

Physical Activity and its Adverse Effects on Health in Adults: an umbrella review

Prepared for the Guideline Development Group for the
WHO Guidelines on physical activity and sedentary behaviour for children and
adolescents, adults and older adults

Authors:

Matthias Lang, Bart Cillekens, Evert Verhagen, Willem van Mechelen, Pieter Coenen

**Department of Public and Occupational Health, Amsterdam Public Health Research Institute, Amsterdam
University Medical Centres, location VUmc, 1081 BT Amsterdam, The Netherlands.**

1. Background:

The Global Action Plan on Physical Activity 2018-2030 identified the need to update the 2010 guidelines on physical activity (PA) in youth, adults, and older adults. At the 142nd meeting of the World Health Organization (WHO) Executive Board (January, 2018), the Member States requested the WHO to prioritize the updating of the 2010 Global Recommendations on PA for Health (1). This request was added to the draft resolution in preparation for the World Health Assembly in May 2019. Over the past decade, there has been a large increase in the amount and quality of evidence on the different types, volumes, and durations of PA and their interrelationship with health. Particular areas of new evidence include the impact of PA on mental wellbeing and cognitive health outcomes, health outcomes in older adults and in sub-populations, such as those living with chronic conditions or impairments. Updated global guidelines on PA are an essential part of the policy framework and underpin national policy, and global and national surveillance of PA. The guidelines under development will also provide recommendations on PA and sedentary behavior for adults living with chronic conditions, including HIV, pregnant women, and adults living with specific impairments.

The information provided in this report is intended to inform the WHO Guideline Development Group (GDG) in order to update the WHO Global Recommendations on PA for Health.

2. Leisure-Time Physical Activity (LTPA)

The current guidelines for adults (18-64 years) recommend at least 150 minutes of moderate-intensity PA per week or at least 75 minutes of vigorous-intensity PA per week (1). Most of the research has been focused on the beneficial health effects of PA, but the adverse health effects of PA have so far not been integrated in these guidelines. Recreational sports, which is also often referred to as leisure time physical activity (LTPA), are competitive and non-competitive physical activities of LTPA, with leisure time being defined as discretionary time. Such recreational sports are participated in for health, enjoyment, amusement, or pleasure (2) and are considered to be “fun”, as opposed to professional sports. Research on the negative health-outcomes of leisure-time PA (LTPA), exclusive of occupational PA, has increased substantially over the last decade, leading to several systematic reviews related to the topic of adverse health outcomes of LTPA (i.e. recreational running, soccer, rugby).

The current umbrella review aims to review the evidence on the relationship between LTPA and a number of adverse health outcomes.

3. Aim

To review the evidence on the relationship between LTPA and adverse health outcomes, specifically regarding: injuries, osteoarthritis, erectile dysfunction, and exposure to pollution during LTPA.

4. Review Question

What is the relationship between LTPA and adverse health outcomes?

Sub questions

- a. Is there a dose-response relationship (total volume and/or duration, frequency, intensity) between LTPA and adverse health outcomes?

- b. Does this relationship vary by type of LTPA?

5. Methods:

This umbrella review was a-priori registered in PROSPERO. The PROSPERO registration number is not known yet, but will be added to this review as soon as it has been received (). This umbrella review synthesized and combined relevant data from systematic reviews or meta-analyses, in order to inform the WHO in their development of Guidelines for PA.

6. Inclusion and exclusion criteria:

Peer-reviewed reviews were eligible for inclusion in this umbrella review if they met all of the following inclusion criteria:

- A measure of LTPA was reported;
- An assessment of a relevant adverse health outcome was reported, examples of which are described below;
- Full-text systematic reviews were available, based on more than one paper (preferably containing meta-analyses), describing studies with an intervention-based, cross-sectional or longitudinal design.

Reviews were excluded if:

- They covered samples of elite or professional sports persons only (including paid, sponsored, and/or scholarship athletes);
- In case of mixed samples of non-elite and elite athletes: data regarding the non-elite samples could not be extracted separately;
- They focused on a clinical population, which cannot be generalized to the general population;
- They had falls as a risk or adverse event, with a focus on the elderly population (this outcome is considered by another WHO review);
- They were published in a language other than English.

7. Population-Exposure-Control-Outcome

The WHO Guideline Development Group decided to use PECO (Population-Exposure-Control-Outcome) questions to define the scope of their guidelines.

Population: Adults 18 years of age and older

Exposure: Duration, frequency and/or intensity of LTPA (dose of LTPA), or a composite score reflecting total volume of LTPA

Comparison: No LTPA, or LTPA of a lesser duration, frequency and/or intensity, or composite score of total volume of LTPA

Outcomes: Adverse health effects (especially injury, osteoarthritis, erectile dysfunction, and exposure to air pollution)

8. Search and Selection

In order to identify relevant evidence, a search for existing systematic reviews (preferably with meta-analyses) was conducted. The following databases were searched for systematic reviews that met the inclusion criteria: PubMed, SPORTDiscus, and Embase. Systematic searches were conducted in December 2019, limited from 2009 onwards (since this was an update of the WHO guidelines from

2009 (1)) and contained the following sets of key terms: harms and injuries, physical activity, and publication type. The full search strategy can be found in Supplementary file 3. Snowball searches by screening reference lists of included studies and by consultation experts were used, to identify additional reviews that were not found in the three databases mentioned above. Final search results were exported to Endnote reference manager and the final counts were captured in a PRISMA flow diagram (3).

9. Study Selection

Two reviewers (BC and ML) assessed the title and abstract of each identified study in a first selection round, assessing the in- or exclusion criteria for each article as described above. This was done blinded in the Rayyan web app (4). If no consensus could be reached, conflicts between the independent reviewers were resolved in a consensus meeting, with the help of a third reviewer (EV).

10. Full Text Search Selection

After obtaining full text articles, two researchers (BC and ML) performed a full text screening of the remaining studies after the initial study selection. If the inclusion criteria were met, the systematic review was included in the subsequent assessments. This was done blinded in the Rayyan web app also (4). Any conflicts were resolved with the help of the third reviewer (EV). The final numbers were updated in a final version of the PRISMA flow diagram, as part of the final report.

11. Assessing Bias in Systematic Reviews

The included systematic reviews were assessed for quality using the AMSTAR 2 (i.e. A Measurement Tool To Assess Systematic Reviews) (5). This is a 16-point assessment tool for assessing the methodological quality of systematic reviews. AMSTAR 2 has a good inter-rater agreement, test-retest reliability, and content validity. The rating values are High, Moderate, Low, and Critically Low. The cutoff values of properly addressing each of the 16 points were 100%, ≥75%, ≥50%, and below 50%, respectively. One reviewer (ML) assessed the risk of bias of the studies included. A second reviewer (BC) reviewed the initial assessment and in case of disagreement, consensus was reached through discussion. If a review was rated Critically Low, this review was excluded since it was judged that the review outcome would not provide an accurate summary of the available evidence.

12. Data extraction

We extracted author's name, year of publication, number of studies included (by study design), and summarized effect estimates from each included systematic review. In case a review dealt with a mixed population of non-elite and elite sports and when possible, we extracted effect estimates only for the non-elite population.

13. Grading the body of Evidence (GRADE)

For each included review the Grading of Recommendations, Assessment, Development and Evaluation (GRADE) method (6) was used to rate the certainty of evidence for each of the outcomes of interest.

The GRADE system rates the quality of evidence in one of four levels (high, moderate, low, and very low). The starting point for the quality of the evidence was 'high' for longitudinal studies that sought to confirm independent associations between the prognostic factor and the outcome (7, 8). The

rating may be decreased if the reviews showed: risk of bias (e.g. selection, performance, detection, attrition, and/or reporting bias), inconsistency of results (e.g. unexplained heterogeneity, if the I^2 statistics was provided a cutoff of >50% was used for an inconsistent result), indirectness of evidence (differences in populations, difference in intervention, difference in outcomes measures, indirect comparisons), imprecision (e.g. 95% Confidence Interval includes Risk Ratio of 1.0) or reporting bias (9). The rating of evidence could be increased by one level if there was a large magnitude of the effect (RR or OR was >2.0 or <0.5) or a dose-response gradient (10).

If the quality of the evidence was rated high, further research is very unlikely to change the rated confidence regarding the estimate of the effect. If the quality was rated very low, any estimate of the effect is very uncertain (6).

14. Results

The PRISMA flowchart of the search and selection process is presented in Figure 1. The database searches generated 991 references, leaving 741 articles for title & abstract screening after duplicates were removed. This left 64 articles for the eligibility assessment after another 677 articles were excluded based on their title and abstract screening. After the full text screening, another 49 records were excluded (see supplementary file 4 for details). A total of 15 systematic reviews were identified that examined an association between LTPA and adverse health outcomes. The most commonly reported outcome was injury ($n = 12$), followed by osteoarthritis ($n = 2$) and erectile dysfunction ($n = 1$). No systematic reviews were found that evaluated LTPA and an association with exposure to air pollution (Table 1). Since the search strategy was specific for the outcomes stated in the review question, no other outcomes were found nor passed the inclusion criteria. AMSTAR scores, reflecting methodological quality of the included reviews is shown in supplementary file 1. The quality of three reviews was considered critically low; these reviews were not considered for further analyses. Extracted information and GRADE scores of included reviews are shown in supplementary file 2. An overview of the evidence identified for the relationship between LTPA and adverse health outcomes is provided in the next section. Evidence from moderate, low, or very low quality reviews pointed towards LTPA (mainly running) being related with musculoskeletal injuries as the adverse health outcome. Also from very low quality evidence reviews it was suggested that LTPA has a protective effect against knee and hip OA and a reduced risk of undergoing surgery due to knee OA.

