Supplementary Materials

Nature exposure and mental health during the COVID-19 pandemic: A Navigation Guide systematic review with meta-analysis

Muhammad Mainuddin Patwary, MSc ; Mondira Bardhan, MSc; Asma Safia Disha, MSc; Angel M. Dzhambov, PhD; Colby Parkinson, MSc; Matthew H E M Browning, PhD; S M Labib, PhD; Lincoln R Larson, PhD; Md Zahidul Haque, MSc; Md Atiqur Rahman; Md Ashraful Alam, PhD; Md Faysal Tareq, MSc; Faysal Kabir Shuvo, PhD

TABLE OF CONTENTS

S1. SEARCH STRATEGY DETAILS

Last updated April 30, 2022

PubMed

1. **Related to nature contact: 663,089**

“Nature contact”[tw] OR “greenness”[tw] OR “greenery”[tw] OR “bluespace*”[tw] OR “blue space*”[tw] OR “lakes”[MeSH] OR “lake*”[tw] OR “rivers”[MeSH] OR “river*” OR "water resources"[MeSH] OR "water resources"[tw] OR “water body”[tw] OR “wetlands”[MeSH] OR "wetland*"[tw] OR “ponds”[MeSH]" OR “pond*”[tw] OR “coast*”[tw] OR “beach*”[tw] OR “Parks, Recreational”[MeSH] OR “recreational parks”[tw] OR “park*”[tw] OR “recreation”[MeSH] OR “recreation*”[tw] OR "green space*"[tw] OR “greenspace*[tw] OR "green area*"[tw] OR "greenway*"[tw] OR "green belt*"[tw] OR "green corridor*"[tw] OR “indoor environment*”[tw] OR “natural environment*"[tw] OR “outdoors”[tw] OR "open space*"[tw] OR “naturalness”[tw] OR “gardens”[MeSH] OR “garden*”[tw] OR “playground*”[tw] OR “canopy”[tw] OR “trees”[MeSH] OR “tree*” OR “forests”[MeSH] OR “forest*”[tw] OR “woodland*”[tw] OR "green roof*"[tw] OR "roof garden*"[tw] OR “arboretum”[tw] OR "urban nature"[tw] OR “protected area”[tw] OR “vegetati*”[tw] OR “green land cover*”[tw] OR “biodiversity”[MeSH] OR “biodiversity”[tw] OR “Normalized Difference Vegetation Index”[tw] OR “NDVI”[tw] OR “Leaf Area Index”[tw] OR “LAI”[tw] OR “horticultural therapy”[MeSH] OR “horticultural therapy”[tw] OR “sunlight”[MeSH] OR “sunlight”[tw] OR "nature-based"[tw] OR “ecosystem”[MeSH] OR “ecosystem”[tw]

2. **Related to human health/exposure: 3,069,621**

“wellbeing”[tw] OR "well-being"[tw] OR “wellness”[tw] OR "emotional health"[tw] OR "psychological health"[tw] OR "mental health"[MeSH] OR "mental health"[tw] OR “mental disorders”[MeSH] OR “mental disorder*”[tw] OR “mood disorders”[MeSH] OR “mood disorder*”[tw] OR “depression”[MeSH] OR “depression”[tw] OR “worry”[tw] OR “fear”[MeSH] OR “fear”[tw] OR “anxiety”[MeSH] OR “anxiety”[tw] OR “anxiety disorders”[MeSH] OR “anxiety disorders”[tw] OR “Stress Disorders, Traumatic, Acute”[MeSH] OR “stress”[tw] OR “Stress disorder*”[tw] OR “perceived stress”[tw] OR “irritable mood”[MeSH] OR “irritability”[tw] OR “happiness”[MeSH] OR “happiness”[tw] OR “psychological disturbance”[tw] OR “psychological disorder”[tw] OR "sleep wake disorders"[MeSH] OR “sleep disturbance” OR “loneliness”[MeSH] OR “loneliness”[tw] OR “boredom”[MeSH] OR “boredom”[tw] OR "sleep initiation and maintenance disorders"[MeSH] OR “insomnia”[tw] OR “Stress Disorders, Post-Traumatic”[MeSH] OR “post-traumatic disorder”[tw] OR “trauma”[tw] OR “psychological distress”[MeSH] OR “psychological distress”[tw] OR "stress, psychological"[MeSH] OR “psychological stress”[tw] OR “Social isolation”[MeSH] OR “Social isolation”[tw]

3. **Related to coronavirus: 215,325**

“COVID-19”[MeSH] OR “COVID-19”[tw] OR “SARS-COV-2”[tw] OR “2019-ncov”[tw] OR “2019 novel coronavirus”[tw] OR “novel coronavirus”[tw] OR “nCoV”[tw] OR "coronavirus"[MeSH] OR “Coronavirus”[tw] OR "pandemics"[MeSH] OR “pandemic*”[tw]

#1 AND #2 AND #3

Total = 447

Web of Science

1. Related to nature contact: (1,112,706)
TI=(“Nature contact” OR “greenness” OR “greenery” OR “bluespace*” OR “blue space*” OR “lake*” OR “river*” OR "water resources" OR “water body” OR "wetland*" OR “pond*” OR “coast*” OR “beach*” OR “recreational parks” OR “park*” OR “recreation*” OR "green space*" OR “greenspace* OR "green area*" OR "greenway*" OR "green belt*" OR "green corridor*" OR “indoor environment*” OR “natural environment*" OR “outdoors” OR "open space*" OR “naturalness” OR “garden*” OR “playground*” OR “canopy” OR “tree*” OR “forest*” OR “woodland*” OR "green roof*" OR "roof garden*" OR “arboretum” OR "urban nature" OR “protected area” OR “vegetati*” OR “green land cover*” OR “biodiversity” OR “Normalized Difference Vegetation Index” OR “NDVI” OR “Leaf Area Index” OR “LAI” OR “horticultural therapy” OR “sunlight” OR "nature-based" OR “ecosystem”) OR TS=(“Nature contact” OR “greenness” OR “greenery” OR “bluespace*” OR “blue space*” OR “lake*” OR “river*” OR "water resources" OR “water body” OR "wetland*" OR “pond*” OR “coast*” OR “beach*” OR “recreational parks” OR “park*” OR “recreation*” OR "green space*" OR “greenspace* OR "green area*" OR "greenway*" OR "green belt*" OR "green corridor*" OR “indoor environment*” OR “natural environment*" OR “outdoors” OR "open space*" OR “naturalness” OR “garden*” OR “playground*” OR “canopy” OR “tree*” OR “forest*” OR “woodland*” OR "green roof*" OR "roof garden*" OR “arboretum” OR "urban nature" OR “protected area” OR “vegetati*” OR “green land cover*” OR “biodiversity” OR “Normalized Difference Vegetation Index” OR “NDVI” OR “Leaf Area Index” OR “LAI” OR “horticultural therapy” OR “sunlight” OR "nature-based" OR “ecosystem”)

2. Related to human health/exposure: (3,290,369)
TI=(“wellbeing” OR "well-being" OR “wellness” OR "emotional health" OR "psychological health" OR "mental health" OR “mental disorder*” OR “mood disorder*” OR “depression” OR “worry” OR “fear” OR “anxiety” OR “anxiety disorders” OR “stress” OR “Stress disorder*” OR “perceived stress” OR “irritable mood” OR “irritability” OR “happiness” OR “psychological disturbance” OR “psychological disorder” OR “sleep disturbance” OR “loneliness” OR “boredom” OR “insomnia” OR “post-traumatic disorder” OR “trauma” OR “psychological distress” OR “psychological stress” OR “Social isolation”) OR TS=(“wellbeing” OR "well-being" OR “wellness” OR "emotional health" OR "psychological health" OR "mental health" OR “mental disorder*” OR “mood disorder*” OR “depression” OR “worry” OR “fear” OR “anxiety” OR “anxiety disorders” OR “stress” OR “Stress disorder*” OR “perceived stress” OR “irritable mood” OR “irritability” OR “happiness” OR “psychological disturbance” OR “psychological disorder” OR “sleep disturbance” OR “loneliness” OR “boredom” OR “insomnia” OR “post-traumatic disorder” OR “trauma” OR “psychological distress” OR “psychological stress” OR “Social isolation”)

3. Related to coronavirus: (226,460)
TI=(“COVID-19” OR “SARS-COV-2” OR “2019-ncov” OR “2019 novel coronavirus” OR “novel coronavirus” OR “nCoV” OR “Coronavirus” OR “pandemic*”) OR TS=(“COVID-19” OR “SARS-COV-2” OR “2019-ncov” OR “2019 novel coronavirus” OR “novel coronavirus” OR “nCoV” OR “Coronavirus” OR “pandemic*”)

#1 AND #2 AND #3
Total = 305

Scopus

1. Related to nature contact: 3,715,934
TITLE-ABS-KEY(“Nature contact” OR “greenness” OR “greenery” OR “bluespace*” OR “blue space*” OR “lake*” OR “river*” OR "water resources" OR “water body” OR "wetland*" OR “pond*” OR “coast*” OR “beach*” OR “recreational parks” OR “park*” OR “recreation*” OR "green space*" OR “greenspace*” OR "green area*" OR "greenway*" OR "green belt*" OR "green corridor*" OR “indoor environment*” OR “natural environment*" OR “outdoors” OR "open space*" OR “naturalness” OR “garden*” OR “playground*” OR “canopy” OR “tree*” OR “forest*” OR “woodland*” OR "green roof*" OR "roof garden*" OR “arboretum” OR "urban nature" OR “protected area” OR “vegetati*” OR “green land cover*” OR “biodiversity” OR “Normalized Difference Vegetation Index” OR “NDVI” OR “Leaf Area Index” OR “LAI” OR “horticultural therapy” OR “sunlight” OR "nature-based" OR “ecosystem”)

2. Related to human health/exposure: 4,698,599

TITLE-ABS-KEY(“wellbeing” OR "well-being" OR “wellness” OR "emotional health" OR "psychological health" OR "mental health" OR “mental disorder*” OR “mood disorder*” OR “depression” OR “worry” OR “fear” OR “anxiety” OR “anxiety disorders” OR “stress” OR “Stress disorder*” OR “perceived stress” OR “irritable mood” OR “irritability” OR “happiness” OR “psychological disturbance” OR “psychological disorder” OR “sleep disturbance” OR “loneliness” OR “boredom” OR “insomnia” OR “post-traumatic disorder” OR “trauma” OR “psychological distress” OR “psychological stress” OR “Social isolation”)

3. Related to coronavirus: 271,700
TITLE-ABS-KEY(“COVID-19” OR “SARS-COV-2” OR “2019-ncov” OR “2019 novel coronavirus” OR “novel coronavirus” OR “nCoV” OR “Coronavirus” OR “pandemic*”)

#1 AND #2 AND #3
Total = 1782

CINAHL

1. Related to nature contact: “Nature contact” OR “greenness” OR “greenery” OR “bluespace*” OR “blue space*” OR “lake*” OR “river*” OR "water resources" OR “water body” OR "wetland*" OR “pond*” OR “coast*” OR “beach*” OR “recreational parks” OR “park*” OR MH “recreation” OR “recreation*” OR "green space*" OR “greenspace* OR "green area*" OR "greenway*" OR "green belt*" OR "green corridor*" OR “indoor environment*” OR “natural environment*" OR “outdoors” OR "open space*" OR “naturalness” OR “garden*” OR “playground*” OR “canopy” OR “tree*” OR “forest*”[tw] OR “woodland*” OR "green roof*" OR "roof garden*" OR “arboretum” OR "urban nature" OR “protected area” OR “vegetati*” OR “green land cover*” OR “biodiversity” OR “Normalized Difference Vegetation Index” OR “NDVI” OR “Leaf Area Index” OR [MB1] “LAI” OR “horticultural therapy” OR MH “sunlight” OR “sunlight” OR "nature-based" OR MH “ecosystem” OR “ecosystem”

