Table B.1.h. Adverse events

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.

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.

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)

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 *(78)*) 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 *(79)*.

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 *(80)*. 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).

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 *(80)*. 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.

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) *(81)*. 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 cut-off 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.