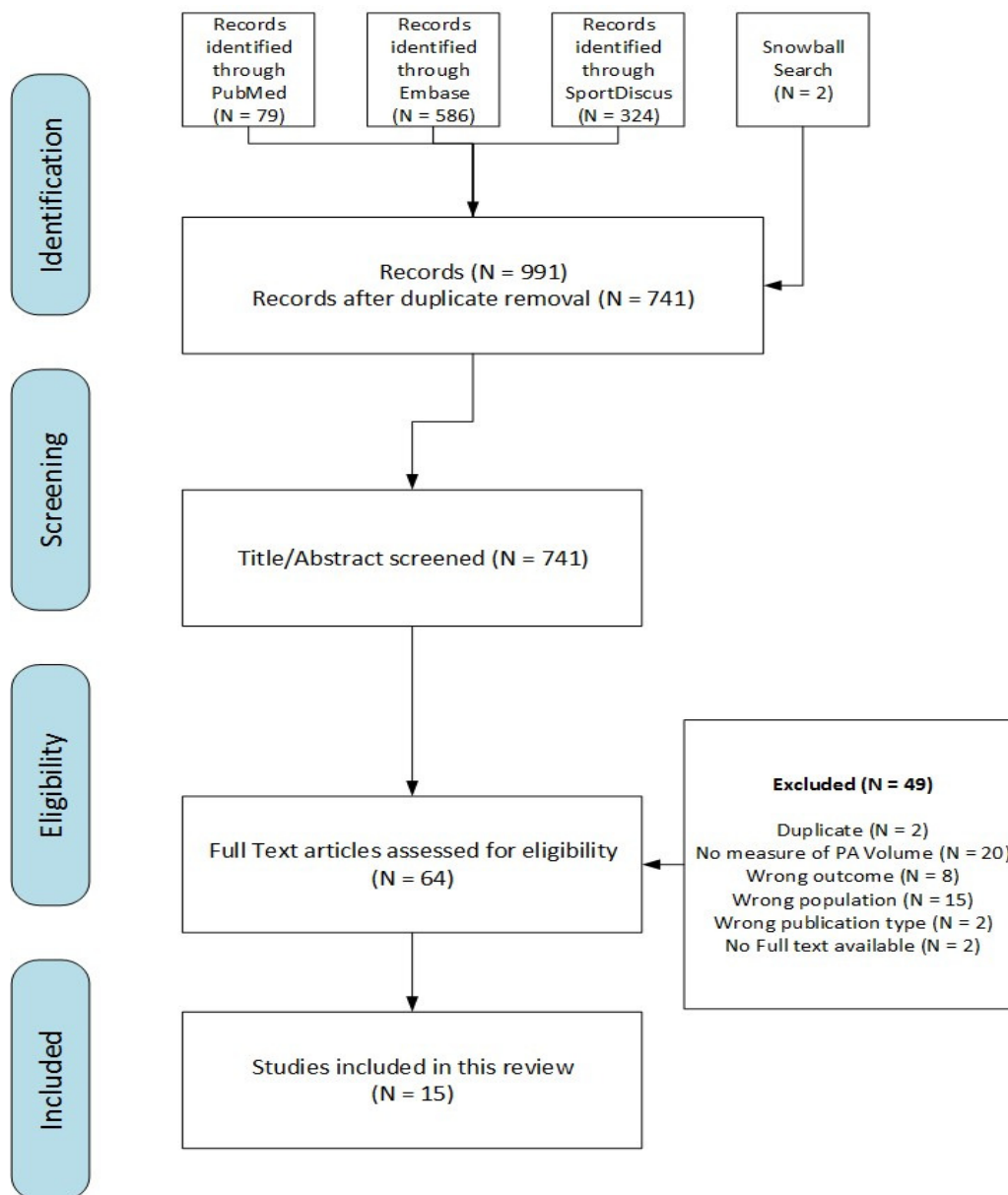


Figure 1 - PRISMA process flowchart of the Search & Selection Strategy

The applicable outcome of each included systematic review is shown in the table below.

Table 1 - Summary of the Outcomes of the included systematic reviews

Author, Year	Injury	Osteoarthritis	Erectile Dysfunction	Exposure to Pollution
Alentorn-Geli, 2017 (11)		X		
Borel, 2019 (12)	X			
Damsted, 2018 (13)	X			
Johnston, 2018 (14)	X			
Lopes, 2012 (15)	X			
Neubauer, 2016 (16)	X			
Nielsen, 2012 (17)	X			
Qu, 2014 (18)	X			
Saragiotto, 2014 (19)	X			
Sobhani, 2013 (20)	X			
Sommer, 2016 (21)			X	
Timmins, 2017 (22)		X		
Tonoli, 2010 (23)	X			
Van der Worp, 2015 (24)	X			
Videbaek, 2015 (25)	X			

15. Summary of the Grade Evidence

Evidence Portfolio – Injury

What is the relationship between leisure-time physical activity and adverse health outcomes?

Sub questions

- a. Is there a dose response relationship (total volume and/or duration, frequency, intensity)?
- b. Does the relationship vary by type of LTPA?

Sources of Evidence: Existing Systematic Reviews, Meta-Analyses

Conclusion statements and grades

Moderate evidence indicates that there is a relationship between LTPA and (musculoskeletal) injuries. **Grade: Moderate.**

Low evidence suggests that there is a relationship between LTPA and injury incidence rates and injury prevalence. **Grade: Low.**

Low evidence suggests that there is a protective relationship between LTPA and risk of fracture. **Grade: Low.**

Low evidence suggests that there is a dose response relationship (e.g. total distance, frequency, and duration) for musculoskeletal injuries. **Grade: Low.**

Insufficient evidence is available to determine whether the relationship varies by type of LTPA. **Grade: Not assignable.**

Description of the Evidence

A search for systematic reviews, meta-analyses, and pooled analyses found moderate level evidence for the guiding question. Low level evidence was found for the first sub-question. However, the search did not identify sufficient literature to fully answer the second sub-question.

A search for systematic reviews, meta-analyses, and pooled analyses found low level evidence for the guiding question and first sub-question (e.g. dose response relationship). However, the search did not identify sufficient literature to fully answer the second sub-question (e.g. variation by type of LTPA).

Existing Systematic Reviews and Meta-Analyses

Overview

A total of 10 existing reviews were included: 3 meta-analyses (12, 18, 25) and 7 systematic reviews (13-15, 17, 19, 20, 24). The reviews were published from 2012 to 2019.

The meta-analyses included a range of 10 to 23 studies that addressed running-related musculoskeletal injuries and risk of fracture. The Meta-analyses covered the following timeframe: from inception to submission of the publication (12, 25), and from 1950 to February 2013 (18).

The systematic reviews included a range of 4 to 28 studies that addressed musculoskeletal injury, ankle and foot overuse injury, and risk factors for injury. Reviews covered the following timeframes: from inception to within one year of publication (13, 14, 17, 20), from 1947 to October 2011 (15), from 1946 to December 2012 (19), and from 1806 to January 2013 (24).

Exposures

The included reviews examined running (12, 13, 15, 17, 19, 24, 25), endurance sports (14), various team sports (20), and high vs. low categories of PA (18) as measures of LTPA.

Outcomes

The included reviews addressed various injuries as an outcome, including running-related (musculoskeletal) injuries, pain, risk of fracture, risk factors for injury, and ankle and foot overuse injuries.

Evidence Portfolio – Osteoarthritis

What is the relationship between leisure-time physical activity and adverse health outcomes?

Sub questions

- a. Is there a dose response relationship (total volume and/or duration, frequency, intensity)?
- b. Does the relationship vary by type of LTPA?

Sources of Evidence: Existing Systematic Reviews, Meta-Analyses

Conclusion statements and grades

Very low evidence suggests that LTPA has a protective effect for the onset of knee and/or hip OA.

Grade: Very Low.

Very low evidence suggests that there is a dose response relationship (e.g. exposure time) for the onset of knee and/or hip OA. **Grade: Low.**

Insufficient evidence is available to determine whether the relationship varies by type of LTPA.

Grade: Not assignable.

Description of the Evidence

A search for systematic reviews, meta-analyses, and pooled analyses found low level evidence for the guiding question and first sub-question (e.g. dose response relationship). However, the search did not identify sufficient literature to fully answer the second sub-question (e.g. variation by type of LTPA).

Existing Systematic Reviews and Meta-Analyses

Overview

A total of 2 existing reviews were included: 2 meta-analyses (11, 22). The reviews were published in 2017.

The meta-analyses included a range of 3 to 17 studies that addressed OA. All Meta-analyses covered an extensive timeframe: from inception to within one year of publication (11, 22).

Exposures

All of the included reviews examined running and orienteering as LTPA and included a comparison group.

Outcomes

All included reviews addressed OA as an outcome.

Evidence Portfolio – Other adverse health outcomes, including erectile dysfunction and exposure to air pollution

What is the relationship between leisure-time physical activity and adverse health outcomes?

Sub questions

- a. Is there a dose response relationship (total volume and/or duration, frequency, intensity)?
- b. Does the relationship vary by type of LTPA?