2. Related to human health/exposure: “wellbeing” OR "well-being" OR “wellness” OR "emotional health" OR "psychological health" OR MH "mental health” OR "mental health" OR MH “mental disorders” OR “mental disorder*” OR “mood disorder*” OR MH “depression” OR “depression” OR “worry” OR MH “fear” OR “fear” OR MH “anxiety” OR “anxiety” OR MH “anxiety disorders” OR “anxiety disorders” OR MH “social anxiety disorders” OR “social anxiety disorders” OR MH “Stress Disorders, Post-Traumatic” OR MH “stress” OR “stress” OR “Stress disorder*” OR “perceived stress” OR “irritability” OR MH “happiness” OR “happiness” OR “psychological disturbance” OR “psychological disorder” OR MH "sleep disorders" OR “sleep disturbance” OR MH “loneliness” OR “loneliness” OR “boredom” OR MH "insomnia" OR “insomnia” OR “post-traumatic disorder” OR MH “trauma” OR “trauma” OR MH “psychological distress” OR “psychological distress” OR MH "stress, psychological” OR “psychological stress” OR MH “Social isolation” OR “Social isolation”

3. Related to coronavirus: MH “COVID-19” OR “COVID-19” OR MH “SARS-COV-2” OR “SARS-COV-2” OR “2019-ncov” OR “2019 novel coronavirus” OR “novel coronavirus” OR “nCoV” OR MH "coronavirus" OR “Coronavirus” OR MH "COVID-19 pandemic" OR “COVID-19 pandemic*”

#1 AND #2 AND #3
Total = 1459

PsychInfo

NOFT(“Nature contact” OR “greenness” OR “greenery” OR “bluespace*” OR “blue space*” OR “lake*” OR “river*” OR "water resources" OR “water body” OR "wetland*" OR “pond*” OR “coast*” OR “beach*” OR “recreational parks” OR “park*” OR “recreation*” OR "green space*" OR greenspace* OR "green area*" OR "greenway*" OR "green belt*" OR "green corridor*" OR “indoor environment*” OR “natural environment*" OR “outdoors” OR "open space*" OR “naturalness” OR “garden*” OR “playground*” OR “canopy” OR “tree*” OR “forest*” OR “woodland*” OR "green roof*" OR "roof garden*" OR “arboretum” OR "urban nature" OR “protected area” OR “vegetati*” OR “green land cover*” OR “biodiversity” OR “Normalized Difference Vegetation Index” OR “NDVI” OR “Leaf Area Index” OR “LAI” OR “horticultural therapy” OR “sunlight” OR "nature-based" OR “ecosystem”) AND NOFT(“wellbeing” OR "well-being" OR “wellness” OR "emotional health" OR "psychological health" OR "mental health" OR “mental disorder*” OR “mood disorder*” OR “depression” OR “worry” OR “fear” OR

"anxiety" OR "anxiety disorders" OR "social anxiety disorders" OR "stress" OR "Stress disorder*" OR "perceived stress" OR "irritability" OR "happiness" OR "psychological disturbance" OR "psychological disorder" OR "sleep disturbance" OR "loneliness" OR "boredom" OR "insomnia" OR "post-traumatic disorder" OR "trauma" OR "psychological distress" OR "psychological stress" OR "Social isolation") AND NOFT("COVID-19" OR "SARS-COV-2" OR "2019-ncov" OR "2019 novel coronavirus" OR "novel coronavirus" OR "nCoV" OR "Coronavirus" OR "COVID-19 pandemic*")
Total = 471

S2. DETAILED METHOD OF ASSESSING RISK OF BIAS

The risk of bias assessment in individual studies has been developed following the steps from previously published systematic reviews protocol (Luque-García et al., 2022; Uwak et al., 2021).

Please, rate the questions within each domain with the following answers and provide details or justification:

- Low risk (equivalent to "yes")
- Probably low risk (equivalent to "probably yes")
- Probably high risk (equivalent to "probably no")
- High risk (equivalent to "no")

Was the strategy for recruiting participants consistent across study groups?

LOW risk of bias (i.e., answer: "YES"): Protocols for recruitment and inclusion/exclusion criteria were applied similarly across study groups, and any one of the following:

- Study participants were recruited from the same population at the same time frame.
- Study participants were not all recruited from the same population, but proportions of participants from each population in each study group are uniform.

PROBABLY LOW risk of bias (i.e., answer: "Probably Yes"): There is insufficient information about participant selection to permit a judgment of 'YES', but there is indirect evidence that suggests that participant recruitment and inclusion/exclusion criteria was consistent, as described by the criteria for a judgment of 'YES'.

PROBABLY HIGH risk of bias (i.e., answer: "Probably No"): There is insufficient information about participant selection to permit a judgment of 'NO', but there is indirect evidence that suggests that participant recruitment or inclusion/exclusion criteria was inconsistent, as described by the criteria for a judgment of 'NO'.

HIGH risk of bias (i.e., answer: "No"): Any of the following:

- Protocols for recruitment or inclusion/exclusion criteria were applied differently across study groups.
- Study participants were recruited at different time frames.
- Study participants were recruited from different populations and proportions of participants from each population in each study group are not uniform.
- Differential loss to follow-up between groups.
- Reported refusal/non-response is uniform between groups.

NOT APPLICABLE (risk of bias domain is not applicable to study): There is evidence that participant selection is not an element of study design capable of introducing risk of bias in the study.

Was knowledge of the exposure adequately prevented during the study?

LOW risk of bias (i.e., answer: "YES"): Any of the following:

No blinding, but the review authors judge that the outcome and the outcome measurement as well as the exposure and exposure measurement are not likely to be influenced by lack of blinding (such as differential outcome assessment where the outcome is assessed using different measurement or estimation metrics across exposure groups, or differential exposure assessment where exposure is assessed using different measurement or estimation metrics across diagnostic or outcome groups).

Blinding of key study personnel was ensured, and it is unlikely that the blinding could have been broken.
Some key study personnel were not blinded, but exposure and outcome assessment was blinded and the non-blinding of others is unlikely to introduce bias.

PROBABLY LOW risk of bias (i.e., answer: "Probably Yes"): There is insufficient information about blinding to permit a judgment of low risk of bias, but there is indirect evidence that suggests the study was adequately blinded, as described by the criteria for a judgment of low risk of bias. For example, investigators were effectively blinded to the exposure and/or outcome groups, for example, if the exposure was measured by a separate entity and the outcome was obtained from a hospital record.

PROBABLY HIGH risk of bias (i.e., answer: "Probably No"): There is insufficient information about blinding to permit a judgment of high risk of bias, but there is indirect evidence that suggests the study was not adequately blinded, as described by the criteria for a judgment of high risk of bias.

HIGH risk of bias (i.e., answer: "No"): Any of the following:

- No blinding or incomplete blinding, and the outcome or outcome measurement or exposure and exposure measurement is likely to be influenced by lack of blinding (i.e., differential outcome or exposure assessment).
- Blinding of key study personnel attempted, but likely that the blinding could have been broken to introduce bias.
- Some key study personnel were not blinded, and the non-blinding of others was likely to introduce bias.

NOT APPLICABLE (risk of bias domain is not applicable to study): There is evidence that blinding is not an element of study design capable of introducing risk of bias in the study.

Were exposure assessment methods robust?

The following scoring list represents a collection of factors that may potentially influence the internal validity of the exposure assessment (Luque-García et al., 2022). The score list is composed of 6 questions, each with two possible answers and scores ("0" or "1"). The maximum obtainable score is 6 points, and the minimum is 0 points. The risk of bias is determined based on the obtained points.

- **LOW risk of bias (i.e., answer: "Yes"):** 5-6 points.
- **PROBABLY LOW risk of bias (i.e., answer: "Probably No"):** 3-4 points.
- **PROBABLY HIGH risk of bias (i.e., answer: "Probably No"):** 2 points.
- **HIGH risk of bias (i.e., answer: "No"):** 0-1 points.

Greenness exposure scoring list:

How many exposure metrics were considered?
- One (0)
- Multiple (1)

Did the exposure metrics differentiate between types of greenness (e.g.: forests, urban parks, gardens, trees, grassland, bushes, etc.)?
- No (0)
- Yes (1)

Was the exposure assessed more than once over time?
- No (0)
- Yes (1)

Were different buffer distances or areas (e.g.: neighborhood, school, home, etc.) considered for the exposure measurement?
- No (0)
- Yes (1)

Was quality of greenness or accessibility assessed (e.g.: survey, lad use classification, shortest distance nearest green space etc.)?
- No (0)

- Yes (1)

Was use of green space or exposure time measured (e.g.: GPS tracking, surveys, etc.)?
- No (0)
- Yes (1)

Was confounding adequately addressed?

Before the evaluation of studies, coauthors collectively developed the following list with the most important confounders. A list with two potential mediators was also included, to penalize studies that over-adjusted their models by including mediators together with confounders in the model.

Confounders

Individual-level socioeconomic status (e.g., household income, parental education, race/ethnicity, marital status, maternal age at birth, parental employment status, etc.)

Neighborhood-level socioeconomic status (e.g., neighborhood household income, neighborhood deprivation, neighborhood safety, population density, etc.)

Mediators
- Air Pollution
- Noise
- Heat
- Physical Activity
- Social Cohesion
- Stress
- Attention

LOW risk of bias (i.e., answer: "Yes"): The study accounted for both individual-level and neighborhood-level socioeconomic status, or reported that potential confounders were evaluated and omitted because inclusion did not substantially affect the results.

PROBABLY LOW risk of bias (i.e., answer: "Probably yes"): The study accounted for only one of the two potential confounders, individual-level or neighborhood-level socioeconomic status. This lack of accounting is not expected to introduce substantial bias.

PROBABLY HIGH risk of bias (i.e., answer: "Probably no"): The study accounted for one or both potential confounders, individual-level or neighborhood-level socioeconomic status, but also included potential mediators, air pollution or noise, in the model. Overadjustment of the model may have introduced substantial bias.

HIGH risk of bias (i.e., answer: "No"): The study did not account for none of our listed potential confounders.

Were incomplete outcome data adequately addressed?

LOW risk of bias (i.e., answer: "Yes"): Participants were followed long enough to obtain outcome measurements, or any one of the following:

- No missing outcome data.
- Reasons for missing outcome data unlikely to be related to true outcome (for survival data, censoring unlikely to introduce bias.
- Attrition or missing outcome data balanced in numbers across exposure groups, with similar reasons for missing data across groups.
- For dichotomous outcome data, the proportion of missing outcomes compared with observed event risk not enough to have a relevant impact on the intervention effect estimate.
- For continuous outcome data, plausible effect size (difference in means or standardized difference in means) among missing outcomes not enough to have a relevant impact on the observed effect size.
- Missing data have been imputed using appropriate methods.

PROBABLY LOW risk of bias (i.e., answer: "Probably yes"): There is insufficient information about incomplete outcome data to permit a judgment of low risk of bias, but there is indirect evidence that suggests incomplete outcome data was adequately addressed, as described by the criteria for a judgment of low risk of bias.

PROBABLY HIGH risk of bias (i.e., answer: "Probably no"): There is insufficient information about incomplete outcome data to permit a judgment of high risk of bias, but there is indirect evidence that suggests incomplete outcome data was not adequately addressed, as described by the criteria for a judgment of high risk of bias.

