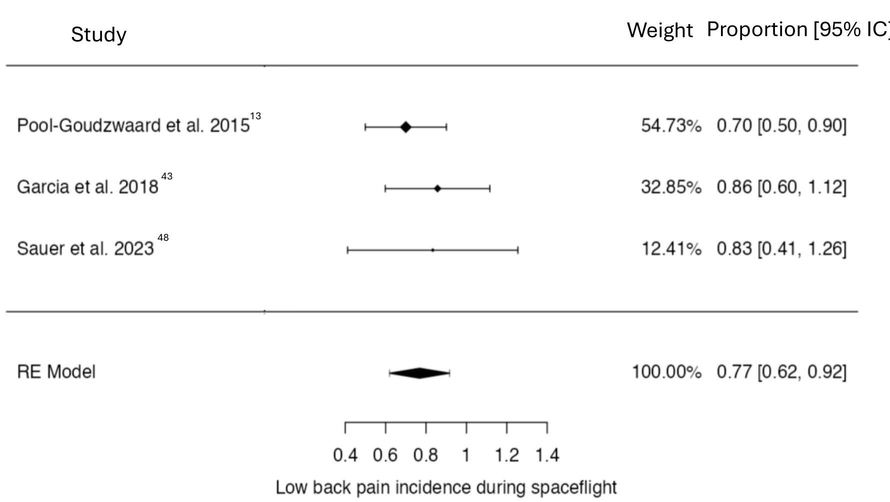

Figure 8 Forest Plot of the Incidence of Low Back Pain During Spaceflights.

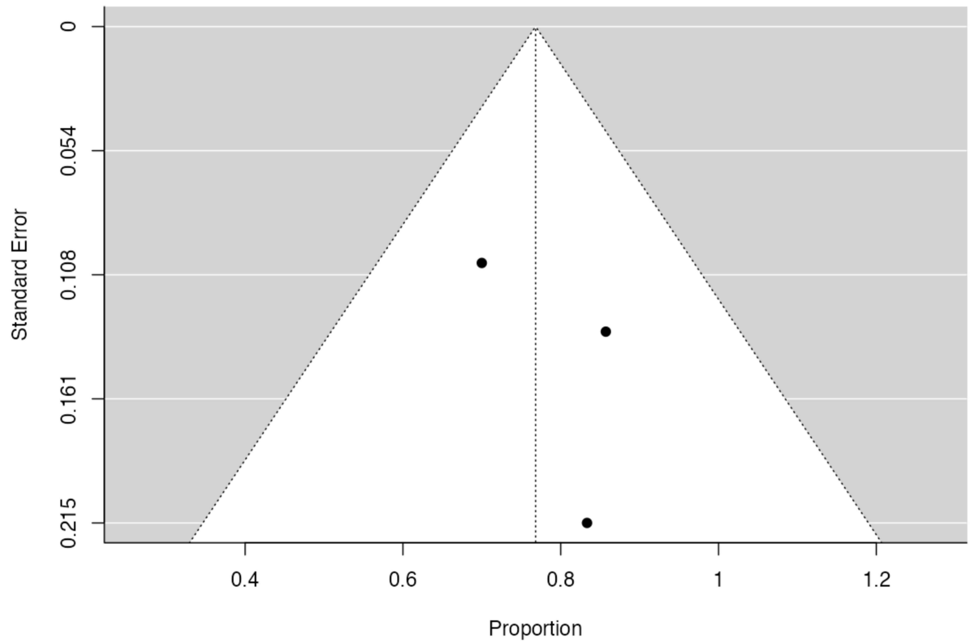

Figure 9 Funnel Plot Assessing Publication Bias in Studies on the Incidence of Low Back Pain During Spaceflights.

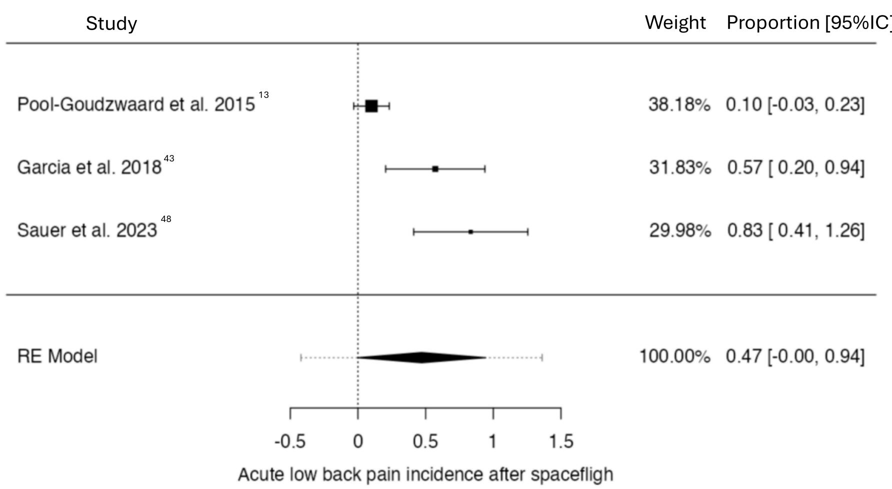

Figure 10 Forest Plot of the Incidence of Acute Low Back Pain After Spaceflights.

seventeen days in space, stating: "I didn't notice any shoulder pain at all during the seventeen days I was in space. That was great!". However, upon returning to Earth, they reported that the shoulder pain had returned "to where it was before". Additionally, they mentioned experiencing muscle pain and stiffness, especially in their calves, immediately after return, likening it to having had intense calf training. They also described lower back pain, characterized as "spasms", although seven days after their return to Earth, they reported that muscle pain and lower back pain had mostly decreased, commenting: "my muscles are getting used to carrying my body in 1G".