Sources of Evidence: Existing Systematic Reviews

Conclusion statements and grades

Insufficient evidence is available to determine whether there is a significant relationship between LTPA and erectile dysfunction. **Grade: Not assignable.**

No evidence is available to determine whether there is a significant relationship between LTPA and exposure to air pollution. **Grade: Not assignable.**

Supplementary file 1. Outcome-specific AMSTAR 2 summary rating of the included systematic reviews

Author, Year	PECO ¹	A priori Methods ²	Study Design Selection ³	Search Strategy ⁴	Study Selection ⁵	Data Extraction ⁶	Excluded Studies ⁷	Included Studies ⁸	RoB Assessment ⁹	Funding Sources ¹⁰	Statistical Methods ¹¹	Impact of RoB ¹²	RoB Results ¹³	Heterogeneity ¹⁴	Publication Bias ¹⁵	COI ¹⁶	Overall Rating ¹⁷
Injury																	
Borel, 2019 (12)	Y	PY	N	PY	PY	PY	PY	N	Y	N	PY	PY	PY	Y	Y	PY	Low
Damsted, 2018 (13)	Y	N	PY	Y	PY	PY	PY	Y	Y	N	N/A	N/A	N/A	N/A	N/A	PY	Low
Johnston, 2018 (14)	PY	Y	PY	Y	PY	PY	PY	Y	Y	Y	N/A	N/A	N/A	N/A	N/A	PY	Low
Lopes, 2012 (15)	Y	N	Y	Y	PY	N	PY	PY	Y	Y	N/A	N/A	N/A	N/A	N/A	Y	Low
Neubauer, 2016 (16)	PY	N	N	PY	PY	N	N	N	N	N	N/A	N/A	N/A	N/A	N/A	PY	Critically Low
Nielsen, 2012 (17)	Y	N	PY	Y	PY	N	Y	PY	Y	N	N/A	N/A	N/A	N/A	N/A	N	Low
Qu, 2014 (18)	Y	N	PY	PY	Y	PY	N	Y	PY	Y	PY	PY	N	N	Y	Y	Low
Saragiotto, 2014 (19)	PY	N	Y	PY	PY	PY	PY	PY	Y	Y	N/A	N/A	N/A	N/A	N/A	Y	Low
Sobhani, 2013 (20)	PY	N	N	Y	Y	N	Y	Y	Y	Y	N/A	N/A	N/A	N/A	N/A	PY	Low
Tonoli, 2010 (23)	PY	N	N	PY	N	N	N	N	PY	N	N/A	N/A	N/A	N/A	N/A	N	Critically Low
Van der Worp, 2015 (24)	PY	N	PY	Y	PY	Y	N	Y	Y	Y	N/A	N/A	N/A	N/A	N/A	Y	Low
Videbaek, 2015 (25)	PY	N	PY	Y	PY	N	PY	PY	Y	Y	PY	Y	Y	Y	PY	Y	Low
Osteoarthritis																	
Alentorn-Geli, 2017 (11)	PY	N	PY	Y	PY	PY	PY	Y	Y	Y	Y	N	PY	Y	N	PY	Low

DRAFT review prepared for the WHO Guideline Development Group
FOR CONSULTATION ONLY

Timmins, 2017 (22)	Y	Y	PY	Y	PY	PY	PY	Y	Y	Y	Y	Y	Y	Y	Y	Y	Moderate
Erectile Dysfunction																	
Sommer, 2016 (21)	PY	N	N	N	N	N	N	PY	N	N	N/A	N/A	N/A	N/A	N/A	PY	Critically Low

Abbreviations: COI = conflict of interest; N = no; PECO = population, exposure, comparator, outcome; PY = partial yes; RoB = risk of bias; Y = yes

¹ Did the research questions and inclusion criteria for the review include the components of PECO?

² Did the report of the review contain an explicit statement that the review methods were established prior to the conduct of the review and did the report justify any significant deviations from the protocol?

³ Did the review authors explain their selection of the study designs for inclusion in the review?

⁴ Did the review authors use a comprehensive literature search strategy?

⁵ Did the review authors perform study selection in duplicate?

⁶ Did the review authors perform data extraction in duplicate?

⁷ Did the review authors provide a list of excluded studies and justify the exclusions?

⁸ Did the review authors describe the included studies in adequate detail?

⁹ Did the review authors use a satisfactory technique for assessing the risk of bias (RoB) in individual studies that were included in the review?

¹⁰ Did the review authors report on the sources of funding for the studies included in the review?

¹¹ If meta-analysis was performed did the review authors use appropriate methods for statistical combination of results?

¹² If meta-analysis was performed, did the review authors assess the potential impact of RoB in individual studies on the results of the meta-analysis or other evidence synthesis?

¹³ Did the review authors account for RoB in individual studies when interpreting/ discussing the results of the review?

¹⁴ Did the review authors provide a satisfactory explanation for, and discussion of, any heterogeneity observed in the results of the review?

¹⁵ If they performed quantitative synthesis did the review authors carry out an adequate investigation of publication bias (small study bias) and discuss its likely impact on the results of the review?

¹⁶ Did the review authors report any potential sources of conflict of interest, including any funding they received for conducting the review?

¹⁷ Shea et al. 2017. AMSTAR 2: a critical appraisal tool for systematic reviews that include randomised or non-randomised studies of healthcare interventions, or both. (5) For each Yes, 1.0 point was given, and for each Partial Yes, 0.5 points was given. The total sum was then divided by the number of questions answered (i.e. 11 or 16). The rating ranges were described in the Methods above.

DRAFT review prepared for the WHO Guideline Development Group
FOR CONSULTATION ONLY

DRAFT review prepared for the WHO Guideline Development Group
FOR CONSULTATION ONLY

DRAFT - for consultation

DRAFT review prepared for the WHO Guideline Development Group
FOR CONSULTATION ONLY

Supplementary file 2. GRADE: Grading the body of evidence

Injury

Population: Adults 18 years of age and older
Exposure: Duration, frequency and/or intensity of LTPA, or a composite score reflecting total volume of LTPA
Comparison: No LTPA, or LTPA of a lesser duration, frequency and/or intensity, or composite score of total volume of LTPA.
Outcome: Adverse health outcomes (especially injury, osteoarthritis, erectile dysfunction, and exposure to pollution)

Certainty assessment							Summary Effect	Certainty	Importance
No of studies	Study design	Risk of bias	Inconsistency	Indirectness	Imprecision	Other considerations			

Prevalence of injuries in Brazilian recreational street runners: meta-analysis (Borel et al., 2019) (12)

DRAFT review prepared for the WHO Guideline Development Group
FOR CONSULTATION ONLY

Certainty assessment							Summary Effect	Certainty	Importance
No of studies	Study design	Risk of bias	Inconsistency	Indirectness	Imprecision	Other considerations			
23 ^a	20 cross-sectional 3 prospective cohort	Serious ^b	Not serious	Not serious	Not serious	None	Adverse health outcomes The pooled prevalence of injuries was 36.5% (95% CI 30.8-42.5%). Based on a total population of 3,786 runners. Injury prevalences were given by location and type: - knee injury: 32.9% (95% CI 26.7-39.6%). - ankle injury: 17.7% (95% CI 11.2-26.9%). - hip injury: 13.3% (95% CI 6.9-24.1%). - muscle injuries (including strains and contractures): 27.9% (95% CI 18.2-40.1%). - The prevalence of ligament injuries (e.g. sprains and dislocations): 27.8% (95% CI 19.4-38.1%). - inflammatory lesions (i.e. plantar fasciitis, tendinitis, synovitis, bursitis, and medial tibial stress syndrome): 26.5% (95% CI 14.9-40.1%). - bone injuries (i.e. fracture, chondromalacia patella and bone edema): 5.6% (95% CI 1.8-16.3%). The following LTPA exposure relationships were reported: - six studies showed a relationship between running distance of 20km or more per week and the occurrence of injury. - five studies showed a relationship between running experience of more than five years and the occurrence of injury. - four studies showed a relationship between a training frequency > 3 days/week and the occurrence of injury. Dose-response relationship No effects to report, due to heterogeneity no meta-analysis was performed and hence no ORs or RRs were provided. Type of LTPA No effects to report, while only one type of LTPA (i.e. running) was included in the review.	Moderate ^{rr}	CRITICAL

Is there evidence for an association between changes in training load and running-related injuries? A systematic review (Damsted et al., 2018) (13)

DRAFT review prepared for the WHO Guideline Development Group
FOR CONSULTATION ONLY

Certainty assessment							Summary Effect	Certainty	Importance
No of studies	Study design	Risk of bias	Inconsistency	Indirectness	Imprecision	Other considerations			
4	1 randomized controlled trial ^c 1 cross-sectional ^d 2 prospective cohort ^e	Serious ^f	Serious ^g	Not serious	Not serious	None	Adverse health outcomes The following LTPA exposure relationships were reported: 2 out of 4 reviews found an association between sudden increases in training load (>30%) and an increased risk of running-related injury: i.e. increasing the weekly running distance by more than 30% compared to a less than 10% change in the previous week Evidence from two studies defined sudden increases in training load as: - Changing one or more of the running variables (velocity, distance, , frequency, or volume) compared with the non-injured runners - The mean difference between the increase in the running distance the week before the onset of an injury and the average weekly increase during other weeks was found to be 86% Dose-response relationship No effects to report, due to heterogeneity no meta-analysis was performed and hence no ORs or RRs were provided. Type of LTPA No effects to report, only one type of LTPA (i.e. running) was included in the review.	Low ^{gs}	CRITICAL