HIGH risk of bias (i.e., answer: "No"): Participants were not followed long enough to obtain outcome measurements, or any one of the following:

- Reason for missing outcome data likely to be related to true outcome, with either imbalance in numbers or reasons for missing data across exposure groups.
- For dichotomous outcome data, the proportion of missing outcomes compared with observed event risk enough to induce biologically relevant bias in intervention effect estimate.
- For continuous outcome data, plausible effect size (difference in means or standardized difference in means) among missing outcomes enough to induce biologically relevant bias in observed effect size.
- Potentially inappropriate application of imputation.

NOT APPLICABLE (risk of bias domain is not applicable to study): There is evidence that incomplete outcome data is not capable of introducing risk of bias in the study.

Does the study report appear to have been comprehensive in its outcome reporting?

LOW risk of bias (i.e., answer: "Yes"): All the study's pre-specified (primary and secondary) outcomes outlined in the protocol, methods, abstract, and/or introduction that are of interest in the review have been reported in the pre-specified way.

PROBABLY LOW risk of bias (i.e., answer: "Probably yes"): There is insufficient information about selective outcome reporting to permit a judgment of low risk of bias, but there is indirect evidence that suggests the study was free of selective reporting, as described by the criteria for a judgment of low risk of bias.

PROBABLY HIGH risk of bias (i.e., answer: "Probably no"): There is insufficient information about selective outcome reporting to permit a judgment of high risk of bias, but there is indirect evidence that suggests the study was not free of selective reporting, as described by the criteria for a judgment of high risk of bias.

HIGH risk of bias (i.e., answer: "No"): Any one of the following:

- Not all of the study's pre-specified primary outcomes (as outlined in the protocol, methods, abstract, and/or introduction) have been reported.
- One or more primary outcomes is reported using measurements, analysis methods or subsets of the data (e.g. subscales) that were not pre-specified.
- One or more reported primary outcomes were not pre-specified (unless clear justification for their reporting is provided, such as an unexpected effect).
- One or more outcomes of interest are reported incompletely.

NOT APPLICABLE (risk of bias domain is not applicable to study): There is evidence that selective outcome reporting is not capable of introducing risk of bias in the study.

Is the study free of support from any company, study author, or other entity having a financial interest in any of the exposures studied?

LOW risk of bias (i.e., answer: "Yes"): The study did not receive support from a company, study author, or other entity having a financial interest in the outcome of the study. Examples include the following:

- Funding source is limited to government, non-profit organizations, or academic grants funded by government, foundations and/or non-profit organizations.
- Chemicals or other treatments used in study were purchased from a supplier.
- Company affiliated staff are not mentioned in the acknowledgments section.
- Authors were not employees of a company with a financial interest in the outcome of the study.

- Company with a financial interest in the outcome of the study was not involved in the design, conduct, analysis, or reporting of the study and authors had complete access to the data.
- Study authors make a claim denying conflicts of interest.
- Study authors are unaffiliated with companies with a financial interest, and there is no reason to believe a conflict of interest exists.
- All study authors are affiliated with a government agency (are prohibited from involvement in projects for which there is a conflict of interest or an appearance of conflict of interest).

PROBABLY LOW risk of bias (i.e., answer: "Probably yes"): There is insufficient information to permit a judgment of low risk of bias, but there is indirect evidence that suggests the study was free of support from a company, study author, or other entity having a financial interest in the outcome of the study, as described by the criteria for a judgment of low risk of bias.

PROBABLY HIGH risk of bias (i.e., answer: "Probably no"): There is insufficient information to permit a judgment of high risk of bias, but there is indirect evidence that suggests the study was not free of support from a company, study author, or other entity having a financial interest in the outcome of the study, as described by the criteria for a judgment of high risk of bias.

HIGH risk of bias (i.e., answer: "No"): The study received support from a company, study author, or other entity having a financial interest in the outcome of the study. Examples of support include:

- Research funds.
- Chemicals, equipment or testing provided at no cost.
- Writing services.
- Author/staff from study was employee or otherwise affiliated with company with financial interest.
- Company limited author access to the data.
- Company was involved in the design, conduct, analysis, or reporting of the study.
- Study authors claim a conflict of interest.

NOT APPLICABLE (risk of bias domain is not applicable to study): There is evidence that conflicts of interest are not capable of introducing risk of bias in the study.

Did the study appear to be free of other problems that could put it at a risk of bias?

LOW risk of bias (i.e., answer: "Yes"): The study appears to be free of other sources of bias.

PROBABLY LOW risk of bias (i.e., answer: "Probably yes"): There is insufficient information to permit a judgment of low risk of bias, but there is indirect evidence that suggests the study was free of other threats to validity.

PROBABLY HIGH risk of bias (i.e., answer: "Probably no"): There is insufficient information to permit a judgment of high risk of bias, but there is indirect evidence that suggests the study was not free of other threats to validity, as described by the criteria for a judgment of high risk of bias.

HIGH risk of bias (i.e., answer: "No"): There is at least one important risk of bias. For example, the study:

- Had a potential source of bias related to the specific study design used.
- Stopped early due to some data-dependent process (including a formal-stopping rule).
- The conduct of the study is affected by interim results (e.g. recruiting additional participants from a subgroup showing greater or lesser effect).
- Has been claimed to have been fraudulent.
- Had some other problems.

S3. DETAILED METHOD OF ASSESSING QUALITY OF EVIDENCE

The risk of bias assessment across studies has been developed following the steps from a previously published systematic review protocol (Luque-García et al., 2022; Uwak et al., 2021). Each of the categories to consider in downgrading or upgrading the evidence is described in detail below. Please record your results on the chart at the end of each category, including a brief explanation for your ratings.

Downgrade categories

Category 1. Quality of study limitations

Possible ratings: 0=no change; -1 or -2 downgrade 1 or 2 levels

When assessing the quality of evidence from studies, it is important to take into account the potential for bias in the studies. Studies that have a high risk of bias can lower the overall rating of the evidence. The risk of bias is evaluated by looking at the outcomes across all studies, and the limitations of individual studies are summarized in heat maps.

According to GRADE, when determining the overall quality of evidence, one should not simply average the limitations of individual studies. Instead, one should carefully consider the contribution of each study, with a focus on high-quality studies[1].

[1] Note: Limitations to GRADE's risk of bias assessments as stated by GRADE: "First, empirical evidence supporting the criteria is limited. Attempts to show systematic difference between studies that meet and do not meet specific criteria have shown inconsistent results. Second, the relative weight one should put on the criteria remains uncertain. The GRADE approach is less comprehensive than many systems, emphasizing simplicity and parsimony over completeness. GRADE's approach does not provide a quantitative rating of risk of bias. Although such a rating has advantages, we share with the Cochrane Collaboration methodologists a reluctance to provide a risk of bias score that, by its nature, must make questionable assumptions about the relative extent of bias associated with individual items and fails to consider the context of the individual items."

This evaluation should take into account the size of the study sample and the number of outcome events. Larger studies with more events will have a greater impact on the estimate of the magnitude of effect. Additionally, one should be cautious when rating down due to risk of bias, and only do so if there is substantial risk of bias across the majority of available evidence.

It is also important to consider the risk of bias in the context of other limitations. If there is a close call between two quality issues (such as risk of bias and precision), GRADE suggests rating down for at least one of them.

When faced with close-call situations, reviewers should acknowledge the uncertainty and clearly explain the reasons for their ultimate judgment.

Rating for Risk of Bias (Study Limitations) 0 no change -1 decrease quality 1 level -2 decrease quality 2 levels		**Rationale for your judgment**
Anxiety	-2	5 out of 15 studies have a high risk of bias, 4 rated as probably high risk of bias and the remaining one (6/16) a low risk of bias. The studies with high risk and probably high risk of bias have a large sample size.
Depression	-2	7 out of 15 studies have a high risk of bias, 5 rated as probably high risk of bias and the remaining one (3/15) a low risk of bias. The high risk of bias studies has a large sample size.
Stress	-2	3 out of 13 studies have high risk of bias, 6 out of 13 have probably high risk of bias and the remaining 4 studies have low risk of bias. Studies with probably high risk have large sample sizes.
General mental health problem	-1	5 out of 13 studies have high risk of bias, 2 out of 13 have probably high risk of bias and the remaining 6 have low risk bias. The studies with low risk bias have large sample size.
Sleep disturbances	-2	Four studies with a high risk of bias and the remaining two studies have a low risk of bias. The study with a high rating has the largest sample size.
Mood	-2	3 out of 4 studies have high risk, and the remaining one has probably high risk of bias. The study with high risk of bias have the largest sample size.
Positive and negative affect	-1	The studies with a high (1/3) and probably high risk of bias (1/3) do not have a large sample size.
Loneliness	-2	All studies have either a high (1/2) or probably high risk of bias (1/2).
Emotional distress	0	One has high and other has low risk of bias. The study with low risk has large sample

		size.
Mental well-being	-1	Although most studies have a high (6/17) or a probably high (5/17) risk of bias rating, the studies with low risk (6/18) have a large sample size.
Happiness	-2	1 out of 4 studies has a high risk of bias, 2/4 rated as probably high risk of bias and the remaining one (1/4) a low risk of bias. The probably high risk of bias studies has a large sample size.
Life satisfaction	-2	The studies with a high (1/3) or a probably high (1/3) risk of bias rating have a large sample size.

Category 2. Indirectness of evidence

Possible ratings: 0=no change; -1 or -2 downgrade 1 or 2 levels

Quality of evidence (your confidence in estimates of effect) may decrease when substantial differences exist between the population, the exposure, or the outcomes measured in research studies under consideration in the review.

Evidence is direct when it directly compares the exposures in which we are interested when applied to the populations in which we are interested and measures outcomes important to the study question (in GRADE the outcomes must be important to patients). Based on GRADE (Guyatt et al., 2011b), evidence can be indirect in one of three ways.[2]

The population studied differs from the population of interest (the term applicability is often used for this form of indirectness). GRADE states that in general, one should not rate down for population differences unless one has compelling reason to think that the biology in the population of interest is so different than the population tested that the magnitude of effect will differ substantially. According to GRADE, most often, this will not be the case.

The intervention (exposure) tested may differ from the exposure of interest, i.e., a difference in the chemical, route and/or dose. Decisions regarding indirectness of populations and exposure depend on an understanding of whether biological or social factors are sufficiently different that one might expect substantial differences in the magnitude of effect. GRADE also states, "As with all other aspects of rating quality of evidence, there is a continuum of similarity of the intervention that will require judgment. It is rare, and usually unnecessary, for the intended populations and interventions to be identical to those in the studies, and we should only rate down if the differences are considered sufficient to make a difference in outcome likely."

Outcomes may differ from those of primary interest; for instance, surrogate outcomes that are not themselves important, but measured in the presumption that changes in the surrogate reflect changes in an important outcome. The difference between desired and measured outcomes may relate to time

[2] GRADE includes a fourth type of indirectness that occurs when there are no direct (i.e., head-to-head) comparisons between two or more interventions of interest. This criterion is not relevant to our study question; it could be relevant to future case studies.

frame. When there is a discrepancy between the time frame of measurement and that of interest, whether to rate down by one or two levels will depend on the magnitude of the discrepancy. Another source of indirectness related to measurement of outcomes is the use of substitute or surrogate endpoints in place of the exposed population's important outcome of interest. In general, the use of a surrogate outcome requires rating down the quality of evidence by one, or even two, levels. Consideration of the biology, mechanism and natural history of the disease can be helpful in making a decision about indirectness. Surrogates that are closer in the putative causal pathway to the adverse outcomes warrant rating down by only one level for indirectness. GRADE states that rarely, surrogates are sufficiently well established that one should choose not to rate down quality of evidence for indirectness. In general, evidence based on surrogate outcomes should usually trigger rating down, whereas the other types of indirectness will require a more considered judgment.