Astronaut 2 described previous episodes of common pain, usually related to injuries such as a tibia fracture four years prior.[13] During space training and aboard the ISS, they experienced predictable discomfort in their tibia, exacerbated by exercise and prolonged standing. In microgravity, they reported significant lower back pain from the outset and also mentioned pain in their left iliotibial band. Upon returning to Earth, they continued to feel pain in their distal tibia, especially when initiating movement after prolonged periods of rest.

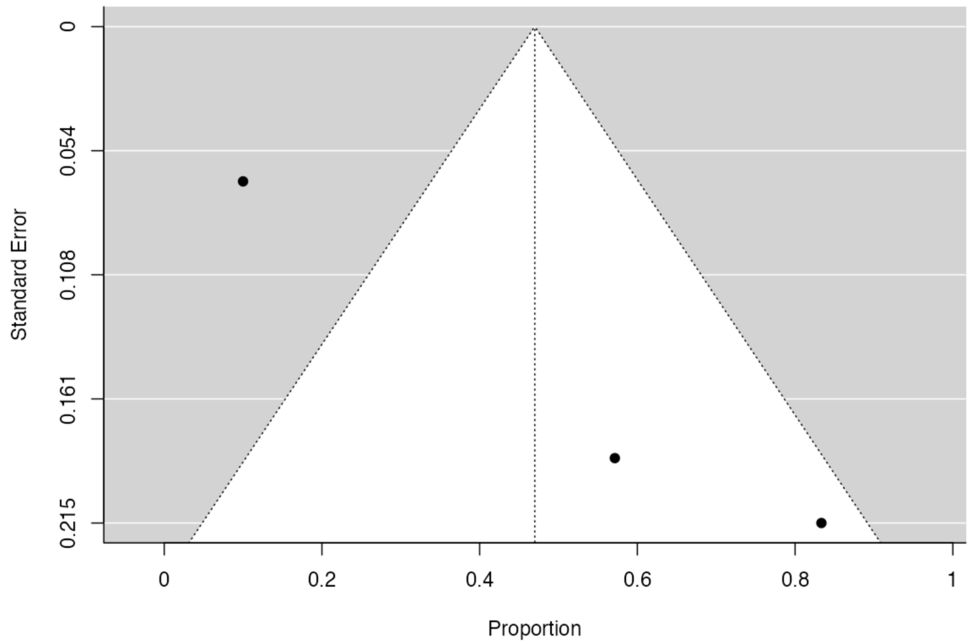

Figure 11 Funnel Plot Assessing Publication Bias in Studies on the Incidence of Acute Low Back Pain After Spaceflights.

Finally, sensory changes in both astronauts included increased thresholds for mechanical touch detection, temporal summation of pain, heat pain thresholds, and differences in conditioned pain modulation after the flight.[13] Therefore, this study suggests that spaceflight can affect various aspects of sensory perception and regulation in astronauts.

Risk of Bias in Studies Included

The risk of bias was analyzed with specific scales that were used for studies in humans and animals with an adequate level of agreement among examiners (Kappa index = 0.77). In the studies involving humans,[1,13,14,41,43–48] we found that all the studies indicated a low risk of bias in the "Prognostic Factors Measurement" and "Outcome Measurement" domains. The majority of issues were found in "Study Attrition", "Study Confounding", and "Statistical Analysis" where the risk of bias was moderate or high. In the article involving animals,[42] evaluated by SYRCLE,[20] the authors did not randomize the sample and did not blind the evaluator of the variables (Table 5).

Quality of Studies Included

The quality of the studies was evaluated with an inter-evaluator agreement of a Kappa index of 0.75. All the studies analyzed outcomes before the flight and used a valid and reliable scale for measuring these outcomes. None of the studies reported sample size calculations, all assessed the exposure only once over time, and none analyzed potential confounding factors such as sex or age in relation to pain (Table 6).

Quality of Evidence

The GRADE-based evaluations[25,26] were previously pilot tested with 2 reviewers for 30% of outcomes, achieving a substantial level of agreement (Kappa = 0.80). Study limitations were finally assessed by 2 independent reviewers, who had lower interrater reliability than in the pilot test, but still acceptable (Kappa = 0.71). The GRADE Assessment reveals low-quality evidence in "Trunk muscle atrophy", and moderate-quality in "Spinal pathologies", "Disc size", "Paraspinal muscle atrophy", "Range of motion". (Table 7).

Table 5 Risk of Bias of Studies Included

Animals	**SYRCLE's (SYstematic Review Centre for Laboratory animal Experimentation) RoB (Risk of Bias) tool**									
Author, year	**1. Was the allocation sequence adequately generated and applied?**	**2. Were the groups similar at baseline or were they adjusted for confounders in the analysis?**	**3. Was the allocation to the different groups adequately concealed during?**	**4. Were the animals randomly housed during the experiment?**	**5. Were the caregivers and/or investigators blinded from knowledge which intervention each animal received during the experiment?**	**6. Were animals selected at random for outcome assessment?**	**7. Was the outcome assessor blinded?**	**8. Were incomplete outcome data adequately addressed?**	**9. Are reports of the study free of selective outcome reporting?**	**10. Was the study apparently free of other problems that could result in high risk of bias?**
Bailey, 2014[42]	Yes	Yes	No	No	No	No	No	Yes	No	No
Humans	**Quality In Prognosis Studies (QUIPS) Risk of Bias tool**									
Author, year	**1. Study Participation**	**2. Study Attrition**	**3. Prognostic Factor Measurement**	**4. Outcome Measurement**	**6.Study Confounding**	**8. Statistical Analysis and Reporting**				
Pool-Goudzwaard et al, 2015[14]										
Chang et al,2016[1]										
Hides et al,2016[45]		NA			**NA**					
Bailey et al, 2018[43]										
Garcia et al,2018[44]										
Burkhart et al, 2019[48]										
Hides et al, 2021[46]										
Bailey et al, 2022[41]										
Coulombe et al, 2023[47]										

(*Continued*)

https://doi.org/10.2147/JPR.S491060

Table 5 (Continued).