The associations between training load and baseline characteristics on musculoskeletal injury and pain in endurance sport populations: a systematic review (Johnston et al., 2018) (14)

DRAFT review prepared for the WHO Guideline Development Group
FOR CONSULTATION ONLY

Certainty assessment							Summary Effect	Certainty	Importance
No of studies	Study design	Risk of bias	Inconsistency	Indirectness	Imprecision	Other considerations			
10	10 prospective cohorts ^a	Serious ⁱ	Serious ^j	Serious ^k	Not serious	None	Adverse health outcomes No effects to report; this review only looked at dose-response relationships. Dose-response relationship • medium effect size association between high total training distances per week/month (not specified in the review) and increased rate of injury and/or pain • medium effect size association between low training frequency (<2 sessions/week) and increased rate of injury and/or pain. • large effect size association between short training duration (<2 hours/week) and increased rate of injury and/or pain; however, this finding may be less applicable given that recreational endurance participation was defined by some studies as a training frequency of three to six training sessions/week (26, 27) and training duration of two to four hours/week (28). Definitions of the effect sizes: • Small: OR ≥ 1.5 or RR ≥ 2 • Medium: OR ≥ 2 or RR ≥ 3 • Large: OR ≥ 3 or RR ≥ 4 Type of LTPA Running and Triathlon as LTPA were included in the review.	Very Low ⁱⁱ	CRITICAL

What are the main running-related musculoskeletal injuries? A systematic review (Lopes et al., 2012) (15)

DRAFT review prepared for the WHO Guideline Development Group
FOR CONSULTATION ONLY

Certainty assessment							Summary Effect	Certainty	Importance
No of studies	Study design	Risk of bias	Inconsistency	Indirectness	Imprecision	Other considerations			
8	2 prospective cohorts ^a 1 clinical trial ^m 2 retrospective cohorts ^b 3 cross-sectional ^c	Serious ^p	Serious ^q	Not serious	Not serious	Electronic searches were conducted only in the main databases related to the sports injuries field	Adverse health outcomes - Prevalence and incidence numbers based on a total population of 3,500 runners - The highest incident rate of running-related musculoskeletal injury (RRMIs) was found for patellar tendinopathy (22.7%) - The most prevalent general RRMi was found for plantar fasciitis (17.5%) - The most prevalent RRMi during ultra-marathon races was found for ankle dorsi-flexors tendinopathy (29.6%). - The most frequently general RRMIs reported were: - medial tibial stress syndrome (incidence rate ranging from 13.6% to 20.0%; prevalence rate of 9.5%) - Achilles tendinopathy (incidence rate ranging from 9.1% to 10.9%; prevalence rate ranging from 6.2% to 9.5%) - plantar fasciitis (incidence rate ranging from 4.5% to 10.0%; prevalence rate ranging from 5.2% to 17.5%) - For RRMIs sustained during ultra-marathon races, the most frequently reported injuries were: - Achilles tendinopathy (prevalence rate ranging from 2.0% to 18.5%) - Patellofemoral syndrome (prevalence rate ranging from 7.4% to 15.6%) Dose-response relationship No effects to report; this review only looked at the adverse health outcomes. Type of LTPA No effects to report; only one type of LTPA (i.e. running) was included in the review.	Low ^{uu}	CRITICAL

Stress fractures of the femoral neck: a review (Neubauer et al., 2016) (16)

							Due to Critically Low rating on the AMSTAR 2 scale, this systematic review did NOT qualify for assessing the GRADE evidence.		CRITICAL
--	--	--	--	--	--	--	--	--	----------

DRAFT review prepared for the WHO Guideline Development Group
FOR CONSULTATION ONLY

DRAFT review prepared for the WHO Guideline Development Group
FOR CONSULTATION ONLY

Certainty assessment							Summary Effect	Certainty	Importance
No of studies	Study design	Risk of bias	Inconsistency	Indirectness	Imprecision	Other considerations			

Training errors and running related injuries: a systematic review (Nielsen et al., 2012) (17)

28	7 Retrospective cohort ^f 12 Prospective cohort ^g 6 Case-control ^f 3 Randomized controlled trial ^h	Serious ^v	Serious ^w	Not serious	Not serious	None	Adverse health outcomes - In several studies, an increased risk, relative risk, or odds ratio for sustaining an RRI was reported when the weekly running frequency increased: persons running 6-7 times per week had the highest risk of RRI. - Based on the studies reviewed, it was not possible to identify which training errors were related to running related injuries. Dose-response relationship No effects to report, due to heterogeneity no meta-analysis was performed and hence no ORs or RRs were provided. Type of LTPA No effects to report; only one type of LTPA was included in the review.	Low ^{wv}	CRITICAL
----	--	----------------------	----------------------	-------------	-------------	------	--	-------------------	----------

Association between physical activity and risk of fracture (Qu et al., 2014) (18)

DRAFT review prepared for the WHO Guideline Development Group
FOR CONSULTATION ONLY

Certainty assessment							Summary Effect	Certainty	Importance
No of studies	Study design	Risk of bias	Inconsistency	Indirectness	Imprecision	Other considerations			
15	15 Prospective Cohort ^a	Serious ^y	Serious ^z	Not serious	Not serious	None	Adverse health outcomes - In the meta-analysis, the participants (total of 1,235,768) with the highest category of LTPA had an approximately 29% lower risk of total fractures compared to those in the lowest category, with 95% CIs of 0.63–0.80 ($p<0.01$) and moderate heterogeneity across studies ($p=0.0$, $I^2=74.2\%$). - The analysis for subtypes of fracture showed a statistically significant inverse relationship between higher category of LTPA and risk of hip or wrist fracture (39% and 28% lower risk, respectively); i.e. among individuals in the highest category of LTPA, compared to those in the lowest category (95% CIs were 0.54–0.69 and 0.49–0.96, respectively, all $p<0.01$). - The association between LTPA and vertebral fracture risk was not statistically significant (RR, 0.87; 95% CI, 0.72–1.03; $p<0.01$). - Regarding age: the stratified analysis found an RR of 0.76 (95% CI, 0.60–0.91; $p=0.52$; $I^2=85.0\%$) for adults <62 years old. Dose-response relationship The review did not conduct a dose-response analysis, the existence of a dose-response relationship between LTPA and fracture risk remains unknown. Type of LTPA No effects to report; the review did not specify the type of LTPA in their high vs. low or moderate vs. sedentary comparisons.	Low ^{www}	CRITICAL

What are the main risk factors for running-related injuries? (Saragiotto et al., 2014) (19)

9	9 Prospective Cohort ^{aa}	Serious ^{bb}	Serious ^{cc}	Not serious	Not serious	Even though the electronic search was conducted in the main databases related to the sports-injuries field, it is possible that eligible articles have been published in journals not indexed in any of the searched databases.	Adverse health outcomes No effects to report; this review only looked at risk factors for adverse health outcomes. Two of the five studies that investigated weekly distance as a risk factor identified that training for more than 64 km a week was a risk factor for lower extremity injuries. Dose-response relationship No effects to report, due to heterogeneity no meta-analysis was performed and hence no ORs or RRs were provided. Type of LTPA No effects to report; only one type of LTPA was included in the review.	Low ^{xx}	CRITICAL
---	------------------------------------	-----------------------	-----------------------	-------------	-------------	---	---	-------------------	----------

DRAFT review prepared for the WHO Guideline Development Group
FOR CONSULTATION ONLY

Certainty assessment							Summary Effect	Certainty	Importance
No of studies	Study design	Risk of bias	Inconsistency	Indirectness	Imprecision	Other considerations			

Epidemiology of ankle and foot overuse injuries in sports: a systematic review (Sobhani et al., 2013) (20)

6	4 Prospective cohort ^{dd} 2 Cross-sectional ^{ee}	Serious ^{ff}	Serious ^{gg}	Serious ^{hh}	None	None	Adverse health outcomes The highest incidences of ankle and foot injury, expressed per 1000 athletes per season, were reported for: - sports dance (ballet), 338.5 (95% CI: 283.2-401.4) - running, 250.0 (95% CI: 100.5-515.1), - gymnastics, 188.7 (95% CI: 90.5-347.0) Dose-response relationship No effects to report, due to heterogeneity no meta-analysis was performed and hence no ORs or RRs were provided. Type of LTPA No effects to report; multiple sports were included in the review, but all summarizing conclusions were based on elite sports.	Very Low ^{yy}	CRITICAL
---	---	-----------------------	-----------------------	-----------------------	------	------	---	------------------------	----------

Incidence, risk factors and prevention of running related injuries in long-distance running: a systematic review (Tonoli et al., 2010) (23)

							Due to Critically Low rating on the AMSTAR 2 scale, this systematic review did NOT qualify for assessing the GRADE evidence.		CRITICAL
--	--	--	--	--	--	--	--	--	----------

Injuries in runners: a systematic review on risk factors and sex differences (van der Worp et al., 2015) (24)

DRAFT review prepared for the WHO Guideline Development Group
FOR CONSULTATION ONLY