Rating for Indirectness 0 no change -1 decrease quality 1 level -2 decrease quality 2 levels		**Rationale for your judgment**
Anxiety	0	Studies measure the population, exposure and the outcome of interest.
Depression	0	Studies measure the population, exposure and the outcome of interest.
Stress	0	Studies measure the population, exposure and the outcome of interest.
General mental health problem	0	Studies measure the population, exposure and the outcome of interest.
Sleep disturbances	0	Studies measure the population, exposure and the outcome of interest.
Mood	0	Studies measure the population, exposure and the outcome of interest.
Positive and negative affect	0	Studies measure the population, exposure and the outcome of interest.
Loneliness	0	Studies measure the population, exposure and the outcome of interest.
Emotional distress	0	Studies measure the population, exposure and the outcome of interest.
Mental well-being	0	Studies measure the population, exposure and the outcome of interest.
Happiness	0	Studies measure the population, exposure and the outcome of interest.
Life satisfaction	0	Studies measure the population, exposure and the outcome of interest.

Category 3. Inconsistency of evidence

Possible ratings: 0=no change; -1 or -2 downgrade 1 or 2 levels

According to Cochrane, "when studies yield widely differing estimates of effect (heterogeneity or variability in results) investigators should look for robust explanations for that heterogeneity…When heterogeneity exists and affects the interpretation of results, but authors fail to identify a plausible explanation, the quality of the evidence decreases."

Based on GRADE (Guyatt et al., 2011c), **a body of evidence is not rated up in quality if studies yield consistent results**, **but may be rated down in quality if inconsistent**. Their stated reason is that a consistent bias will lead to consistent, spurious findings.

GRADE suggests rating down the quality of evidence if large inconsistency (heterogeneity) in study results remains after exploration of a priori hypotheses that might explain heterogeneity. Judgment of the extent of heterogeneity is based on similarity of point estimates, extent of overlap of confidence intervals, and statistical criteria. GRADE's recommendations refer to inconsistencies in effect size, specifically to relative measures (risk ratios and hazard ratios or odds ratios), not absolute measures. Based on GRADE, reviewers should consider rating down for inconsistency when:

- Point estimates vary widely across studies;
- Confidence intervals (CIs) show minimal or no overlap;
- The statistical test for heterogeneity-which tests the null hypothesis that all studies in a meta-analysis have the same underlying magnitude of effect-shows a low P-value;
- The I^2 -which quantifies the proportion of the variation in point estimates due to among-study differences-is large. (I.e., the I^2 index quantifies the degree of heterogeneity in a meta-analysis).

GRADE states that inconsistency is important **only when it reduces confidence in results in relation to a particular decision**. Even when inconsistency is large, it may not reduce confidence in results regarding a particular decision. For example, studies that are inconsistently related to the magnitude of a beneficial or harmful effect (but are in the same direction) would not be rated down; in instances when results are inconsistent as to whether there is a benefit or harm of treatment, GRADE would rate down the quality of evidence as a result of variability in results, because the meaning of the inconsistency is so relevant to the decision to treat or not to treat.

Rating for Inconsistency of Evidence 0 no change -1 decrease quality 1 level -2 decrease quality 2 levels		**Rationale for your judgment**
Anxiety	-1	The reported confidence intervals were partly overlapped.
Depression	-1	The reported confidence intervals were partly overlapped.
Stress	-1	The reported confidence intervals were partly overlapped.
General mental health problem	-2	High heterogeneity exists among studies, in particular, studies included in meta-analysis.
Sleep disturbances	-1	The reported confidence intervals were partly overlapped.
Mood	-2	Few studies available for overall outcome.

Positive and negative affect	0	Low heterogeneity
Loneliness	0	Low heterogeneity
Emotional distress	-1	The confidence intervals were partly overlapped.
Mental well-being	-1	The reported confidence intervals were partly overlapped.
Happiness	-1	The reported confidence intervals were partly overlapped.
Life satisfaction	-1	Two studies showed positive association and one study showed negative association. Heterogeneity exists.

Category 4. Imprecision of evidence

Possible ratings: 0=no change; -1 or -2 downgrade 1 or 2 levels

Cochrane states that when studies have few participants and few events, and thus have wide confidence intervals (CIs), authors can lower their rating of the quality of evidence. These ratings of precision are made as judgments by review authors. The ratings are made by looking across studies, or, if available, on the results of a meta-analysis.

GRADE defines evidence quality differently for systematic reviews and guidelines. For systematic reviews, quality refers to confidence in the estimates of effect. For guidelines, quality refers to the extent to which confidence in the effect estimate is adequate to support a particular decision (Guyatt et al., 2011a). For the purpose of step 3 of Navigation Guide, we will use the systematic review definition, because the decision phase does not occur until step 4 when recommendations for prevention are made. Thus, when reviewing the data for imprecision, evaluate your confidence in the estimate of the effect.

According to GRADE, to a large extent, CIs inform the impact of random error on evidence quality. Thus, when considering imprecision, the issue is whether the CI around the estimate of exposure effect is sufficiently narrow. If it is not, GRADE rates down the evidence quality by one level (for instance, from high to moderate). If the CI is very wide, GRADE might rate down by two levels.

Rating for Imprecision of Evidence 0 no change -1 decrease quality 1 level -2 decrease quality 2 levels		**Rationale for your judgment**
Anxiety	0	No support for other factors that may lead to downgrade of quality of evidence.
Depression	0	No support for other factors that may lead to downgrade of quality of evidence.

Stress	0	No support for other factors that may lead to downgrade of quality of evidence.
General mental health problem	0	No support for other factors that may lead to downgrade of quality of evidence.
Sleep disturbances	0	No support for other factors that may lead to downgrade of quality of evidence.
Mood	0	No support for other factors that may lead to downgrade of quality of evidence.
Positive and negative affect	-1	Decision-based on each side of the confidence intervals was associated with different judgments.
Loneliness	-2	Few studies (less than 3 studies) to make decision on imprecision.
Emotional distress	-2	Few studies (less than 3 studies) to make decision on imprecision.
Mental well-being	0	No support for other factors that may lead to downgrade of quality of evidence.
Happiness	-1	Decision-based on each side of the confidence intervals was associated with different judgments.
Life satisfaction	0	No support for other factors that may lead to downgrade of quality of evidence.

Category 5. Publication bias

Possible ratings: 0=no change; -1 or -2 downgrade 1 or 2 levels

GRADE (Guyatt et al., 2011d) and Cochrane (Higgins and Green, 2011) assess publication bias in a similar manner. Whereas "selective outcome reporting" is assessed for each study included in the review as part of the risk of bias assessment, "publication bias" is assessed on the body of evidence. GRADE states that "when an entire study remains unreported and the results relate to the size of the effect- publication bias- one can assess the likelihood of publication bias only by looking at a group of studies."

Cochrane's definition of publication bias is "the *publication* or *non-publication* of research findings depending on the nature and direction of the results." Cochrane and GRADE are primarily concerned with *overestimates* of true effects of treatments or pharmaceuticals, especially related to "small studies effects", i.e., the tendency for estimates of an intervention to be more beneficial in smaller studies. There is empirical evidence in the clinical sciences that publication and other reporting biases result in over estimating the effects of interventions (Higgins and Green, 2011).

In contrast, in environmental health, we are primarily concerned with *underestimating* the true effects of a chemical exposure, since in many cases population wide exposure has already occurred. We are also concerned that studies finding no association are less likely to be published because journals are less likely to publish "negative" findings.

Applying this inverted concern to GRADE's assessment for publication bias, leads to these considerations when rating publication bias:

- Early *negative* studies, particularly if small in size, are suspect. (GRADE is concerned with early *positive* studies).

- Authors of systematic reviews should suspect publication bias when studies are uniformly small, particularly when sponsored by the industry. (Same as GRADE)
- Empirical examination of patterns of results (e.g., funnel plots) may suggest publication bias but should be interpreted with caution. (Same as GRADE)
- More compelling than any of these theoretical exercises is authors' success in obtaining the results of some unpublished studies and demonstrating that the published and unpublished data show different results. (Same as GRADE)
- Comprehensive searches of the literature, including unpublished studies, i.e., the grey literature, and a search for research in other languages are important to addressing publication bias. Note that Cochrane also states "comprehensive searching is not sufficient to prevent some substantial potential biases."

Rating for Publication Bias 0 no change -1 decrease quality 1 level -2 decrease quality 2 levels		**Rationale for your judgment**
Anxiety	0	Inadequate studies to evaluate publication bias. However, given the comprehensive search and size of the sample in the published studies we decided little or even no publication bias.
Depression	0	Inadequate studies to evaluate publication bias. However, given the comprehensive search and size of the sample in the published studies we decided little or even no publication bias.
Stress	0	Inadequate studies to evaluate publication bias. However, given the comprehensive search and size of the sample in the published studies we decided little or even no publication bias.
General mental health problem	0	Inadequate studies to evaluate publication bias. However, given the comprehensive search and size of the sample in the published studies we decided little or even no publication bias.
Sleep disturbances	0	Inadequate studies to evaluate publication bias. However, given the comprehensive search and size of the sample in the published studies we decided little or even no publication bias.
Mood	0	Inadequate studies to evaluate publication bias. However, given the comprehensive search and size of the sample in the published studies we decided little or even no publication bias.
Positive and negative affect	0	Inadequate studies to evaluate publication bias. However, given the comprehensive search and size of the sample in the published studies we decided little or even no publication bias.
Loneliness	0	Inadequate studies to evaluate publication bias. However, given the comprehensive search and size of the sample in the published studies we decided little or even no publication bias.
Emotional distress	0	Inadequate studies to evaluate publication bias. However, given the comprehensive search and size of the sample in the published studies we decided little or even no publication bias.

Mental well-being	0	Inadequate studies to evaluate publication bias. However, given the comprehensive search and size of the sample in the published studies we decided little or even no publication bias.
Happiness	0	Inadequate studies to evaluate publication bias. However, given the comprehensive search and size of the sample in the published studies we decided little or even no publication bias.
Life satisfaction	0	Inadequate studies to evaluate publication bias. However, given the comprehensive search and size of the sample in the published studies we decided little or even no publication bias.

Upgrade categories

GRADE states that the circumstances for upgrading likely occur infrequently and are primarily relevant to observational and other non-randomized studies. Although it is possible to rate up results from randomized controlled trials, GRADE has yet to find a compelling circumstance for doing so (Guyatt et al., 2011e). GRADE specifies 3 categories for increasing the quality of evidence (Guyatt et al., 2011e)

Category 6. Large magnitude of effect

Possible ratings: 0=no change; +1 or +2 upgrade 1 or 2 levels

Modeling studies suggest that confounding (from non-random allocation) alone is unlikely to explain associations with a relative risk (RR) greater than 2 (or less than 0.5), and very unlikely to explain associations with an RR greater than 5 (or less than 0.2). Thus, these are the definitions of "large magnitude of effect" used by GRADE to upgrade 1 or 2 levels, respectively. Also, GRADE is more likely to rate up if the effect is rapid and out of keeping with prior trajectory; usually supported by indirect evidence. GRADE presents empirical evidence to support these conclusions, and states that "although further research is warranted, both modeling and empirical work suggest the size of bias from confounding is unpredictable in direction but bounded in size. Hence, the GRADE group has previously suggested guidelines for rating quality of evidence up by one category (typically from low to moderate) for associations greater than 2, and up by two categories for associations greater than 5."