Humans	**Quality In Prognosis Studies (QUIPS) Risk of Bias tool**									
Sauer et al, 2023[13]					NA	NA				
Low										
Moderate										
High										

Abbreviation: NA = Not Applicable.

https://doi.org/10.2147/JPR.S491060

Table 6 Quality of Studies Included

Animals	**The Collaborative Approach to Meta-Analysis and Review of Animal Data from Experimental Studies (CAMARADES) tool**													
Author, year	**1. Sample size calculation**	**2. Random allocation to treatment or control**	**3. Blinded induction of ischemia**	**4. Blinded assessment of outcome**	**5. Appropriate animal model**	**6. Use of anesthetic without significant intrinsic neuroprotective activity**	**7. Statement of control of temperature***	**8. Compliance with animal welfare regulations**	**9.Peer-reviewed publication**	**10. Statement of potential conflict of interests**				
Bailey et al, 2014[42]	No	No	NA	No	Yes	NA	Yes	Yes	Yes	Yes				
Humans	**NIH Quality assessment tool**													
Author, year	**1. Was the research question or objective in this paper clearly stated?**	**2. Was the study population clearly specified and defined?**	**3. Was the participation rate of eligible persons at least 50%?**	**4. Were all the subjects selected or recruited from the same or similar populations (including the same time period)? Were inclusion and exclusion criteria for being in the study prespecified and applied uniformly to all participants?**	**5. Was a sample size justification, power description, or variance and effect estimates provided?**	**6. For the analyses in this paper, were the exposure (s) of interest measured prior to the outcome (s) being measured?**	**7. Was the timeframe sufficient so that one could reasonably expect to see an association between exposure and outcome if it existed?**	**8. For exposures that can vary in amount or level, did the study examine different levels of the exposure as related to the outcome (eg, categories of exposure, or exposure measured as continuous variable)?**	**9. Were the exposure measures (independent variables) clearly defined, valid, reliable, and implemented consistently across all study participants?**	**10. Was the exposure (s) assessed more than once over time?**	**11. Were the outcome measures (dependent variables) clearly defined, valid, reliable, and implemented consistently across all study participants?**	**12. Were the outcome assessors blinded to the exposure status of participants**	**13. Was loss to follow-up after baseline 20% or less?**	**14. Were key potential confounding variables measured and adjusted statistically for their impact on the relationship between exposure (s) and outcome (s)?**
Bailey et al, 2018[43]	Yes	Yes	NR	Yes	Not	Yes	Yes	Not	Yes	Not	Yes	Not	Yes	Not
Chang et al,2016[1]	Yes	Yes	NR	Yes	Not	Yes	Yes	Not	Yes	Not	Yes	Not	Yes	Not
Garcia et al,2018[44]	Yes	Yes	NR	Yes	Not	Yes	Yes	Not	Yes	Not	Yes	Not	NR	Not
Hides et al,2016[45]	Yes	Yes	NA	NA	NA	Yes	Yes	Not	Yes	Not	Yes	NA	NA	NA
Hides et al,2021[46]	Yes	Yes	NR	Not	Not	Yes	Yes	Not	Yes	Not	Yes	Not	Yes	Not
Pool-Goudzwaard et al, 2015[14]	Yes	Yes	NR	Not	Not	Yes	Yes	Not	Yes	Not	Yes	Not	NR	Not

(*Continued*)

Table 6 (Continued).

Bailey et al, 2022[41]	Yes	Yes	Yes	Yes	Not	Yes	Yes	Not	Yes	Not	Yes	Not	NR	Not
Coulombe et al, 2023[47]	Yes	Yes	Yes	Yes	Not	Yes	Yes	Yes	Yes	Not	Yes	Not	Yes	Not
Burkhart et al, 2019[48]	Yes	Yes	Yes	Yes	Not	Yes	Yes	Not	Yes	Not	Yes	Not	Yes	Not
Sauer et al, 2023[13]	Yes	Yes	Yes	Yes	NA	Yes	Yes	Not	Yes	Not	Yes	Not	Yes	Not

Abbreviations: CD, cannot determine; NA, not applicable; NR, not reported.

Table 7 Quality of Evidence of Changes on Spinal Pathologies in Low Back After Space Flights Assessed by GRADE System

Factors	Number of participants	Number of studies	Phase	Study limitations	Inconsistency	Indirectness	Imprecision	Publication bias	Overall quality
Spinal pathologies	25	3 [2; 3; 6]	1	x	x	x	x	✓	++
Disc size	24	2 [1; 5]	1	x	✓	x	x	✓	+
Water content	18	2 [2; 3]	1	✓	x	x	x	✓	+
Lumbar lordosis	6	1[2]	1	x	NA	x	x	✓	++
Muscle atrophy	46	6 [2–5; 7; 8]	1 and 2	✓✓	✓	✓	x	✓	+++
Range of motion	18	2 [2; 3]	1	x	✓	✓	x	✓	+++

Notes: ✓ = no serious limitations, x = serious limitations, xx = very serious limitations; NA Not Applicable. +++=moderate, ++=low, +=very low.