Certainty assessment							Summary Effect	Certainty	Importance
No of studies	Study design	Risk of bias	Inconsistency	Indirectness	Imprecision	Other considerations			
13	11 Prospective cohort ⁱ 2 Retrospective cohort ⁱⁱ	Serious ^{kk}	Serious ^l	Serious ^{mm}	Not Serious	Selection & Publication bias ^{kk}	Adverse health outcomes This review only looked at risk factors for adverse health outcomes. Factors that increased the risk of running-related injuries in women were: • older age: not specified • previous participation in non-axial sports (e.g. cycling, swimming, etc.) • participating last year in a marathon • running on concrete surface • longer weekly running distance (> 48–63.8 km) • wearing the same running shoes for > 4 to 6 months Men were at greater risk of running-related injuries if they had: • restarted running • history of previous injuries • running experience of 0–2 years • weekly running distance between 32–47.8 km • weekly running distance more than 64 km per week Dose-response relationship No effects to report, due to heterogeneity no meta-analysis was performed and hence no ORs or RRs were provided. Type of LTPA No effects to report; only one type of LTPA (i.e. running) was included in the review.	Very Low ^{zz}	CRITICAL

Incidence of Running-Related injuries per 1000 h of running in different types of runners: a systematic review and meta-analysis (Videbaek et al., 2015) (25)

12	7 Prospective cohort ⁿⁿ 5 Randomized controlled trials ^{oo}	Serious ^{pp}	Serious ^{qq}	Not Serious	Not Serious	None	Adverse health outcomes The weighted estimates showed that novice runners faced a significantly greater injury rate of 17.8 (95 % CI 16.7–19.1; population of 2,480) than recreational runners, who sustained 7.7 (95 % CI 6.9–8.7; population of 946) running-related injuries per 1000 hours of running. The weighted estimate showed an injury incidence for all types of runners combined of 1.07 (95 % CI 1.01–1.13) per 1000 km of running. (Based on a total population of 2,248 runners). Dose-response relationship No effects to report; the meta-analysis only looked at the adverse health outcomes relationship. Type of LTPA No effects to report; only one type of LTPA (i.e. running) was included in the review.	Low ^{aaa}	CRITICAL
----	--	-----------------------	-----------------------	-------------	-------------	------	--	--------------------	----------

PA = physical activity; OR = odds ratio; CI = confidence interval; RRMI = running-related musculoskeletal injury; RRI = running-related injury; RR = relative risk;

DRAFT review prepared for the WHO Guideline Development Group
FOR CONSULTATION ONLY

DRAFT review prepared for the WHO Guideline Development Group
FOR CONSULTATION ONLY

a: Authors did not provide breakdown of the cross-sectional and prospective cohort studies. Abiko et al., 2017; Araujo et al., 2015; Campos et al., 2016; Fernandes et al., 2014; Ferreira et al., 2012; Hespanhol Junior et al., 2012; Hespanhol Junior et al., 2013; Hino et al., 2009; Ishida et al., 2013; Lopes et al., 2011; Oliveira et al., 2012; Oliveira EGA, Santos-Filho SD, 2018; Pazin et al., 2008; Pileggi et al., 2010; Purim et al., 2014; Rangel et al., 2016; Rios et al., 2017; Rolim et al. 2015; Salicio et al., 2017; Saragiotto et al., 2016; Souza et al., 2014; Yamato et al., 2011

b: Most studies did not use a standard definition of injury. The limitations of the study include the moderate quality of evidence, use of self-administered questionnaires that can lead to memory bias, and lack of standard collected information, thus compromising a more detailed interpretation of the data.

c: Buist et al., 2008

d: Cantidio Ferreira et al., 2012

e: Nielsen et al., 2013; Nielsen et al., 2014

f: Among the non-randomized studies, the most frequent reasons for decreased quality scores were: low external validity, a follow-up period shorter than 12 weeks, and lack of reporting a measure of association, while the risk of bias was more related to the absence of blinding procedures in the included randomized trial.

g: Due to the heterogeneity observed in the study designs, the runners' profiles, as well as the methods used for data collection and analysis of changes in training load, comparison of the results of the four studies included in the present systematic review must be performed with caution.

h: Bovens et al., 1989; Hein et al., 2014; Hespanhol Junior et al., 2013; Lysholm and Wiklander, 1987; Malisoux et al., 2015; Nielsen et al., 2013; Nielsen et al., 2014; Taunton et al., 2003; van Middelkoop et al., 2007; Zwingerberger et al., 2014

i: There was variability in definitions of injury and/or pain, external training load, baseline assessments, data collection and statistical analysis.

j: No heterogeneity I^2 tests were performed or provided.

k: The generalizability of results should be considered given that nine studies involved a recreational ESP whereas three involved an elite ESP.

l: Lysholm, Wiklander, 1987; Pileggi et al., 2010

m: Jakobsen et al., 1994

n: Jacobs, Berson, 1986; McKean et al., 2006

o: Fallon 1996; Hutson 1984; Scheer, Murray, 2011

p: (Outcome) Although most of the studies have a clear definition of RRMIs the definitions always differ between studies. In terms of the participants, only one study performed a random sample selection and two studies sampled the entire target population of runners. There are only three prospective studies that could enable the assessment of the loss to a follow-up criterion, and all of them fulfilled this criterion, which indicated a lower risk of bias in these studies.

q: No heterogeneity I^2 tests were performed or provided.

r: Koplan 1982; Koplan 1995; Marti 1988a; Marti 1988b; McKean 2006; Valliant 1981; Wen 1997

s: Bovens 1989; Fields 1990; Hootman 2002; Kelsey 2007; Lysholm 1987; Macera 1989; Middelkoop 2008; Pollock 1977; Satterthwaite 1999; Taunton 2003; Walter 1989; Wen 1998

t: Colbert 2000; Duffey 2000; McCrory 1999; Messier 1995; Messier 1991; Messier 1988

u: Buist 2008; Jakobsen 1994; Mechelen 1993

v: The types of participants (novice, recreational, and elite), and the injury definition used varied considerably between the studies.

w: No heterogeneity I^2 tests were performed or provided.

x: Mussolino et al., 1998; Hoidrup et al., 2001; Lau et al., 2001; Feskanich et al., 2002; Roy et al., 2003; Samelson et al., 2006; Thorpe et al., 2006; Michaelsson et al., 2007; Robbins et al., 2007; Appleby et al., 2008; Lee et al., 2010; Trimpou et al., 2010; Armstrong et al., 2011; Nikander et al., 2011; Morseth et al., 2012

y: The quality of individual studies varied; some of these may have had limited adjustments for potential statistical confounders. The classification of the quantity of physical activity is difficult to evaluate, a fact that inevitably weakens the strength of the identified association. Differences in methodology between studies may also introduce heterogeneities.

z: Most of the I^2 estimates calculated in this meta-analysis were assessed as moderate. The overall I^2 value of 74.2% is considered as substantial.

aa: Bredeweg et al., 2012; Pileggi et al., 2010; Buist et al., 2009; Lun et al., 2004; Taunton et al., 2003; Wen et al., 1998; Fields et al., 1990; Macera et al., 1989; Walter et al., 1989

bb: The inconsistencies among studies complicate inter-study comparisons and prevent us from confirming the relationship between all risk factors and running injuries. In addition, relatively few prospective studies were identified in this review, reducing the overall ability to detect risk factors.

cc: We found a great heterogeneity of statistical methods between studies, which prevented us from performing a meta-analysis. No heterogeneity I^2 tests were performed or provided.

dd: Dannenberg et al., 1996; Olsen et al., 2006; Seil et al., 1998; van Ginckel et al., 2009

ee: Weiss 1985; Tuffery 1989

ff: Methodological information was missing or provided poorly in most studies. Lack of adequate description of population characteristics, sampling method, and participation rate makes it impossible to generalize results to relevant populations.

gg: Due to the heterogeneity across studies in terms of population characteristics, overuse definitions, assessment tools and sampling methods, data pooling and a meta-analysis were not possible. No heterogeneity I^2 tests were

DRAFT review prepared for the WHO Guideline Development Group
FOR CONSULTATION ONLY

DRAFT review prepared for the WHO Guideline Development Group
FOR CONSULTATION ONLY

performed or provided.

hh: Incidence and Prevalence rates ranged considerably across studies.

ii: Thijs et al., 2011; Buist et al., 2010; Buist et al., 2010; Hesar et al., 2009; van Ginckel et al., 2009; van Middelkoop et al., 2008; Thijs et al., 2008; Lun et al., 2004; Taunton et al., 2003; Wen et al., 1998; Macera et al., 1989

jj: McKean et al., 2006; Wen et al., 1997

kk: By our inclusion criteria (e.g. long-distance runners recreational and/or competitive) for selecting the original studies, a broad spectrum in the type of runners (novice, track and field, etc.) was selected. Although we performed an extensive literature search, it is likely that both selection and publication bias influenced the results.

ll: The heterogeneity in study populations, in operationalization of both outcomes and risk factors, and time to follow-up prevented us from following a formal meta-analytical approach. No heterogeneity I^2 tests were performed or provided.

mm: Indirect comparisons = A point of concern is that many of the included studies did not clearly describe the participation rate of the target group, which limits the generalizability of findings.

nn: Bovens et al., 1989; Buist et al., 2010; Nielsen et al., 2013; Malisoux et al., 2015; Wen et al., 1998; Krabak et al., 2011; Lysholm et al., 1987

oo: Bredeweg et al., 2012; Buist et al., 2008; Jakobsen et al., 1994; Theisen et al., 2014; van Mechelen et al., 1993

pp: The definition of injury varies considerably across studies. Second, runners from the included studies were classified into four groups according to the type of runner, enabling relevant intergroup comparison. Third, the method of gathering data on exposure time may be questionable. Further, some studies specified the premise that the same runner was included and was contributing exposure time, if running was resumed after an injury occurrence (e.g. the can contribute two injuries from one individual).

qq: No heterogeneity I^2 tests were performed or provided.

rr: Certainty was downgraded from High to Moderate because of serious risk of bias.

ss: Certainty was downgraded from High to Low because of serious risk of bias and inconsistency.

tt: Certainty was downgraded from High to Very Low because of serious risk of bias, inconsistency, and indirectness.

uu: Certainty was downgraded from High to Low because of serious risk of bias and inconsistency.

vv: Certainty was downgraded from High to Low because of serious risk of bias and inconsistency.

ww: Certainty was downgraded from High to Low because of serious risk of bias and inconsistency.

xx: Certainty was downgraded from High to Low because of serious risk of bias and inconsistency.

yy: Certainty was downgraded from High to Very Low because of serious risk of bias, inconsistency, and indirectness.

zz: Certainty was downgraded from High to Very Low because of serious risk of bias, inconsistency, and indirectness.

aaa: Certainty was downgraded from High to Low because of serious risk of bias and inconsistency.