Applying the GRADE definitions of large magnitude of effect i.e., RR greater than 2 or 0.5 is problematic in environmental health because for dichotomous outcomes RR is a function of the exposure comparator; these definitions also are not applicable to results from continuous variables. At present, we do not have an empirically defined "large magnitude of effect." Therefore, for the purpose of this case study, co-authors should assess whether the results indicate a large magnitude of effect using their expert judgment of "large effects" in environmental health and state their definition for discussion by the group.

Rating for Large Magnitude of Effect 0 no change -1 decrease quality 1 level -2 decrease quality 2 levels		**Rationale for your judgment**
Anxiety	0	In all reported studies, the magnitude of the effect did not meet the prespecified criteria.

Depression	0	In all reported studies, the magnitude of effect did not meet the prespecified criteria.
Stress	0	In all reported studies, the magnitude of effect did not meet the prespecified criteria.
General mental health problem	0	In all reported studies, the magnitude of effect did not meet the prespecified criteria.
Sleep disturbances	0	In all reported studies, the magnitude of effect did not meet the prespecified criteria.
Mood	0	In all reported studies, the magnitude of effect did not meet the prespecified criteria.
Positive and negative affect	0	In all reported studies, the magnitude of effect did not meet the prespecified criteria.
Loneliness	0	In all reported studies, the magnitude of effect did not meet the prespecified criteria.
Emotional distress	0	In all reported studies, the magnitude of effect did not meet the prespecified criteria.
Mental well-being	0	In all reported studies, the magnitude of effect did not meet the prespecified criteria.
Happiness	0	In all reported studies, the magnitude of effect did not meet the prespecified criteria.
Life satisfaction	0	In all reported studies, the magnitude of effect did not meet the prespecified criteria.

Category 7: Dose response

Possible ratings: 0=no change; +1 or +2 upgrade 1 or 2 levels

Possible considerations include consistent dose-response gradients in one or multiple studies, and/or dose-response across studies, depending on the overall relevance to the body of evidence.

Rating for Dose-response 0 no change -1 decrease quality 1 level -2 decrease quality 2 levels		**Rationale for your judgment**
Anxiety	0	No support for other factors that may lead to upgrade of quality of evidence.
Depression	0	No support for other factors that may lead to upgrade of quality of evidence.
Stress	0	No support for other factors that may lead to upgrade of quality of evidence.
General mental health problem	0	No support for other factors that may lead to upgrade of quality of evidence.
Sleep disturbances	0	No support for other factors that may lead to upgrade of quality of evidence.
Mood		No support for other factors that may lead to upgrade of quality of evidence.

Positive and negative affect	0	No support for other factors that may lead to upgrade of quality of evidence.
Loneliness	0	No support for other factors that may lead to upgrade of quality of evidence.
Emotional distress	0	No support for other factors that may lead to upgrade of quality of evidence.
Mental well-being	0	No support for other factors that may lead to upgrade of quality of evidence.
Happiness	0	No support for other factors that may lead to upgrade of quality of evidence.
Life satisfaction	0	No support for other factors that may lead to upgrade of quality of evidence.

Category 8. Confounding minimizes effect

Possible ratings: 0=no change; +1 or +2 upgrade 1 or 2 levels

All plausible residual confounders or biases would reduce a demonstrated effect, or suggest a spurious effect when results show no effect. GRADE provides the following example of grading up evidence when observational studies have failed to demonstrate an association.

Rating for Confounding Minimizes Effect 0 no change -1 decrease quality 1 level -2 decrease quality 2 levels		**Rationale for your judgment**
Anxiety	0	No support for other factors that may lead to upgrade of quality of evidence.
Depression	0	No support for other factors that may lead to upgrade of quality of evidence.
Stress	0	No support for other factors that may lead to upgrade of quality of evidence.
General mental health problem	0	No support for other factors that may lead to upgrade of quality of evidence.
Sleep disturbances	0	No support for other factors that may lead to upgrade of quality of evidence.
Mood	0	No support for other factors that may lead to upgrade of quality of evidence.
Positive and negative affect	0	No support for other factors that may lead to upgrade of quality of evidence.
Loneliness	0	No support for other factors that may lead to upgrade of quality of evidence.
Emotional distress	0	No support for other factors that may lead to upgrade of quality of evidence.
Mental well-being	0	No support for other factors that may lead to upgrade of quality of evidence.
Happiness	0	No support for other factors that may lead to upgrade of quality of evidence.
Life satisfaction	0	No support for other factors that may lead to upgrade of quality of evidence.

The results of the reviewers' ratings by population will be compiled and discussed leading to a final decision on overall quality of human evidence. The rationale for the decision will be fully documented.

Final decision ratings

- High
- Moderate
- Low
- Very Low

TABLE S1. DATA RETRIEVED FROM CORRESPONDING AUTHORS

SL	Author	Data requested	Data received
1	Cheng et al. 2021	Exposures used in Table 1 for adjusted logistic regressions	(model 1) TCC, Distance to nearest surrounding green space and Distance to nearest surrounding blue space (model 2) TCC, Distance to nearest surrounding green space and Distance to nearest surrounding blue space, TCC & Epidemic
2	Huerta and Utomo, 2021	95% CI or exact p-values of logistic regression for exposure–outcome in Table 4	**UGS use during COVID-19 (stopped using vs. started using), OR (95% CI):** started using UGS = 1.460 (1.21-1.76), p<0.001; T**ime to closest UGS (min), OR (95% CI):** 0 to 10 (ref.); 11 to 20 = 1.119 (0.87-1.44); 21+ = 1.047 (0.80-1.37).
3	Lehberger et al. 2021	Standardized beta value of Table 4	**Being outside vs Life satisfaction:** (reference = ≤4 hours), > 4 to ≤10 hours (2nd quartile) = 0.068, > 11 to ≤18 hours (3rd quartile) = 0.072, > 18 hours (4th quartile) = 0.076, Change in time spend outside for sport = 0.051, Change in time spend outside for leisure = 0.149; **Garden owner** (yes vs. no) = 0.095 **Being outside vs Mental well-being:** (reference = ≤4 hours), > 4 to ≤10 hours (2nd quartile) = 0.149, > 11 to ≤18 hours (3rd quartile) = 0.171, > 18 hours (4th quartile) = 0.161, Change in time spend outside for sport = 0.078, Change in time spend outside for leisure = 0.158;

			Garden owner (yes vs. no) = 0.068
4	Lõhmus et al., 2021	Standardized beta value for mental health estimate and NDVI (Suppl Table S5.C and Table 8)	**Mental health estimate (Std beta)** **50m** **100m** **300m** **500m** Mental health score (RAND36) 0.049 0.047 0.047 0.038 Vitality score (RAND36) 0.034 0.031 0.026 0.024 Anxiety score (SCL90) -0.055 -0.052 -0.039 -0.038 Depression score (SCL90) -0.036 -0.034 -0.039 -0.033 Perceived Stress Scale (PSS) -0.042 -.039 -0.047 -0.046 The Cognitive Stress Score (COPSOQ) -0.044 -0.045 -0.054 -0.054
5	Marques et al., 2021	Standardized beta value and 95% CI/SE value for exposure–outcome of Suppl Table (S1-S5)	**Parameter** **Coefficient** **SE** **95% CI** **t(164)** (Intercept) 60.36 4.48 [51.52, 69.20] 13.48 frequency of visit to parks -3.9 1.02 [-5.93, -1.88] -3.81 estimated green view -1.32 0.67 [-2.64, -0.01] -1.99 home garden (yes) -10.92 3.93 [-18.67, -3.16] -2.78 gender (male) -13.22 4.42 [-21.94, -4.49] -2.99 household distance form parks 2.1 1.93 [-1.72, 5.92] 1.08 leaving home for non-essential activities 0.95 0.89 [-0.80, 2.70] 1.07 gardening activities -0.95 0.82 [-2.56, 0.67] -1.16
6	P´erez-Urrestarazu et al., 2020	OR and 95% CI and also standardized b coefficients for the effect of frequency of visits and house plants	Not responded
7	Poortinga et al. 2021	Standardized betas and OR with 95% CI for exposure–outcome of Table 4	Coefficients: **Estimate** **Std. Error** **t value** **Pr(>|t|)** (Intercept) 2.62326 0.05686 46.139 < 2e-16 *** green_space5-10 min -0.17183 0.03312 -5.188 2.20e-07 *** green_space>10 min -0.37253 0.04501 -8.277 < 2e-16 *** garden_privateYes 0.17852 0.05164 3.457 0.00055 *** Male 0.31091 0.03131 9.930 < 2e-16 *** age_centred 0.17456 0.01356 12.869 < 2e-16 *** Retired 0.10976 0.03898 2.816 0.00488 ** Unemployed -0.56084 0.05834 -9.613 < 2e-16 *** married_or_living_togetherMarried, civil partnership or living together 0.16451 0.03097 5.312 1.12e-07 *** Signif. codes: 0 ‘***’ 0.001 ‘**’ 0.01 ‘*’ 0.05 ‘.’ 0.1 ‘ ’ 1
8	Robinson et al., 2021	Standardized OR (95% CI) for exposure–outcome of Table 4 & 5	Not responded

TABLE S2. RISK OF BIAS RATINGS BY STUDY

	Author	Recruitment	Blinding	Exposure Assessment	Confounding	Incomplete Outcome Data	Selective Reporting	Conflicts Of Interest	Other Bias	Overall Risk Of Bias
1	Amerio et al. 2021	**LOW**	**LOW**	**HIGH**	**Probably LOW**	**LOW**	**LOW**	**LOW**	**LOW**	**HIGH**
2	Asim et al. 2021	**LOW**	**LOW**	**HIGH**	**HIGH**	**LOW**	**LOW**	**LOW**	**LOW**	**HIGH**
3	Basu et al. 2021	**Probably HIGH**	**LOW**	**Probably HIGH**	**Probably LOW**	**LOW**	**LOW**	**LOW**	**LOW**	**Probably HIGH**
4	Bourion-Beds et al. 2020	**LOW**	**LOW**	**Probably LOW**	**Probably LOW**	**LOW**	**LOW**	**LOW**	**LOW**	**LOW**
5	Bourion-Beds et al. 2021	**LOW**	**LOW**	**Probably LOW**	**Probably LOW**	**LOW**	**LOW**	**LOW**	**LOW**	**LOW**
6	Browning et al. 2020	**LOW**	**LOW**	**Probably HIGH**	**Probably LOW**	**LOW**	**LOW**	**LOW**	**LOW**	**Probably HIGH**
7	Bu et al. 2020	**LOW**	**LOW**	**HIGH**	**Probably LOW**	**LOW**	**LOW**	**LOW**	**LOW**	**HIGH**
8	Chen and Liu 2021	**Probably LOW**	**LOW**	**HIGH**	**Probably HIGH**	**LOW**	**LOW**	**LOW**	**LOW**	**HIGH**
9	Cheng et al. 2021	**Probably LOW**	**LOW**	**Probably LOW**	**Probably HIGH**	**LOW**	**Probably LOW**	**LOW**	**LOW**	**Probably HIGH**
10	Corley et al. 2020	**Probably LOW**	**LOW**	**Probably LOW**	**Probably LOW**	**LOW**	**LOW**	**LOW**	**LOW**	**LOW**
11	Dzhambov et al. 2020	**LOW**	**LOW**	**Probably LOW**	**Probably LOW**	**LOW**	**LOW**	**LOW**	**LOW**	**LOW**
12	Friedman et al., 2021	**LOW**	**LOW**	**High**	**Probably LOW**	**LOW**	**LOW**	**LOW**	**LOW**	**High**
13	Garrido-Cumbrera et al. 2021	**LOW**	**LOW**	**HIGH**	**Probably LOW**	**LOW**	**LOW**	**LOW**	**LOW**	**HIGH**
14	Gola et al. 2021	**Probably LOW**	**LOW**	**Probably LOW**	**HIGH**	**LOW**	**LOW**	**LOW**	**LOW**	**HIGH**
15	Groot et al. 2021	**Probably LOW**	**LOW**	**Probably LOW**	**Probably LOW**	**LOW**	**LOW**	**LOW**	**HIGH**	**HIGH**
16	Hansmann et al. 2021	**LOW**	**LOW**	**HIGH**	**Probably LOW**	**LOW**	**LOW**	**LOW**	**LOW**	**HIGH**
17	Heo et al. 2021	**LOW**	**LOW**	**Probably LOW**	**Probably LOW**	**LOW**	**LOW**	**LOW**	**LOW**	**LOW**