Discussion

In this study we summarize the physiological changes in low back after spaceflights, and we describe the incidence of inflight and chronic LBP of astronauts. The results of 11 studies show several changes in the spine after short space flights and acute and chronic pain in both humans and mice.[1,13,14,41,43–48] Those studies included 93 astronauts that represent the 20% of the astronauts population in the history of spaceflights (500 people).[49]

Spinal Pathologies and Disc Dimensions

On Earth, intervertebral discs maintain their health by bearing loads and retaining water, thanks to proteoglycans, essential components that help absorb impacts.[50–55] During a prolonged stay in space, the intervertebral discs of astronauts undergo significant changes due to the lack of gravity.[1,13,14,41–48]

Space flights are associated with several changes in spine structure, particularly in the intervertebral discs. The absence of compressive loads in space, causes a decrease in the proteoglycan content of the discs.[50,52–55] Discs are continuously remodeled according to daily mechanical loads,[56] with less load (or total absence in space), the numbers of proteoglycans are reduced, and therefore the discs retain less water, affecting their ability to absorb impacts and making them more prone to herniation and pain.[50,52–55]

When astronauts return to Earth and their discs start bearing weight again, proteoglycan levels can recover since the intervertebral disc's extracellular matrices are continuously remodeled.[56] Studies have observed that discs recover their proteoglycan content after fourteen days of normal physiological loading, which could explain the adaptation of intervertebral discs to the microgravity environment of space and their subsequent adaptation to the terrestrial environment after landing.[52] However, we identified a significant gap in the literature regarding mid-term evaluations post-landing (eg, 30 days), which could help better understand these findings.

Muscle Atrophy

Muscle atrophy associated with space flights represents a significant challenge for the health and performance of astronauts. This phenomenon not only affects load distribution and spinal stability upon returning to Earth's gravity, but also has prolonged implications for recovery and muscle function.[47,48,57,58] It is worth noting that muscle atrophy has been the most studied factor in the studies included in this review, demonstrating significant concern. However, its effects on space flights are also beginning to be understood.[1,41,43,45,46,48]

Microgravity results in mechanical unloading of the muscles which may produce muscle atrophy during space flights.[57] On Earth, gravity provides a constant load that muscles must resist to maintain posture and perform movements. In space, the absence of this constant load leads to a decrease in muscle activity and consequently, a reduction in muscle mass and strength.[1,41,43,45,46,48] This atrophy is exacerbated by the duration of the flight, reduced physical activity, and changes in nutrition and metabolism.[57,58]

The muscles of the lumbar region, including the spinal erectors, multifidus, psoas, quadratus lumborum, transversus abdominis, and internal oblique, are crucial for lumbar spine stability.[59,60] Multifidus atrophy, particularly pronounced at levels L3, L4, and L5 after spaceflight,[41,45,46] is of particular concern due to its central role in lumbar stabilization[61,62]. The atrophy of lumbar region muscles during and after spaceflight may compromise the spine's ability to maintain necessary rigidity, increasing the risk of injuries and lower back pain.[57,61–63]

Among patients with chronic lower back pain, multifidus atrophy is particularly notable, suggesting a relationship between muscle loss and pain perception.[57,61–63] This finding is consistent with pain localized in the lumbar region during and after spaceflight, indicating that muscular atrophy may affects the physical function of astronauts.[13,58]

Muscle attenuation refers to a reduction in muscle quality or density, which can make muscles less efficient in their functions.[63] The muscle function is affected by a decrease in attenuation in the psoas, erector spinae, multifidus and quadratus lumborum muscles.[1,41,43,45,46,48] This attenuation, occurring independently of atrophy, suggests changes in muscle composition, possibly related to loss of muscle density or alterations in contractile properties.[45,46,48,57] The persistence of this attenuation, even after rehabilitation, underscores the potential need for the development of effective recovery postflight programs.

During spaceflights, the muscle atrophy and reduction of muscle strength, associated with the lack of gravity leads to a significant decrease in the ability of these muscles to generate force.[63] This loss of strength is further compounded by the infiltration of fat into the muscles, a process that further reduces their functional capacity and affects the muscle's efficiency in supporting the spine and protecting it from mechanical stress.[57,61–63] This reduction in the muscle's ability to generate force not only compromises general physical function during and after the flight but also makes it more difficult to recover spinal stability, prolonging the rehabilitation process in the medium and long term.[57,58] In rehabilitation, it is crucial to design specific programs that focus on restoring muscle strength to regain the ability to bear loads and perform functional movements. Without appropriate intervention, the loss of strength can perpetuate muscle dysfunction and extend the time needed for full recovery, affecting astronauts' effective reintegration into their daily and professional activities.

Range of Motion and Lumbar Lordosis

The absence of gravitational load on the spine alters the normal patterns of compression and tension experienced by the muscles and structures of the lumbar spine.[4,57] This can lead to deconditioning of the paravertebral and abdominal muscles, manifested as atrophy and reduced muscle attenuation.[1,41,43,45,46,48,57] The highest levels of muscle atrophy in the multifidus muscle occurred at the L4-L5 and L5-S1 levels,[41,46] coinciding with the greatest changes in ROM in the lumbar region.[41,43] Therefore, it appears that atrophy of the lumbar muscles, especially the lumbar multifidus, will result in decreased mobility in that region. This is because the muscles will have reduced contractile capacity and mobility, thus reducing the ROM in the affected vertebral segments.[41,45,46,48,61,62]

ROM assesses the functional capacity of the spine to move without restrictions, serving as a key measure for identifying limitations caused by muscle atrophy following exposure to microgravity.[41,45,46,48,61,62] These limitations in ROM may contribute to increased LBP, as reduced mobility impairs the proper distribution of loads in the spine, exacerbating the risk of injury.[1,41,43,45,46,48,57] In rehabilitation, improving ROM is essential not only for reducing pain but also for restoring functional movement patterns. Addressing ROM limitations early in the recovery process can enhance the effectiveness of strength training and reduce the overall time needed for rehabilitation.