Osteoarthritis

Population: Adults 18 years of age and older
Exposure: Duration, frequency and/or intensity of LTPA, or a composite score reflecting total volume of LTPA
Comparison: No LTPA, or LTPA of a lesser duration, frequency and/or intensity, or composite score of total volume of LTPA.
Outcome: Adverse health outcomes (especially injury, osteoarthritis, erectile dysfunction, and exposure to pollution)

Certainty assessment							Summary Effect	Certainty	Importance
No of studies	Study design	Risk of bias	Inconsistency	Indirectness	Imprecision	Other considerations			

The association of recreational and competitive running with hip and knee osteoarthritis: a systematic review and meta-analysis (Alentorn-Geli et al., 2017) (11)

DRAFT review prepared for the WHO Guideline Development Group
FOR CONSULTATION ONLY

Certainty assessment							Summary Effect	Certainty	Importance
No of studies	Study design	Risk of bias	Inconsistency	Indirectness	Imprecision	Other considerations			
25	13 case-control ^a 5 cross-sectional ^b 7 prospective cohort ^c	Serious ^d	Serious ^e	Serious ^f	Not serious	None	Adverse health outcomes - Compared with the control group, recreational runners had a significantly lower association with hip and/or knee OA (OR 0.66; 95% CI 0.57-0.76; I² 50%) and knee OA alone (OR 0.72; 95% CI 0.63-0.83; I² 0%) in the overall population and in males (OR 0.78; 95% CI 0.68-0.89; I² 0%; and OR 0.7; 95% CI 0.5-0.97; I² 0% respectively). - Compared with the control group, female recreational runners had a lower association with hip and/or knee OA (OR 0.54; 95% CI 0.41-0.71; I² 43%). - In fact, running at recreational level was even found to have a protective effect on hip and/or knee OA. - It was also not possible to demonstrate the confounding effect of associated risk factors (age, gender, weight, occupational workload and previous injury) on the risk of OA in runners. Dose-response relationship Compared with the control group, individuals with exposure to running of less than 15 years had a lower association with hip and/or knee OA in: - the overall population (OR 0.6; 95% CI 0.49-0.73; I² 47%) - males (OR 0.79; 95% CI 0.68-0.91; I² 0%) - females (OR 0.52; 95% CI 0.47-0.57; I² 0%) Type of LTPA No effects to report; the meta-analysis pooled running and orienteering (e.g. runners that use a topographical map to navigate from point A to point B at speed in unfamiliar terrain) populations at the recreational level, but did not analyze these separately.	Very Low ⁿ	CRITICAL

Running and knee osteoarthritis: a systematic review and meta-analysis (Timmins et al., 2017) (22)

DRAFT review prepared for the WHO Guideline Development Group
FOR CONSULTATION ONLY

Certainty assessment							Summary Effect	Certainty	Importance
No of studies	Study design	Risk of bias	Inconsistency	Indirectness	Imprecision	Other considerations			
10	4 case-control ^g 1 ambispective cohort ^h 2 prospective cohort ⁱ 3 retrospective cohort ^j	Serious ^k	Serious ^l	Serious ^m	Not serious	None	Adverse health outcomes The combined odds ratio of undergoing knee surgery due to OA was 0.46 (95% CI 0.30-0.71) in runners or orienteers when compared with non-runners. The I² was 0%, with 95% CI 0% to 73%. From this evidence, it is not possible to conclude whether running was associated with a diagnosis of knee OA, and studies offered differing conclusions.^p Dose-response relationship No effects to report; this review only looked at the adverse health outcomes relationship for OA. Type of LTPA No effects to report; the meta-analysis pooled running and orienteering populations, but did not analyze these separately.	Very Low ^o	CRITICAL

PA = physical activity; OA = osteoarthritis; OR = odds ratio; CI = confidence interval

a: Oahaghin et al., 2009; Kettunen et al., 1999; Kohatsu and Schurman, 1990; Lau et al., 2000; Lo et al., 2016; Manninen et al., 2001; Marti et al., 1989; Puranen et al., 1975; Sohn and Micheli, 1985; Spector et al., 1996; Wingard et al., 1993; Wingard et al., 1998; Vrezas et al., 2010

b: Konradsen et al., 1990; Lane et al., 1986; Panush et al., 1986; Puranen et al., 1975; Williams 2013

c: Chakravarty et al., 2008; Cheng et al., 2000; Kujala et al., 1994; Kujala et al., 1999; Lane et al., 1993; Lane et al., 1998; Panush et al., 1995

d: The assessment of the risk of bias was conducted using a tool not specifically designed for observational, etiologic association studies and the use of other appraisal tools might therefore provide different insights. 16/25 studies scored High risk on the type of bias detection.

e: Due to high between-studies heterogeneity (high I² statistic), the random-effects model, which can inappropriately weight smaller studies in some instances, was necessary. 30-60% (as captured in summary of effects) may be considered moderate.

f: In some studies, the runners were also exposed to other types of sport (i.e. tennis), the runners included some individuals performing only walking exercise, or involved orienteering running.

g: Kohatsu, 1990; Sandmark, 1999; Manninen, 2001; Thelin, 2006

h: Panush, 1986

i: Lane, 1986; Felson, 2007

j: de Carvalho, 1977; Muhlbauer, 2000; Mosher, 2010

k: Gray literature was not included in the eligibility criteria. As a result, the findings of this review may reflect publication bias. The meta-analysis included only a small number of studies, with odds ratios that represent unadjusted proportions (i.e., odds were not adjusted for confounding factors). Given the nature of observational studies, only low- to moderate-quality evidence could be expected. However, the assessment of potential bias undertaken in this review indicated that many studies would be downgraded to low or very low quality.

l: Although the I² indicated low heterogeneity, the upper 95% CI of the I² is high (73%), and the pooled estimate should be interpreted with caution.

m: The populations under investigation are not the same. The outcomes are differently defined in these studies.

n: Certainty rated from High to Very Low because of serious risk of bias, inconsistency, and indirectness.

o: Certainty rated from High to Very Low because of serious risk of bias, inconsistency, and indirectness.

p: The review considered 5 different outcomes related to knee OA: diagnosis of knee OA, radiographic and imaging markers, arthroplasty for knee OA, knee pain, and knee-associated disability. Due to heterogeneity of outcome definition and measurement of studies, only 1 meta-analysis was appropriate: This combined the case-control studies that identified cases of knee surgery due to OA. Hence overall, there was no conclusion to be made on the relationship between running and knee OA as a long-term adverse health outcome.

DRAFT review prepared for the WHO Guideline Development Group
FOR CONSULTATION ONLY

Erectile Dysfunction

Population: Adults 18 years of age and older
Exposure: Duration, frequency and/or intensity of LTPA, or a composite score reflecting total volume of LTPA
Comparison: No LTPA, or LTPA of a lesser duration, frequency and/or intensity, or composite score of total volume of LTPA.
Outcome: Adverse health outcomes (especially injury, osteoarthritis, erectile dysfunction, and exposure to pollution)

Certainty assessment							Summary Effect	Certainty	Importance
No of studies	Study design	Risk of bias	Inconsistency	Indirectness	Imprecision	Other considerations			

Bicycle riding and erectile dysfunction: a review (Sommer et al., 2010) (21)

							Due to Critically Low rating on the AMSTAR 2 scale, this systematic review did NOT qualify for assessing the GRADE evidence.		HIGH
--	--	--	--	--	--	--	--	--	------

LTPA = leisure-time physical activity

Exposure to Air Pollution

Population: Adults 18 years of age and older
Exposure: Duration, frequency and/or intensity of LTPA, or a composite score reflecting total volume of LTPA
Comparison: No LTPA, or LTPA of a lesser duration, frequency and/or intensity, or composite score of total volume of LTPA.
Outcome: Adverse health outcomes (especially injury, osteoarthritis, erectile dysfunction, and exposure to air pollution)

Certainty assessment							Summary Effect	Certainty	Importance
No of studies	Study design	Risk of bias	Inconsistency	Indirectness	Imprecision	Other considerations			

N/A

							No Systematic Reviews were found in relation to the Exposure to air Pollution outcome for assessing the GRADE evidence.		LOW
--	--	--	--	--	--	--	---	--	-----