18	Hubbard et al. 2021	**LOW**	**LOW**	**HIGH**	**Probably LOW**	**LOW**	**LOW**	**LOW**	**LOW**	**HIGH**
19	Huerta and Utomo, 2021	**LOW**	**LOW**	**Probably LOW**	**Probably LOW**	**LOW**	**LOW**	**LOW**	**LOW**	**LOW**
20	Jackson et al. 2021a	**LOW**	**LOW**	**Probably LOW**	**Probably HIGH**	**LOW**	**LOW**	**LOW**	**LOW**	**Probably HIGH**
21	Jackson et al. 2021b	**LOW**	**LOW**	**Probably LOW**	**Probably HIGH**	**LOW**	**LOW**	**LOW**	**LOW**	**Probably HIGH**
22	Jato-Espino et al. 2021	**Probably LOW**	**LOW**	**LOW**	**Probably HIGH**	**LOW**	**LOW**	**LOW**	**LOW**	**Probably HIGH**
23	Kontsevaya et al. 2020	**LOW**	**LOW**	**Probably HIGH**	**Probably LOW**	**LOW**	**LOW**	**HIGH**	**LOW**	**HIGH**
24	Kou et al. 2021	**Probably LOW**	**LOW**	**HIGH**	**HIGH**	**LOW**	**LOW**	**HIGH**	**LOW**	**HIGH**
25	Lades et al. 2020	**LOW**	**LOW**	**HIGH**	**Probably LOW**	**LOW**	**LOW**	**LOW**	**LOW**	**HIGH**
26	Larson et al. 2021	**LOW**	**LOW**	**Probably LOW**	**LOW**	**LOW**	**LOW**	**LOW**	**LOW**	**LOW**
27	Lee et al. 2021	**LOW**	**LOW**	**Probably LOW**	**Probably LOW**	**LOW**	**LOW**	**LOW**	**LOW**	**LOW**
28	Lehberger et al. 2021	**LOW**	**LOW**	**Probably LOW**	**Probably LOW**	**LOW**	**LOW**	**LOW**	**LOW**	**LOW**
29	Lenaerts et al. 2021	**Probably LOW**	**LOW**	**Probably LOW**	**Probably LOW**	**LOW**	**LOW**	**LOW**	**LOW**	**LOW**
30	Lõhmus et al., 2021	**LOW**	**LOW**	**LOW**	**Probably HIGH**	**LOW**	**LOW**	**LOW**	**LOW**	**Probably HIGH**
31	Marques et al., 2021	**LOW**	**LOW**	**Probably LOW**	**Probably LOW**	**LOW**	**LOW**	**LOW**	**LOW**	**LOW**
32	Mead et al., 2021	**LOW**	**LOW**	**HIGH**	**Probably LOW**	**LOW**	**LOW**	**Probably HIGH**	**LOW**	**HIGH**
33	Millán-Jiménez et al., 2021	**LOW**	**LOW**	**Probably LOW**	**HIGH**	**LOW**	**LOW**	**LOW**	**LOW**	**HIGH**
34	Mintz et al., 2021	**LOW**	**LOW**	**Probably LOW**	**Probably LOW**	**LOW**	**LOW**	**LOW**	**LOW**	**LOW**
35	Okely et al., 2021	**LOW**	**LOW**	**Probably LOW**	**HIGH**	**LOW**	**LOW**	**LOW**	**LOW**	**HIGH**
36	Olszewska-Guizzo et al., 2021	**LOW**	**LOW**	**Probably HIGH**	**HIGH**	**LOW**	**LOW**	**Probably LOW**	**LOW**	**HIGH**
37	Oswald et al., 2021	**LOW**	**LOW**	**Probably LOW**	**Probably LOW**	**LOW**	**LOW**	**LOW**	**LOW**	**LOW**
38	Pearson et al., 2021	**LOW**	**LOW**	**LOW**	**Probably LOW**	**LOW**	**LOW**	**LOW**	**LOW**	**LOW**
39	P´erez-Urrestarazu et al., 2020	**LOW**	**LOW**	**Probably HIGH**	**HIGH**	**LOW**	**LOW**	**LOW**	**LOW**	**HIGH**
40	Poortinga et al. 2021	**LOW**	**LOW**	**Probably LOW**	**Probably LOW**	**LOW**	**LOW**	**Probably LOW**	**LOW**	**LOW**
41	Pouso et al., 2020	**Probably LOW**	**LOW**	**Probably LOW**	**LOW**	**LOW**	**LOW**	**LOW**	**LOW**	**LOW**
42	Riberio et al., 2021	**Probably LOW**	**LOW**	**LOW**	**LOW**	**LOW**	**LOW**	**LOW**	**LOW**	**LOW**
43	Robinson et al., 2021	**Probably LOW**	**LOW**	**Probably LOW**	**Probably LOW**	**LOW**	**LOW**	**Probably HIGH**	**LOW**	**Probably HIGH**
44	Samuelsson et al., 2021	**LOW**	**LOW**	**Probably LOW**	**Probably LOW**	**LOW**	**LOW**	**Probably LOW**	**LOW**	**LOW**
45	Samus et al., 2021	**Probably HIGH**	**LOW**	**Probably LOW**	**Probably LOW**	**LOW**	**LOW**	**LOW**	**LOW**	**Probably HIGH**
46	Sansal et al., 2021	**LOW**	**LOW**	**HIGH**	**HIGH**	**Probably LOW**	**LOW**	**LOW**	**LOW**	**HIGH**
47	Sia et al., 2021	**Probably LOW**	**LOW**	**HIGH**	**Probably LOW**	**LOW**	**LOW**	**LOW**	**LOW**	**HIGH**
48	Soga et al., 2020	**LOW**	**LOW**	**Probably HIGH**	**Probably LOW**	**Probably LOW**	**LOW**	**LOW**	**LOW**	**Probably HIGH**
49	Spano et al., 2021	**LOW**	**LOW**	**Probably LOW**	**Probably LOW**	**LOW**	**LOW**	**LOW**	**LOW**	**LOW**
50	Sundara keeran et al., 2021	**LOW**	**LOW**	**Probably HIGH**	**HIGH**	**LOW**	**LOW**	**LOW**	**LOW**	**HIGH**
51	Theodorou et al., 2021	**LOW**	**LOW**	**High**	**Probably LOW**	**LOW**	**LOW**	**LOW**	**LOW**	**High**
52	Tomasso et al., 2021	**LOW**	**LOW**	**LOW**	**LOW**	**LOW**	**LOW**	**Low**	**LOW**	**Low**
53	Vos et al., 2021	**LOW**	**LOW**	**Probably HIGH**	**Probably LOW**	**LOW**	**LOW**	**LOW**	**LOW**	**Probably HIGH**
54	Wortzel et al., 2021	**LOW**	**LOW**	**Probably HIGH**	**Probably LOW**	**LOW**	**LOW**	**Low**	**LOW**	**Probably HIGH**
55	Xie et al., 2020	**LOW**	**LOW**	**Probably HIGH**	**HIGH**	**LOW**	**LOW**	**LOW**	**LOW**	**HIGH**
56	Yao et al., 2021	**Probably LOW**	**LOW**	**HIGH**	**Probably HIGH**	**LOW**	**LOW**	**LOW**	**LOW**	**HIGH**
57	Young et al., 2021	**LOW**	**LOW**	**Probably HIGH**	**LOW**	**LOW**	**LOW**	**LOW**	**LOW**	**Probably HIGH**
58	Zhang et al., 2021	**LOW**	**LOW**	**LOW**	**Probably LOW**	**LOW**	**LOW**	**LOW**	**LOW**	**LOW**
59	Zhuo and Zacharias, 2021	**LOW**	**LOW**	**Probably HIGH**	**Probably LOW**	**LOW**	**LOW**	**LOW**	**LOW**	**Probably HIGH**

TABLE S3. JUSTIFICATION FOR RISK OF BIAS RATINGS

1. **Narrative justification for (Amerio et al., 2020).**

Bias	Rating	Support for judgment
Recruitment	**Low**	A web-based survey questionnaire was sent by mail from 1 April 2020 to 1 May 2020 to students from a University Institute in Milan, Lombardy region, Italy. The study was performed three weeks after the COVID-19 epidemic outbreak in Italy. The inclusion criteria characteristics are defined (undergraduate students and aged>18 years old) and data is obtained by questionnaires. Participants were recruited from same population at same time frame. Eligibility criteria were explained clearly.
Blinding	**Low**	There was no information about blinding for exposure-outcome, but this does not influence the outcome and exposure measurement.
Exposure assessment	**High**	They assessed multiple greenspace elements (1). Exposure metrics did not differentiate between types of greenness (0). The exposure was not assessed more than once over time (0). No buffer distances or areas were considered for exposure measurements (0). They did not assess the quality of greenness or accessibility (0). They did not measure the use of green space or exposure time (0).
Confounding	**Probably low**	The study accounted for only one potential confounder, individual-level socioeconomic status. Model adjusted for age and gender.
Incomplete outcome data	**Low**	No missing outcome data.
Selective reporting	**Low**	All outcomes were reported.
Conflict of interest	**Low**	The study did not receive support from a company, study author, or other entity having a financial interest in the outcome of the study as authors make a claim denying conflicts of interest.
Other bias	**Low**	Appears free of other sources of bias.

2. **Narrative justification for (Asim et al., 2021).**

Bias	Rating	Support for judgment
Recruitment	**Low**	A cross-sectional study using google form questionnaire were used to collect data from a COVID-19 containment zone with 3000 students of Haridwar district of Himalayan state of Uttarakhand of Indian during April 2021. All participating 432 students in this study were recruited from

		the same containment zone. Participants were recruited from same population at same time frame. Eligibility criteria were explained clearly.
Blinding	**Low**	There was no information about blinding for exposure-outcome, but this does not influence the outcome and exposure measurement.
Exposure assessment	**High**	Assessed multiple exposure metrics (1). Exposure metrics did not differentiate between types of greenness (0). The exposure was not assessed more than once over time (0). No buffer distances or areas were considered for exposure measurements (0). Quality of greenness was not assessed (0) They did not measure the use of green space or exposure time (0).
Confounding	**High**	The study did not account for any potential confounders.
Incomplete outcome data	**Low**	No missing outcome data.
Selective reporting	**Low**	All outcomes were reported.
Conflict of interest	**Low**	The study did not receive support from a company, study author, or other entity having a financial interest in the outcome of the study as authors make a claim denying conflicts of interest.
Other bias	**Low**	Appears free of other sources of bias.