Weightlessness also has a significant impact on the sensory perception and proprioception of astronauts.[64,65] On Earth, gravity provides constant sensory feedback through receptors in the joints, muscles, and skin, which are essential for precise control of movements and posture.[9,10,41–43,57] The lack of gravity reduces or eliminates the gravitational forces acting on the body, particularly on the spine.[6,43,52,57,58] This can lead to a decrease in the sensory stimulation that proprioceptive receptors normally receive, affecting the brain's ability to process and adjust the position and movement of spinal joints.[64,65] As a result, astronauts may experience a reduction in the accuracy and coordination of lumbar spine movements, thereby reducing the ROM in the lumbar spine.[41,43]

Pain During and After Space Flight

The onset of in-flight LBP often occurs within the first twenty-four hours and may persist throughout a mission.[14,44] In-flight surveys conducted over fifteen days found that pain incidence and intensity were highest in the first two days, after which there was a steady decrease in both incidence and intensity and no pain was reported after the ninth day.[14] The astronauts differentiated between types of pain, with the majority experiencing continuous pain (as opposed to intermittent pain) on the first day.[14] The incidence of post-flight LBP was significantly lower after a 15-day flight compared to that of long duration flight (10% vs 40–50%).[6,14] These results corroborate other reports of in-flight LBP primarily resolving after five days and rarely persisting beyond twelve days.[9]

Therefore, there appear to be two classes of LBP in space. On one hand, acute changes occurs within the first 24–48 hours and are responsible for in-flight LBP, coined "space adaptation pain", within the first 9–15 days.[13,14,44,46] Due to their minimal impact on post-flight LBP and disc herniation, it is unlikely that they lead to lasting decreases in load-bearing ability after reintroduction to terrestrial gravity. On the other hand, prolonged changes occur anytime after 12–15 days and appear to be unrelated to the occurrence of in-flight LBP.[1,13,14,41–48] These changes will likely impair spinal loading ability and appear mainly responsible for post-flight symptoms and injury.[1,13,14,41–48]

Microgravity-induced physiological changes have been demonstrated to align with some correlates of LBP on Earth.[1,13,14,41–48] The pathophysiology behind the in-flight LBP, post-flight disc herniation, and post-flight LBP in astronauts requires further study.

Limitations and Future Research

Despite the valuable insights provided by this study, there are several limitations that should be considered. While the number of astronauts included represents approximately 20% of all astronauts historically,[49] the small sample sizes in the studies do not allow for robust causal analysis.[66–68] Future research could focus on increasing sample sizes to enhance statistical power analysis. Additionally, moderate to high risks of selection bias and methodological limitations, particularly in follow-up and statistical analysis, may affect the robustness of conclusions drawn.[69] Addressing these concerns could involve implementing stricter protocols for participant follow-up and employing advanced statistical methods to mitigate bias. On the other hand, the variability in assessment methods across studies, including analyses of repeated measures within the same subjects but at different studies, introduces potential inconsistencies in results and complicates direct comparisons.[26] Future studies should aim for standardized assessment protocols to improve data comparability and reliability. While the role of the multifidus muscle in low back pain has been extensively studied, recent research highlights the critical role of the erector spinae.[70,71] Future studies on the effects of spaceflight on spinal musculature should incorporate a more comprehensive evaluation of both muscle groups in humans and animals. This would provide a deeper understanding of how different back muscles respond to microgravity and contribute to low back pain and post-flight recovery. Lastly, the limited number of studies investigating pain in spaceflight underscores the need for comprehensive research in this area. Further investigations could explore pain experiences across different mission durations and conditions, potentially informing targeted interventions to mitigate lumbar pain in astronauts. These identified limitations highlight avenues for future research aimed at strengthening the evidence base and enhancing understanding of the complex interactions between physiological changes, pain, and the spaceflight environment.

Conclusions

Astronauts experience muscle atrophy in the lumbar multifidus with a moderate to large effect, especially in the L4-L5 and L5-S1 segments, after space flights. Additionally, there is a reduction in the ROM with a moderate effect, along with disc herniations and disc dehydration. Seven out of ten astronauts develop pain during the spaceflight, four out of ten develop acute pain after spaceflight, and three out of ten develop chronic pain. Furthermore, the quality of this evidence ranges from moderate to low, and moderate to high risks of bias were identified in the studies, particularly in areas such as follow-up loss and statistical analysis, underscoring the need to enhance methodological quality in future research.

Abbreviations

LBP, Low back pain; MRI, Magnetic Resonance Imaging; CSA, Cross-sectional area; FCSA, Functional cross-sectional area; ISS, International Space Station; IVD, Intervertebral disc; GAG, Glycosaminoglycan; PSM, Paraspinal muscle size; STROBE, STrengthening the Reporting of OBservational studies in Epidemiology guidelines to observational studies; ROIs, Regions of interest; NP, Nucleus pulposus.

Acknowledgments

The academic and research activities of Dr. Ingelmo are supported by the Louis and Alan Edwards Family Foundation and by the Montreal Children's Hospital Foundation.

The project Pain in Space has been founded by the Montreal Children's Hospital Foundation by a special donation.