LTPA = leisure-time physical activity

References

1. Organization WH. Global recommendations on physical activity for health. 2010. 2015.
2. Craike M, Hibbins R, Cuskelly G. The influence of various aspects of enjoyment on participation in leisure time physical activity. *World Leisure Journal*. 2010;52.
3. Moher D, Liberati A, Tetzlaff J, Altman DG. Preferred reporting items for systematic reviews and meta-analyses: the PRISMA statement. *Annals of internal medicine*. 2009;151(4):264-9.
4. Ouzzani M, Hammady H, Fedorowicz Z, Elmagarmid A. Rayyan—a web and mobile app for systematic reviews. *Systematic reviews*. 2016;5(1):210.
5. Shea BJ, Reeves BC, Wells G, Thuku M, Hamel C, Moran J, et al. AMSTAR 2: a critical appraisal tool for systematic reviews that include randomised or non-randomised studies of healthcare interventions, or both. *Bmj*. 2017;358:j4008.
6. Guyatt GH, Oxman AD, Vist GE, Kunz R, Falck-Ytter Y, Alonso-Coello P, et al. GRADE: an emerging consensus on rating quality of evidence and strength of recommendations. *Bmj*. 2008;336(7650):924-6.
7. Huguet A, Hayden JA, Stinson J, McGrath PJ, Chambers CT, Tougas ME, et al. Judging the quality of evidence in reviews of prognostic factor research: adapting the GRADE framework. *Systematic reviews*. 2013;2(1):71.
8. van der Molen HF, Foresti C, Daams JG, Frings-Dresen MH, Kuijer PPF. Work-related risk factors for specific shoulder disorders: a systematic review and meta-analysis. *Occupational and environmental medicine*. 2017;74(10):745-55.
9. Schünemann H. The GRADE handbook: Cochrane Collaboration; 2013.
10. Guyatt GH, Oxman AD, Kunz R, Vist GE, Falck-Ytter Y, Schünemann HJ. What is “quality of evidence” and why is it important to clinicians? *Bmj*. 2008;336(7651):995-8.
11. Alentorn-Geli E, Samuelsson K, Musahl V, Green CL, Bhandari M, Karlsson JON. The Association of Recreational and Competitive Running With Hip and Knee Osteoarthritis: A Systematic Review and Meta-analysis. *Journal of Orthopaedic & Sports Physical Therapy*. 2017;47(6):373-90.
12. Borel WP, Filho JE, Diz JBM, Moreira PF, Veras PM, Catharino LL, et al. Prevalence of injuries in Brazilian Recreational street runners: meta-analysis. / Prevalencia de lesiones en corredores de calle aficionados Brasileños: META-ANÁLISIS. *Revista Brasileira de Medicina do Esporte*. 2019;25(2):161-7.
13. Damsted C, Nielsen RO, Sørensen H, Glad S, Malisoux L. Is there evidence for an association between changes in training load and running-related Injuries? A Systematic Review. *International Journal of Sports Physical Therapy*. 2018;13(6):931-42.
14. Johnston R, Cahalan R, O'Keeffe M, O'Sullivan K, Comyns T. The associations between training load and baseline characteristics on musculoskeletal injury and pain in endurance sport populations: A systematic review. *Journal of science and medicine in sport*. 2018;21(9):910-8.
15. Lopes AD, Hespanhol LC, Yeung SS, Costa LOP. What are the main running-related musculoskeletal injuries? *Sports medicine*. 2012;42(10):891-905.
16. Neubauer T, Brand J, Lidder S, Krawany M. Stress fractures of the femoral neck in runners: a review. *Research in Sports Medicine*. 2016;24(3):283-97.
17. Oestergaard Nielsen R, Buist I, Sørensen H, Lind M, Rasmussen S. Training errors and running related injuries: A Systematic Review. *International Journal of Sports Physical Therapy*. 2012;7(1):58-75.
18. Qu X, Zhang X, Zhai Z, Li H, Liu X, Li H, et al. Association Between Physical Activity and Risk of Fracture. *Journal of Bone & Mineral Research*. 2014;29(1):202-11.
19. Saragiotto B, Yamato T, Hespanhol Junior L, Rainbow M, Davis I, Lopes A, et al. What are the Main Risk Factors for Running-Related Injuries? *Sports Medicine*. 2014;44(8):1153-63.
20. Sobhani S, Dekker R, Postema K, Dijkstra PU. Epidemiology of ankle and foot overuse injuries in sports: A systematic review. *Scandinavian Journal of Medicine & Science in Sports*. 2013;23(6):669-86.

21. Sommer F, Goldstein I, Korda JB. Bicycle riding and erectile dysfunction: a review. *The journal of sexual medicine*. 2010;7(7):2346-58.
22. Timmins KA, Leech RD, Batt ME, Edwards KL. Running and Knee Osteoarthritis: A Systematic Review and Meta-analysis. *The American journal of sports medicine*. 2017;45(6):1447-57.
23. Tonoli DC, Cumps E, Aerts I, Verhagen E, Meeusen R. Incidence, risk factors and prevention of running related injuries in long-distance running: a systematic review. *Sport & Geneeskunde*. 2010;43(5):12-8.
24. Van Der Worp MP, Ten Haaf DSM, Van Cingel R, De Wijer A, S N-VD, en MWG, et al. Injuries in runners; a systematic review on risk factors and sex differences. *PLoS ONE*. 2015;10(2).
25. Videbæk S, Bueno AM, Nielsen RO, Rasmussen S. Incidence of Running-Related Injuries Per 1000 h of running in Different Types of Runners: A Systematic Review and Meta-Analysis. *Sports medicine (Auckland, NZ)*. 2015;45(7):1017-26.
26. Lopes AD, Costa LOP, Saragiotto BT, Yamato TP, Adami F, Verhagen E. Musculoskeletal pain is prevalent among recreational runners who are about to compete: an observational study of 1049 runners. *Journal of physiotherapy*. 2011;57(3):179-82.
27. Malisoux L, Nielsen RO, Urhausen A, Theisen D. A step towards understanding the mechanisms of running-related injuries. *Journal of science and medicine in sport*. 2015;18(5):523-8.
28. Van Middelkoop M, Van Linschoten R, Berger MY, Koes BW, Bierma-Zeinstra SM. Knee complaints seen in general practice: active sport participants versus non-sport participants. *BMC musculoskeletal disorders*. 2008;9(1):36.

Supplementary file 3: Search strategies.

Search strategy in PubMed (read from bottom-up).

No.	Query
#5	#1 AND #2 AND #3 AND #4
#4	("systematic review"[Publication Type] OR "meta-analysis"[Publication Type] OR "systematic literature review"[Publication Type] OR "umbrella review"[Publication Type] OR "review of reviews"[Publication Type] OR "pooled analysis"[Publication Type])
#3	("2009"[Date - Publication] : "3000"[Date - Publication])
#2	("physical activity"[Title] OR "physical exercise"[Title] OR "exercis*" [Title] OR "sport*" [Title] OR "training" [Title] OR "walking"[Title] OR "intensity activity"[Title] OR "resistance training"[Title] OR "aerobic activity"[Title] OR "physically active"[Title] OR "active transportation"[Title] OR "runn*" [Title] OR "cycling"[Title] OR "bicycling"[Title] OR "leisure-time"[Title] OR "fitness"[Title])
#1	("exercise-induced"[Title] OR "induced by exercise"[Title] OR "due to exercise"[Title] OR "splenosis" [Title] OR "Myoglobinuria"[Title] OR "hernia"[Title] OR "rupture"[Title] OR "fracture*" [Title] OR "Rhabdomyolysis"[Title] OR "injur* Shaken Baby Syndrome"[Title] OR "Cerebrospinal Fluid Ot*" [Title] OR "rhea"[Title] OR "Cerebral Hem"[Title] OR "Brain Hem*" [Title] OR "rhage"[Title] OR "Craniocerebral Trauma"[Title] OR "Decompression Sickness" [Title] OR "Asphyxia" [Title] OR "Frostbite"[Title] OR "Chilblains"[Title] OR "Heat Exhaustion"[Title] OR "Heat Stroke"[Title] OR "Heat Strain"[Title] OR "Sunstroke"[Title] OR "Tendinopathy"[Title] OR "Concussion*" [Title] OR "Hematom*" [Title] OR "Contusion*" [Title] OR "Wound*" [Title] OR "Osteoarthritis"[Title] OR "laceration*" [Title] OR "sprain"[Title] OR "strain"[Title] OR "fall"[Title] OR "Injur*" [Title] OR "overuse"[Title] OR "dislocation"[Title] OR "health risk"[Title] OR "erectile dysfunction"[Title] OR "pollution"[Title] OR "head injuries"[Title] OR "sexual dysfunction"[Title] OR "cycling-related"[Title] OR "urogenital"[Title] OR "collision*" [Title] OR "crash*" [Title])

Search strategy in Embase (read from bottom-up).