3. Narrative justification for (Basu et al., 2021).

Bias	**Rating**	**Support for judgment**
Recruitment	**Probably high**	A cross-sectional study was conducted to collect data via online questionnaires. The study was conducted during the end of May 2020 in different cities in India. Study participants were recruited from different cities and proportions of participants from each population in each study group are not specified.
Blinding	**Low**	There was no information about blinding for exposure-outcome, but this does not influence the outcome and exposure measurement.
Exposure assessment	**Probably high**	They assessed multiple greenspace elements (1). Exposure metrics did not differentiate between types of greenness (0). The exposure was not assessed more than once over time (0). No buffer distances or areas were considered for exposure measurements (0). They did not assess the quality of greenness or accessibility (0). They measured the use of green space or exposure time by asking participants about the time they usually spend working in their home garden i.e., their active interaction with home gardens (1).
Confounding	**Probably low**	The study accounted for only one potential confounder, individual-level socioeconomic status. Model adjusted for age, gender, marital status, educational level, occupation, and average household income level

Incomplete outcome data	**Low**	No missing outcome data.
Selective reporting	**Low**	All outcomes were reported.
Conflict of interest	**Low**	The study did not receive support from a company, study author, or other entity having a financial interest in the outcome of the study as authors make a claim denying conflicts of interest.
Other bias	**Low**	Appears free of other sources of bias.

4. Narrative justification for (Bourion-Bedes et al., 2020).

Bias	**Rating**	**Support for judgment**
Recruitment	**Low**	This is a cross-sectional analysis of data from the observational Feelings and Psychological Impact of the COVID-19 Epidemic among Students in the Grand Est Area (PIMS-CoV19) study. The study was conducted from May 7 to 20, 2020. A sample of students was recruited from the University of Lorraine and the Sciences Po College located in Nancy, Lorraine, Grand Est region, France. Participants were recruited from same population at same time frame. Inclusion and exclusion criteria was explained clearly.
Blinding	**Low**	There was no information about blinding for exposure-outcome, but this does not influence the outcome and exposure measurement.
Exposure assessment	**Probably low**	They assessed multiple greenspace elements (1). Exposure metrics differentiated between types of greenness (1). The exposure was not assessed more than once over time (0). No buffer distances or areas were considered for exposure measurements (0). They assessed the quality of greenness or accessibility (1) They did not assess the time spent in greenspace (0).
Confounding	**Probably low**	The study accounted for only one potential confounder, individual-level socioeconomic status. Model adjusted for age, gender, home location, living arrangement, academic performance and scholarship status.
Incomplete outcome data	**Low**	No missing outcome data.
Selective reporting	**Low**	All outcomes were reported.
Conflict of interest	**Low**	The study did not receive support from a company, study author, or other entity having a financial interest in the outcome of the study as authors make a claim denying conflicts of interest.
Other bias	**Low**	Appears free of other sources of bias.

5. **Narrative justification for (Bourion-Bedes et al., 2021).**

Bias	Rating	Support for judgment
Recruitment	**Low**	This is a cross-sectional analysis of data from the observational Feelings and Psychological Impact of the COVID-19 Epidemic among Students in the Grand Est Area (PIMS-CoV19) study. Every student from the University of Lorraine and the Sciences Po College located in Nancy, Grand Est region, France, was eligible to participate in the study. Students were recruited to participate in an anonymous online survey from May 7 to May 17, 2020. Participants were recruited from same population at same time frame. Inclusion and exclusion criteria was explained.
Blinding	**Low**	There was no information about blinding for exposure-outcome, but this does not influence the outcome and exposure measurement.
Exposure assessment	**Probably low**	They assessed multiple greenspace elements (1). Exposure metrics differentiated between types of greenness (1). The exposure was not assessed more than once over time (0). No buffer distances or areas were considered for exposure measurements (0). They assessed the quality of greenness or accessibility (1). They did not assess the time spent in greenspace (0).
Confounding	**Probably low**	The study accounted for only one potential confounder, individual-level socioeconomic status. Model adjusted for age, gender, home location, living arrangement, academic performance and scholarship status.
Incomplete outcome data	**Low**	No missing outcome data.
Selective reporting	**Low**	All outcomes were reported.
Conflict of interest	**Low**	The study did not receive support from a company, study author, or other entity having a financial interest in the outcome of the study as authors make a claim denying conflicts of interest.
Other bias	**Low**	Appears free of other sources of bias.

6. **Narrative justification for (Browning et al., 2021).**

Bias	Rating	Support for judgment
Recruitment	**Low**	In spring 2020, 14,174 participants were recruited cross-sectionally from representative and targeted samples at seven large state universities, which in sum enrolled more than 238,000 students. A total of 2534 responses had received. Participants were recruited from same population at same time frame. Inclusion and exclusion criteria was explained.
Blinding	**Low**	There was no information about blinding for exposure-outcome, but this does not influence the outcome and exposure measurement.

Exposure assessment	**Probably high**	They assessed single greenspace elements (0). Exposure metrics did not differentiate between types of greenness (0). The exposure was not assessed more than once over time (0). No buffer distances or areas were considered for exposure measurements (0). They did not assess the quality of greenness or accessibility (0). They measured the time spent in greenspace (1).
Confounding	**Probably low**	The study accounted for only one potential confounder, individual-level socioeconomic status. Model adjusted for age, gender, race, class, BMI, academic staus and COVID-19 infection status.
Incomplete outcome data	**Low**	No missing outcome data.
Selective reporting	**Low**	All outcomes were reported.
Conflict of interest	**Low**	The study did not receive support from a company, study author, or other entity having a financial interest in the outcome of the study as authors make a claim denying conflicts of interest.
Other bias	**Low**	Appears free of other sources of bias.

7. Narrative justification for (Bu et al., 2021).

Bias	**Rating**	**Support for judgment**
Recruitment	**Low**	The study recruited participants from the UCL COVID-19 study, a large panel study that collected data of over 50,000 adults in the UK between 21 March 2020 and 23 May 2020 during the COVID-19 pandemic. The study was conducted online on a weekly basis. Participants were recruited from same population at same time frame. Inclusion and exclusion criteria was explained clearly.
Blinding	**Low**	There was no information about blinding for exposure-outcome, but this does not influence the outcome and exposure measurement.
Exposure assessment	**High**	They assessed single greenspace elements (0). Exposure metrics did not differentiate between types of greenness (0). The exposure was not assessed more than once over time (0). No buffer distances or areas were considered for exposure measurements (0). They did not assess the quality of greenness or accessibility (0). They measured the time spent in greenspace (1).
Confounding	**Probably low**	The study accounted for only one potential confounder, individual-level socioeconomic status. Model adjusted for gender, age, income, education, ethnicity and area of living.
Incomplete outcome data	**Low**	No missing outcome data.
Selective reporting	**Low**	All outcomes were reported.
Conflict of interest	**Low**	Funding source is limited to government, non-profit organizations, or academic grants funded by government,

		foundations and/or non-profit organizations. However, funder had no role at any stage of the study.
Other bias	**Low**	Appears free of other sources of bias.

8. Narrative justification for (Chen and Liu, 2021).

Bias	**Rating**	**Support for judgment**
Recruitment	**Probably Low**	There is insufficient information about participant selection to permit a judgment of 'YES', but there is indirect evidence that suggests that participant recruitment and inclusion/exclusion criteria was consistent, as described by the criteria for a judgment of 'YES
Blinding	**Low**	There was no information about blinding for exposure-outcome, but this does not influence the outcome and exposure measurement.
Exposure assessment	**High**	They assessed single greenspace elements (0). Exposure metrics did not differentiate between types of greenness (0). The exposure was not assessed more than once over time (0). No buffer distances or areas were considered for exposure measurements (0). They assessed the nearest distance of park (1). They did not measure the time spent in greenspace (0).
Confounding	**Probably High**	The study accounted for both potential confounder, individual and neighborhood level socioeconomic status but included one potential mediator (e.g., air quality index). Over-adjustment of the model have introduced substantial bias.
Incomplete outcome data	**Low**	No evidence of missing outcome data.
Selective reporting	**Low**	All outcomes were reported.
Conflict of interest	**Low**	Funding source is limited to government, non-profit organizations, or academic grants funded by government, foundations and/or non-profit organizations.
Other bias	**Low**	Appears free of other sources of bias.

9. Narrative justification for (Cheng et al., 2021).

Bias	**Rating**	**Support for judgment**
Recruitment	**Probably Low**	There is insufficient information about participant selection to permit a judgment of 'YES', but there is indirect evidence that suggests that participant recruitment and inclusion/exclusion criteria was consistent, as described by the criteria for a judgment of 'YES

Blinding	**Low**	There was no information about blinding for exposure-outcome, but this does not influence the outcome and exposure measurement.
Exposure assessment	**Probably low**	They assessed single greenspace elements (0). Exposure metrics did not differentiate between types of greenness (0). The exposure was assessed more than once over time (1). Multiple buffer distances or areas were considered for exposure measurements (1). They assessed the quality of greenness or accessibility (1). They did not measure the time spent in greenspace (0).
Confounding	**Probably High**	The study accounted for none of our listed potential confounders but used potential mediators, including land Surface temperature and air quality index.
Incomplete outcome data	**Low**	No evidence of missing outcome data.
Selective reporting	**Probably Low**	There is insufficient information about selective outcome reporting to permit a judgment of low risk of bias, but there is indirect evidence that suggests the study was free of selective reporting, as described by the criteria for a judgment of low risk of bias
Conflict of interest	**Low**	The study did not receive support from a company, study author, or other entity having a financial interest in the outcome of the study.
Other bias	**Low**	Appears free of other sources of bias.

10. Narrative justification for (Corley et al., 2020).

Bias	**Rating**	**Support for judgment**
Recruitment	**Probably Low**	This research examined data from 171 individuals (mean age 84 years) from the Lothian Birth Cohort 1936 study who participated in an online survey around two months after the lockdown began (May/June 2020). Participants were recruited from same population at same time frame. Inclusion and exclusion criteria was explained.
Blinding	**Low**	There was no information about blinding for exposure-outcome, but this does not influence the outcome and exposure measurement.
Exposure assessment	**Probably low**	They assessed multiple greenspace elements (1). Exposure metrics did not differentiate between types of greenness (0). The exposure was assessed more than once over time (1). No buffer distances were considered for exposure measurements (1). They assessed the quality of greenness or accessibility (1). They did not measure the time spent in greenspace (0).
Confounding	**Probably Low**	The study accounted for one potential confounder, individual level socioeconomic status. Models were adjusted for sex, living alone, education, occupational social class, anxiety and depressive symptoms, body mass index, and history of diabetes and cardiovascular disease.
Incomplete outcome data	**Low**	No evidence of missing outcome data.

Selective reporting	**Low**	All outcomes were reported.
Conflict of interest	**Low**	Funding source is limited to government, non-profit organizations, or academic grants funded by government, foundations and/or non-profit organizations.
Other bias	**Low**	Appears free of other sources of bias.

11. Narrative justification for (Dzhambov et al., 2020).

Bias	**Rating**	**Support for judgment**
Recruitment	**Low**	This study conducted an online survey among medical students from two universities in the city of Plovdiv, Bulgaria, between 17 May and 10 June, 2020. Participants were recruited from same population at same time frame. Inclusion and exclusion criteria were explained.
Blinding	**Low**	There was no information about blinding for exposure-outcome, but this does not influence the outcome and exposure measurement.
Exposure assessment	**Probably low**	They assessed multiple greenspace elements (1). Exposure metrics differentiated between types of greenness (1). The exposure did not assess more than once over time (0). No buffer distances were considered for exposure measurements (0). They assessed the quality of greenness or accessibility (1). They did not measure the time spent in greenspace (0).
Confounding	**Probably low**	The study accounted for one potential confounder, individual level socioeconomic status. Models were adjusted for age, sex, ethnicity, income, dwelling type, settlement type, university, and connectedness to nature. The study also accounted for potential mediators, including social support using in a separate structure equation modeling.
Incomplete outcome data	**Low**	No evidence of missing outcome data.
Selective reporting	**Low**	All outcomes were reported.
Conflict of interest	**Low**	The study did not receive support from a company, study author, or other entity having a financial interest in the outcome of the study as authors make a claim denying conflicts of interest.
Other bias	**Low**	Appears free of other sources of bias.