The first author, Guillermo Ceniza-Bordallo was funded with a predoctoral fellowship by the doctoral program in health care of the Faculty of Nursing, Physiotherapy and Podiatry of the Complutense University of Madrid and, Banco Santander (CT58/21-CT59/21).

Disclosure

The authors report no conflicts of interest related to this work.

References

1. Chang DG, Healey RM, Snyder AJ, et al Lumbar spine paraspinal muscle and intervertebral disc height changes in astronauts after long-duration spaceflight on the international space station. *Spine*. 2016;41(24):1917–1924. doi:10.1097/brs.0000000000001873
2. LeBlanc A, Schneider V, Shackelford L, et al Bone mineral and lean tissue loss after long duration space flight. *J Musculoskelet Neuronal Interact*. 2000;1(2):157–160.
3. Shackelford L. *Principles of Clinical Medicine for Space Flight.* Springer New York; 2019.
4. Lee SU, Hargens AR, Fredericson M, Lang PK. Lumbar spine disc heights and curvature: upright posture vs. supine compression harness. *Aviat Space Environ Med*. 2003;74(5):512–516.
5. Sayson JV, Hargens AR. Pathophysiology of low back pain during exposure to microgravity. *Aviat Space Environ Med*. 2008;79(4):365–373. doi:10.3357/asem.1994.2008
6. Johnston SL, Campbell MR, Scheuring R, Feiveson AH. Risk of herniated nucleus pulposus among U.S. astronauts. *Aviat Space Environ Med*. 2010;81(6):566–574. doi:10.3357/asem.2427.2010
7. Fritsch-Yelle JM, Charles JB, Jones MM, Wood ML. Microgravity decreases heart rate and arterial pressure in humans. *J Appl Physiol*. 1996;80(3):910–914. doi:10.1152/jappl.1996.80.3.910
8. Paul AM, Mhatre SD, Cekanaviciute E, et al Neutrophil-to-lymphocyte ratio: a biomarker to monitor the immune status of astronauts. *Front Immunol*. 2020;11:564950. doi:10.3389/fimmu.2020.564950
9. Penchev R, Scheuring RA, Soto AT, Miletich DM, Kerstman E, Cohen SP. Back pain in outer space. *Anesthesiology.* 2021;135(3):384–395. doi:10.1097/aln.0000000000003812
10. Wing PC, Tsang IK, Susak L, Gagnon F, Gagnon R, Potts JE. Back pain and spinal changes in microgravity. *Orthop Clin North Am*. 1991;22(2):255–262. doi:10.1016/S0030-5898(20)31651-5
11. Styf JR, Hutchinson K, Carlsson SG, Hargens AR. Depression, mood state, and back pain during microgravity simulated by bed rest. *Psychosom Med Nov-Dec*. 2001;63(6):862–864. doi:10.1097/00006842-200111000-00002
12. Kerstman EL, Scheuring RA, Barnes MG, DeKorse TB, Saile LG. Space adaptation back pain: a retrospective study. *Aviat Space Environ Med*. 2012;83(1):2–7. doi:10.3357/asem.2876.2012
13. Sauer AK, Vigouroux M, Dougherty PM, Cata JP, Ingelmo PM.Pain experience and sensory changes in astronauts during and after short-lasting commercial spaceflight: a proof-of-concept study. *J Pain Res*.2023;Volume 16:4253–4266. doi:10.2147/JPR.S440630
14. Pool-Goudzwaard AL, Belavý DL, Hides JA, Richardson CA, Snijders CJ. Low back pain in microgravity and bed rest studies. *Aerosp Med Hum Perform*. 2015;86(6):541–547. doi:10.3357/amhp.4169.2015
15. Egger M, Higgins JPT, Davey Smith G. *Systematic Reviews in Health Care: Meta-Analysis in Context*. 3th ed. John Wiley & Sons; 2022.
16. Moher D, Liberati A, Tetzlaff J, Altman DG, Group TP, The PRISMA Group. Preferred reporting items for systematic reviews and meta-analyses: the PRISMA statement. *PLoS Med*. 2009;6(7):e1000097. doi:10.1371/journal.pmed.1000097
17. Ouzzani M, Hammady H, Fedorowicz Z, Elmagarmid A. Rayyan-a web and mobile app for systematic reviews. *Syst Rev.* 2016;5(1):210. doi:10.1186/s13643-016-0384-4