No.	Query
#5	#1 AND #2 AND #3 AND #4
#4	[2009-2020]/py
#3	[cochrane review]/lim OR [systematic review]/lim OR [meta analysis]/lim
#2	'physical activity':ti OR 'physical exercise':ti OR 'exercis*':ti OR 'sport*':ti OR 'training':ti OR 'walking':ti OR 'intensity activity':ti OR 'resistance training':ti OR 'aerobic activity':ti OR 'physically active':ti OR 'active transportation':ti OR 'runn*':ti OR 'cycling':ti OR 'bicycling':ti OR 'leisure-time':ti OR 'fitness':ti
#1	'exercise-induced':ti OR 'induced by exercise':ti OR 'due to exercise':ti OR splenosis:ti OR 'myoglobinuria':ti OR 'hernia':ti OR 'rupture':ti OR 'fracture*':ti OR 'rhabdomyolysis':ti OR 'injur* shaken baby syndrome':ti OR 'cerebrospinal fluid ot*':ti OR 'rhea':ti OR 'cerebral hem':ti OR 'brain hem*':ti OR 'rhage':ti OR 'craniocerebral trauma':ti OR 'decompression sickness':ti OR asphyxia:ti OR 'frostbite':ti OR 'chilblains':ti OR 'heat exhaustion':ti OR 'heat stroke':ti OR 'heat strain':ti OR 'sunstroke':ti OR 'tendinopathy':ti OR 'concussion*':ti OR 'hematom*':ti OR 'contusion*':ti OR 'wound*':ti OR 'osteoarthritis':ti OR 'laceration*':ti OR 'sprain':ti OR 'strain':ti OR 'fall':ti OR 'injur*':ti OR 'overuse':ti OR 'dislocation':ti OR 'health risk':ti OR 'erectile dysfunction':ti OR 'pollution':ti OR 'head injuries':ti OR 'sexual dysfunction':ti OR 'cycling-related':ti OR 'urogenital':ti OR 'collision*':ti OR 'crash*':ti

Search strategy in SPORTDiscus (read from bottom-up).

No.	Query
#5	#1 AND #2 AND #3 AND #4
#4	Published Date: 20090101-20201231
#3	("systematic review" OR "meta-analysis" OR "systematic literature review" OR "umbrella review" OR "review of reviews" OR "pooled analysis")
#2	TI ("physical activity" OR "physical exercise" OR "exercis*" OR "sport*" OR "training " OR "walking" OR "intensity activity" OR "resistance training" OR "aerobic activity" OR "physically active" OR "active transportation" OR "runn*" OR "cycling" OR "bicycling" OR "leisure-time" OR "fitness")
#1	"TI ("exercise-induced" OR "induced by exercise" OR "due to exercise" OR "splenosis" OR "Myoglobinuria" OR "hernia" OR "rupture" OR "fracture*" OR "Rhabdomyolysis" OR "injur* Shaken Baby Syndrome" OR "Cerebrospinal Fluid Ot*" OR "rhea" OR "Cerebral Hem" OR "Brain Hem*" OR "rhage" OR "Craniocerebral Trauma" OR "Decompression Sickness" OR "Asphyxia" OR "Frostbite" OR "Chilblains" OR "Heat Exhaustion" OR "Heat Stroke" OR "Heat Strain" OR "Sunstroke" OR "Tendinopathy" OR "Concussion*" OR "Hematom*" OR "Contusion*" OR "Wound*" OR "Osteoarthritis" OR "laceration*" OR "sprain" OR "strain" OR "fall" OR "Injur*" OR "overuse" OR "dislocation" OR "health risk" OR "erectile dysfunction" OR "pollution" OR "head injuries" OR "sexual dysfunction" OR "cycling-related" OR "urogenital" OR "collision*" OR "crash*")

Supplementary file 4: Excluded articles after full text screening.

Reference (first author, year and title)	Reason
Bosomworth. 2010. Exercise and knee osteoarthritis: Benefit or hazard?	No measure of PA Volume
Abrahams et al. 2014. Risk factors for sports concussion: an evidence-based systematic review	No measure of PA Volume
Alentorn-Geli et al. 2014. Prevention of anterior cruciate ligament injuries in sports-Part I: Systematic review of risk factors in male athletes	No measure of PA Volume
Bahadur. 2017. Does sport and exercise cause osteoarthritis?	No Full text available – abstract only
Baran et al. 2014. Cycling-related sexual dysfunction in men and women: A review	No measure of PA Volume
Bell et al. 2018. Sport specialization and risk of overuse injuries: A systematic review with meta-analysis	Wrong population
Booth et al. 2018. The effect of training loads on performance measures and injury characteristics in rugby league players: A systematic review	No measure of PA Volume
Brindle et al. 2019. Association between spatial temporal parameters and overuse injury history in runners: A systematic review and meta-analysis	No measure of PA Volume
Chan et al. 2016. Epidemiology of sport-related spinal cord injuries: a systematic review	No measure of PA Volume
Cheron et al. 2017. Association between sports type and overuse injuries of extremities in adults: a systematic review	No measure of PA Volume
Clark et al. 2016. Exercise-induced stress behavior, gut-microbiota-brain axis and diet: a systematic review for athletes	No measure of PA Volume
Costa et al. 2017. Systematic review: exercise-induced gastrointestinal syndrome – implications for health and intestinal disease	Wrong outcome
De Souto Barreto et al. 2019. Association of long-term exercise training with risk of falls, fractures, hospitalizations, and mortality in older adults: a systematic review and meta-analysis	Wrong population
Drew et al. 2015. Moderate evidence for the relationship between training or competition load and injury: a systematic review	Duplicate
Drew et al. 2016. The relationship between training load and injury, illness and soreness: a systematic and literature review	Wrong population
Driban et al. 2011. Association between sports participation and the risk of knee osteoarthritis: a systematic review	No measure of PA Volume

Driban et al. 2017. Is participation in certain sports associated with knee osteoarthritis? A systematic review	Wrong population
Duncan et al. 2016. A systematic review to evaluate exercise for anterior cruciate ligament injuries: does this approach reduce the incidence of knee osteoarthritis?	Wrong population
Eckard et al. 2018. The relationship between training load and injury in athletes: a systematic review	Wrong population
Emery et al. 2012. Are joint injury, obesity, physical activity, sport, and occupational risk factors factors for osteoarthritis? A systematic review of the literature	Duplicate
Francis et al. 2019. The proportion of lower limb running injuries by gender, anatomical location and specific pathology: a systematic review	Wrong population
Fuller et al. 2019. Ballet and contemporary dance injuries when transitioning to full-time training or professional level dance: a systematic review	Wrong population
Green et al. 2017. Calf muscle strain injuries in sport: a systematic review of risk factors for injury	No measure of PA Volume
Griffin et al. 2019. The association between the Acute:Chronic Workload Ratio and injury and its application in team sports: a systematic review	Wrong population
Hutton et al. 2016. Catastrophic Cervical spine injuries in contact sports	Wrong population
Jones et al. 2017. Training Load and fatigue marker associations with injury and illness: a systematic review of longitudinal studies	No measure of PA Volume
Jones et al. 2018. Incidence, diagnosis, and management of injury in sport climbing and bouldering: a critical review	No measure of PA Volume
Kemmler et al. 2013. Effects of exercise on fracture reduction in older adults: a systematic review and meta-analysis	Wrong outcome
Kim et al. 2016. Meta-analysis of prognostic implications of exercise-induced ventricular premature complexes in the general population	No measure of PA Volume
King et al. 2014. The incidence of match and training injuries in rugby league: a pooled data analysis of published studies	No measure of PA Volume
King et al. 2019. Match and training injuries in women's rugby union: a systematic review of published studies	Wrong population
Kluitenberg et al. 2015. What are the differences in injury proportions between different populations of runners? A systematic review and meta-analysis	No measure of PA Volume

Knapik et al. 2018. Prevalence and management of coracoid fracture sustained during sporting activities and time to return to sport: a systematic review	No measure of PA Volume
Ko. 2019. Long-term exercise training in older adults is associated with reduced injurious falls and fractures	Wrong publication type
Orchard. 2015. Men at higher risk of groin injuries in elite team sports: a systematic review	Wrong population
Panagodage Perera et al. 2018. Epidemiology of injuries in women playing competitive team bat-or-stick sports: a systematic review and meta-analysis	Wrong population
Pires et al. 2017. Association between exercise-induced hyperthermia and intestinal permeability: a systematic review	Wrong outcome
Polmann et al. 2019. Prevalence of dentofacial injuries among combat sports participants: a systematic review and meta-analysis	No measure of PA Volume
Prien et al. 2018. Epidemiology of head injuries focusing on concussions in team contact sports: a systematic review	Wrong population
Richmond et al. 2013. Are joint injury, sport activity, physical activity, obesity, or occupational activities predictors for osteoarthritis? A systematic review	No measure of PA Volume
Sakaguchi et al. 2019. Metabolomics-based studies assessing exercise-induced alterations of the human metabolome: a systematic review	Wrong outcome
Schultz et al. 2013. Exercise-induced hypertension, cardiovascular events, and mortality in patients undergoing exercise stress testing: a systematic review and meta-analysis	Wrong outcome
Stege et al. 2014. Effect of interventions on potential, modifiable risk factors for knee injury in team ball sports: a systematic review	Wrong outcome
Sugimoto et al. 2016. Critical components of neuromuscular training to reduce ACL injury risk in female athletes: meta-regression analysis	Wrong outcome
Vigdorchik et al. 2017. What is the association of elite sporting activities with the development of hip osteoarthritis?	Wrong population
Visentini. 2017. A systematic review of parameters related to cycling overuse injuries or pain	No Full text available – abstract only
Weiler et al. 2016. Exercise-induced bronchoconstriction update-2016	No measure of PA Volume
Mueller et al. 2015. Health impact assessment of active transportation: a systematic review	Wrong outcome
Hamstra-Wright et al. 2014. Common running-related injuries: methodology and prevention: evidence-based report	Wrong publication type