18. Cohen JA. A coefficient of agreement for nominal scale. *Educ Psychol Meas*. 1960;20. doi:10.1177/001316446002000104
19. Landis JR, Koch GG. The measurement of observer agreement for categorical data. *Biometrics*. 1977;33(1):159–174. doi:10.2307/2529310
20. Hooijmans CR, Rovers MM, de Vries RB, Leenaars M, Ritskes-Hoitinga M, Langendam MW. SYRCLE's risk of bias tool for animal studies. *BMC Med Res Methodol*. 2014;14:43. doi:10.1186/1471-2288-14-43
21. Hayden JA, van der Windt DA, Cartwright JL, Côté P, Bombardier C. Research and reporting methods annals of internal medicine assessing bias in studies of prognostic factors. *Ann Intern Med*. 2013;158(4):6. doi:10.7326/0003-4819-158-4-201302190-00009
22. Macleod MR, O'Collins T, Howells DW, Donnan GA. Pooling of animal experimental data reveals influence of study design and publication bias. *Stroke*. 2004;35(5):1203–1208. doi:10.1161/01.str.0000125719.25853.20
23. Nhlbi N Study Quality Assessment Tools. Available from: https://www.nhlbi.nih.gov/health-topics/study-quality-assessment-tools. Accessed November 25, 2024.
24. Ma LL, Wang YY, Yang ZH, Huang D, Weng H, Zeng XT. Methodological quality (risk of bias) assessment tools for primary and secondary medical studies: what are they and which is better? *Mil Med Res*. 2020;7(1):7. doi:10.1186/s40779-020-00238-8
25. Guyatt GH, Oxman AD, Vist GE, et al GRADE: an emerging consensus on rating quality of evidence and strength of recommendations. *BMJ*. 2008;336(7650):924–926. doi:10.1136/bmj.39489.470347.AD
26. Huguet A, Hayden JA, Stinson J, et al Judging the quality of evidence in reviews of prognostic factor research: adapting the GRADE framework. *Syst Rev.* 2013;2(71). doi:10.1186/2046-4053-2-71
27. The jamovi project (Version 2.3). 2022. Available from: https://www.jamovi.org. Accessed November 25, 2024.
28. Library C. Chapter 10: analysing data and undertaking meta-analyses. *Cochrane Handbook for Systematic Reviews of Interventions*. Chichester: Wiley-Blackwell; 2021.
29. Ferguson CJ, Brannick MT. Publication bias in psychological science: prevalence, methods for identifying and controlling, and implications for the use of meta-analyses. *Psychol Methods*. 2012;17(1):120–128. doi:10.1037/a0024445
30. Viechtbauer W. Conducting meta-analyses in R with the metafor package. *Journal of Statistical Software*. 2010;36:1–48. doi:10.18637/jss.v036.i03
31. Lipsey M, Wilson D. *Practical Meta-Analysis*. SAGE Publications, Inc;; 2001.
32. Stijnen T, Hamza TH, Ozdemir P. Random effects meta-analysis of event outcome in the framework of the generalized linear mixed model with applications in sparse data. *Stat Med*. 2010;29(29):3046–3067. doi:10.1002/sim.4040
33. Cohen JF, Chalumeau M, Cohen R, Korevaar DA, Khoshnood B, Bossuyt PM. Cochran's Q test was useful to assess heterogeneity in likelihood ratios in studies of diagnostic accuracy. *J Clin Epidemiol*. 2015;68(3):299–306. doi:10.1016/j.jclinepi.2014.09.005
34. Higgins JP, Thompson SG. Quantifying heterogeneity in a meta-analysis. *Stat Med*. 2002;21(11):1539–1558. doi:10.1002/sim.1186
35. Takkouche B, Cadarso-Suárez C, Spiegelman D. Evaluation of old and new tests of heterogeneity in epidemiologic meta-analysis. *Am J Epidemiol*. 1999;150(2):206–215. doi:10.1093/oxfordjournals.aje.a009981
36. Thompson SG, Sharp SJ. Explaining heterogeneity in meta-analysis: a comparison of methods. *Stat Med*. 1999;18(20):2693–2708. doi:10.1002/(sici)1097-0258(19991030)18:20<2693::aid-sim235>3.0.co;2-v

37. Hayashino Y, Noguchi Y, Fukui T. Systematic evaluation and comparison of statistical tests for publication bias. *J Epidemiol*. 2005;15(6):235–243. doi:10.2188/jea.15.235
38. Andrade C. Mean difference, standardized mean difference (SMD), and their use in meta-analysis: as simple as it gets. *J Clin Psychiatry*. 2020;81 (5). doi:10.4088/JCP.20f13681
39. Viechtbauer W. Bias and efficiency of meta-analytic variance estimators in the random-effects model. *J Educ Behav Stat*. 2005;30(3):261–293. doi:10.3102/10769986030003261
40. Viechtbauer W, Cheung MW. Outlier and influence diagnostics for meta-analysis. *Res Synth Methods*. 2010;1(2):112–125. doi:10.1002/jrsm.11
41. Bailey JF, Nyayapati P, Johnson GTA, et al Biomechanical changes in the lumbar spine following spaceflight and factors associated with postspaceflight disc herniation. *Spine J*. 2022;22(2):197–206. doi:10.1016/j.spinee.2021.07.021
42. Bailey JF, Hargens AR, Cheng KK, Lotz JC. Effect of microgravity on the biomechanical properties of lumbar and caudal intervertebral discs in mice. *J Biomech*. 2014;47(12):2983–2988. doi:10.1016/j.jbiomech.2014.07.005
43. Bailey JF, Miller SL, Khieu K, et al From the international space station to the clinic: how prolonged unloading may disrupt lumbar spine stability. *Spine J*. 2018;18(1):7–14. doi:10.1016/j.spinee.2017.08.261
44. Garcia KM, Harrison MF, Sargsyan AE, Ebert D, Dulchavsky SA. Real-time ultrasound assessment of astronaut spinal anatomy and disorders on the international space station. *J Ultrasound Med*. 2018;37(4):987–999. doi:10.1002/jum.14438
45. Hides JA, Lambrecht G, Stanton WR, Damann V. Changes in multifidus and abdominal muscle size in response to microgravity: possible implications for low back pain research. *Eur Spine J*. 2016;25 Suppl 1:175–182. doi:10.1007/s00586-015-4311-5
46. Hides JA, Lambrecht G, Sexton CT, et al The effects of exposure to microgravity and reconditioning of the lumbar multifidus and anterolateral abdominal muscles: implications for people with LBP. *Spine J*. 2021;21(3):477–491. doi:10.1016/j.spinee.2020.09.006
47. Coulombe JC, Johannesdottir F, Burkhart KA, Brummer H, Allaire BT, Bouxsein ML. Changes in vertebral bone density and paraspinal muscle morphology following spaceflight and 1 year readaptation on earth. *JBMR Plus*. 2023;7(12):e10810. doi:10.1002/jbm4.10810
48. Burkhart K, Allaire B, Bouxsein ML. Negative effects of long-duration spaceflight on paraspinal muscle morphology. *Spine*. 2019;44(12):879–886. doi:10.1097/brs.0000000000002959
49. Security A, studies Cfsai. International astronaut database.Available from: https://aerospace.csis.org/data/international-astronaut-database/. Accessed July 2, 2024.
50. Urban JP, Roberts S. Degeneration of the intervertebral disc. *Arthritis Res Ther*. 2003;5(3):120–130. doi:10.1186/ar629
51. Urban JP, Holm S, Maroudas A, Nachemson A. Nutrition of the intervertebral disc: effect of fluid flow on solute transport. *Clin Orthop Relat Res*. 1982;170:296–302.
52. Jin L, Feng G, Reames DL, Shimer AL, Shen FH, Li X. The effects of simulated microgravity on intervertebral disc degeneration. *Spine J*. 2013;13 (3):235–242. doi:10.1016/j.spinee.2012.01.022
53. Inoue N, Espinoza Orías AA. Biomechanics of intervertebral disk degeneration. *Orthop Clin North Am*. 2011;42(4):487–99, vii. doi:10.1016/j.ocl.2011.07.001
54. Gawri R, Moir J, Ouellet J, et al Physiological loading can restore the proteoglycan content in a model of early IVD degeneration. *PLoS One*. 2014;9(7):e101233. doi:10.1371/journal.pone.0101233
55. Liu X, Krishnamoorthy D, Lin L, et al A method for characterising human intervertebral disc glycosaminoglycan disaccharides using liquid chromatography-mass spectrometry with multiple reaction monitoring. *Eur Cell Mater*. 2018;35:117–131. doi:10.22203/eCM.v035a09
56. Chan SC, Ferguson SJ, Gantenbein-Ritter B. The effects of dynamic loading on the intervertebral disc. *Eur Spine J*. 2011;20(11):1796–1812. doi:10.1007/s00586-011-1827-1
57. Lee PHU, Chung M, Ren Z, Mair DB, Kim DH. Factors mediating spaceflight-induced skeletal muscle atrophy. *Am J Physiol Cell Physiol*. 2022;322(3):C567–c580. doi:10.1152/ajpcell.00203.2021
58. Patron M, Neset M, Mielkozorova M, et al Markers of tissue deterioration and pain on earth and in space. *J Pain Res*. 2024;Volume 17:1683–1692. doi:10.2147/JPR.S450180.
59. Oliva-Lozano JM, Muyor JM. Core muscle activity during physical fitness exercises: a systematic review. *Int J Environ Res Public Health*. 2020;17 (12):4306. doi:10.3390/ijerph17124306
60. Imai A, Kaneoka K, Okubo Y, et al Trunk muscle activity during lumbar stabilization exercises on both a stable and unstable surface. *J Orthop Sports Phys Ther*. 2010;40(6):369–375. doi:10.2519/jospt.2010.3211
61. Wang K, Deng Z, Chen X, et al The role of multifidus in the biomechanics of lumbar spine: a musculoskeletal modeling study. *Bioengineering*. 2023;10(1). doi:10.3390/bioengineering10010067
62. Beneck GJ, Kulig K. Multifidus atrophy is localized and bilateral in active persons with chronic unilateral low back pain. *Arch Phys Med Rehabil*. 2012;93(2):300–306. doi:10.1016/j.apmr.2011.09.017
63. Chen ZN, Yao XM, Lv Y, et al Morphology of the lumbar multifidus muscle in lumbar disc herniation at different durations and at different ages. *Exp Ther Med*. 2018;15(5):4119–4126. doi:10.3892/etm.2018.5983
64. Clément G, Reschke M, Wood S. Neurovestibular and sensorimotor studies in space and Earth benefits. *Curr Pharm Biotechnol*. 2005;6 (4):267–283. doi:10.2174/1389201054553716
65. Clément G, Wood SJ. Rocking or rolling--perception of ambiguous motion after returning from space. *PLoS One*. 2014;9(10):e111107. doi:10.1371/journal.pone.0111107
66. Snijders T, Bosker R. *Multilevel Analysis: An Introduction to Basic and Advanced Multilevel Modeling*. 2th ed. SAGE Publications; 2011.
67. Hair J, Black W, Babin B, Anderson R. *Multivariate Data Analysis*. 7th ed. Pearson New International; 2014.
68. Cohen J. *Statistical Power Analysis for the Behavioral Sciences*. 2nd ed. Routledge;1998.
69. Grooten WJA, Tseli E, Olov Äng B, et al Elaborating on the assessment of the risk of bias in prognostic studies in pain rehabilitation using QUIPS—aspects of interrater agreement. *Diagnostic Progn Res*. 2019;3(1):11. doi:10.1186/s41512-019-0050-0
70. Ekşi MŞ, Özcan-ekşi EE. Fatty infiltration of the erector spinae at the upper lumbar spine could be a landmark for low back pain. *Pain Pract*. 2024;24(2):278–287. doi:10.1111/papr.13302
71. Ekşi MŞ, Öztaş UO, Topaloğlu F, et al Erector spinae could be the game changer in surgical decision-making in patients with lumbar spondylolisthesis: a cross-sectional analysis of age-, sex-, subtype-, level-matched patients with similar spinopelvic parameters who received surgical or conservative management. *Eur Spine J*. 2024;33:3715–3723. doi:10.1007/s00586-024-08341-3