12. Narrative justification for (Friedman et al., 2021).

Bias	**Rating**	**Support for judgment**
Recruitment	**Low**	Participants from six countries (Australia, China, Italy, Sweden, UK and USA) participated in i-FAMS-Covid, an online survey that was used to collect data. However, respondents from only United Kingdom were included in this research. Parents with a child between the ages of 3- and 7-years

		old responded to the survey with reference to one target child. Participants were recruited from same population at same time frame. Inclusion and exclusion criteria were explained.
Blinding	**Low**	There was no information about blinding for exposure-outcome, but this does not influence the outcome and exposure measurement.
Exposure assessment	**High**	They did not assess multiple greenspace elements (0). Exposure metrics differentiated between types of greenness (0). The exposure did not assess more than once over time (0). No buffer distances were considered for exposure measurements (0). They did not assess the quality of greenness or accessibility (0). They did not measure the time spent in greenspace (0).
Confounding	**Probably low**	The study accounted for one potential confounder, individual level socioeconomic status.
Incomplete outcome data	**Low**	No evidence of missing outcome data.
Selective reporting	**Low**	All outcomes were reported.
Conflict of interest	**Low**	The authors declare that there is no conflict of interés.
Other bias	**Low**	Appears free of other sources of bias.

13. Narrative justification for (Garrido-Cumbrera et al., 2021).

Bias	**Rating**	**Support for judgment**
Recruitment	**Low**	This study is a part of the GreenCOVID study, which was conducted online to collect data of general population in three European countries (Spain, England and Ireland) at first wave of the pandemic. Participants were recruited from same population. Inclusion and exclusion criteria were explained.
Blinding	**Low**	There was no information about blinding for exposure-outcome, but this does not influence the outcome and exposure measurement.
Exposure assessment	**High**	They assessed single greenspace elements, including nature sound (0). Exposure metrics did not differentiate between types of greenness (e.g.: forests, urban parks, gardens, trees, grassland, bushes, etc.) (0). The exposure was assessed more than once over time (1). No buffer distances were considered for exposure measurements (0). They did not assess the quality of greenness or accessibility (0). They did not measure the time spent in greenspace (0).
Confounding	**Probably low**	The study accounted for one potential confounder, individual level socioeconomic status. Models were adjusted for age, sex, education, employment status, alcohol consumption, smoking and urbanization degree.
Incomplete outcome data	**Low**	No evidence of missing outcome data.

Selective reporting	**Low**	All outcomes were reported.
Conflict of interest	**Low**	The study did not receive support from a company, study author, or other entity having a financial interest in the outcome of the study as authors make a claim denying conflicts of interest.
Other bias	**Low**	Appears free of other sources of bias.

14. Narrative justification for (Gola et al., 2021).

Bias	**Rating**	**Support for judgment**
Recruitment	**Probably low**	There is insufficient information about participant selection to permit a judgment of 'Yes', but there is indirect evidence that suggests that participant recruitment or inclusion/exclusion criteria were consistent, as described by the criteria for a judgment of 'Yes'.
Blinding	**Low**	There was no information about blinding for exposure-outcome, but this does not influence the outcome and exposure measurement.
Exposure assessment	**Probably low**	They assessed multiple greenspace elements (1). Exposure metrics differentiate between types of greenness (1). The exposure did not assess more than once over time (0). No buffer distances were considered for exposure measurements (0). They assessed the quality of greenness or accessibility (1). They did not measure the time spent in greenspace (0).
Confounding	**High**	The study did not account for any potential confounder.
Incomplete outcome data	**Low**	No evidence of missing outcome data.
Selective reporting	**Low**	All outcomes were reported.
Conflict of interest	**Low**	The study did not receive support from a company, study author, or other entity having a financial interest in the outcome of the study as authors make a claim denying conflicts of interest.
Other bias	**Low**	Appears free of other sources of bias.

15. Narrative justification for (Groot et al., 2021).

Bias	**Rating**	**Support for judgment**
Recruitment	**Probably Low**	The Danish National Birth Cohort (DNBC) consists of mothers and offspring from approximately 100 000 pregnancies enrolled in the cohort during the years 1996 to 2002. This cohort did follow-up data collection after offspring turned 7, 11 and 18 years of old. As a part of it, this current study

		collected data from the offspring during the third week of Danish COVID-19 lockdown. There is insufficient information about participant selection to permit a judgment of 'YES', but there is indirect evidence that suggests that participant recruitment and inclusion/exclusion criteria was consistent, as described by the criteria for a judgment of 'YES'.
Blinding	**Low**	There was no information about blinding for exposure-outcome, but this does not influence the outcome and exposure measurement.
Exposure assessment	**Probably low**	They assessed multiple greenspace elements, including having direct access to a common yard only, a balcony only, a garden only, or multiple outdoor spaces/other (1). Exposure metrics differentiated between types of greenness (1). The exposure did not assess more than once over time (0). No buffer distances were considered for exposure measurements (0). They assessed the quality of greenness or accessibility (1). They did not measure the time spent in greenspace (0).
Confounding	**Probably low**	The study accounted for one potential confounder, individual level socioeconomic status, sex, current educational enrolment, part-time work, moving during the lockdown and geographical región.
Incomplete outcome data	**Low**	No evidence of missing outcome data.
Selective reporting	**Low**	All outcomes were reported.
Conflict of interest	**Low**	Funding source is limited to government, non-profit organizations, or academic grants funded by government, foundations and/or non-profit organizations. Funders had no role in study design.
Other bias	**High**	The authors reported that misclassification of housing conditions may have influenced the results.

16. Narrative justification for (Hansmann et al., 2021).

Bias	**Rating**	**Support for judgment**
Recruitment	**Low**	A cross-sectional online study was conducted from 8 April-10 May, 2020 and collected data from 5932 of adult respondents. Participants were recruited from the same population at same time frame. Inclusion and exclusion criteria was explained.
Blinding	**Low**	There was no information about blinding for exposure-outcome, but this does not influence the outcome and exposure measurement.
Exposure assessment	**High**	They assessed single greenspace elements, including access to outdoor spaces (0). Exposure metrics did not differentiate between types of greenness (0). The exposure did not assess more than once over time (0). No buffer distances were considered for exposure measurements (0). They assessed the quality of greenness or accessibility (1). They did not measure the time spent in greenspace (0).
Confounding	**Probably low**	The study accounted for one potential confounder, individual level socioeconomic status, age, sex, education level partnership status, and employment status at the time of the survey.
Incomplete outcome data	**Low**	No evidence of missing outcome data.

Selective reporting	**Low**	All outcomes were reported.
Conflict of interest	**Low**	Funding source is limited to government, non-profit organizations, or academic grants funded by government, foundations and/or non-profit organizations. Funders had no role in study design.
Other bias	**Low**	Appears free of other sources of bias.

17. Narrative justification for (Heo et al., 2021).

Bias	**Rating**	**Support for judgment**
Recruitment	**Low**	Data were collected through an online anonymous survey that was conducted for adults age 19 or older living in South Korea recruited from social media platforms between 21 September and 7 December 2020. Participants were recruited from the same population at the same time frame. Inclusion and exclusión criteria were explained clearly.
Blinding	**Low**	There was no information about blinding for exposure-outcome, but this does not influence the outcome and exposure measurement.
Exposure assessment	**Probably Low**	They assessed multiple greenspace elements (1). Exposure metrics differentiated between types of greenness (1). The exposure was assessed more than once over time (1). No buffer distances were considered for exposure measurements (0). They did not assess the quality of greenness or accessibility (0). They did not measure the time spent in greenspace (0).
Confounding	**Probably low**	The study accounted for one potential confounder, individual level socioeconomic status such as age, gender, smoking, marital status, depression in the year 2019, anxiety in the past 12 months, frequency of visiting greenspace in the year 2019, job-related or financial concerns, health-related concerns, BMI
Incomplete outcome data	**Low**	No evidence of missing outcome data.
Selective reporting	**Low**	All outcomes were reported.
Conflict of interest	**Low**	Funding source is limited to government, non-profit organizations, or academic grants funded by government, foundations and/or non-profit organizations. Funders had no role in study design.
Other bias	**Low**	Appears free of other sources of bias.

18. Narrative justification for (Hubbard et al., 2021).

Bias	**Rating**	**Support for judgment**
Recruitment	**Low**	Data were collected from CHARIS Project, a serial, weekly, nationally representative, cross-sectional, observational study of randomly selected adults in Scotland. Participants were recruited from same population at same time frame. Eligibility criteria were clearly explained.
Blinding	**Low**	There was no information about blinding for exposure-outcome, but this does not influence the outcome and exposure measurement.

Exposure assessment	**Low**	They assessed multiple greenspace elements (1). Exposure metrics differentiated between types of greenness (1). The exposure assessed more than once over time (1). No buffer distance was considered for exposure measurements (0). They assessed the quality of greenness or accessibility (1). They measured the time spent in greenspace (1).
Confounding	**Probably high**	The study accounted for one potential confounder, individual level socioeconomic status, age, gender, shielding category of government, COVID-19 illness representation and threat perception.
Incomplete outcome data	**Low**	Few outcome data are missing that addressed adequately.
Selective reporting	**Low**	All outcomes were reported.
Conflict of interest	**Low**	Funding source is limited to government, non-profit organizations, or academic grants funded by government, foundations and/or non-profit organizations. Funders had no role in study design.
Other bias	**Low**	Appears free of other sources of bias.

19. Narrative justification for (Huerta and Utomo, 2021).

Bias	**Rating**	**Support for judgment**
Recruitment	**Low**	Data was collected through a online survey from May 26 to June 15, 2020 (20 days) through targeted advertisements in Mexico. Targeted participants were adults (18 years and older) whose profile indicated that they lived in one of the 16 municipalities of Mexico City. Participants were recruited from same population at same time frame. Eligibility criteria was nicely explained.
Blinding	**Low**	There was no information about blinding for exposure-outcome, but this does not influence the outcome and exposure measurement.
Exposure assessment	**Probably low**	They assessed multiple greenspace elements (1). Exposure metrics differentiated between types of greenness (1). The exposure assessed more than once over time (1). No buffer distances were considered for exposure measurements (0). They assessed the quality of greenness or accessibility (1). They did not measure the time spent in greenspace (0).
Confounding	**Probably low**	The study accounted for one potential confounder, individual level socioeconomic status, age, gender, income, employment status and municipal features.
Incomplete outcome data	**Low**	No evidence of missing outcome data.
Selective reporting	**Low**	All outcomes were reported.
Conflict of interest	**Low**	Funding source is limited to government, non-profit organizations, or academic grants funded by government, foundations and/or non-profit organizations. Funders had no role in study design.

Other bias	**Low**	Appears free of other sources of bias.

20. Narrative justification for (Jackson et al., 2021a).

Bias	Rating	Support for judgment
Recruitment	**Low**	Data was collected through a online survey from 30 April 2020 and closed 15 June 2020 in USA. The targeted participants were adolescents (18 years and older). Participants were recruited from same population at the same time frame. Eligibility criteria was nicely explained.
Blinding	**Low**	There was no information about blinding for exposure-outcome, but this does not influence the outcome and exposure measurement.
Exposure assessment	**Probably low**	They assessed multiple greenspace elements (1). Exposure metrics differentiated between types of greenness (1). The exposure assessed more than once over time (1). No buffer distances were considered for exposure measurements (0). They did not assess the quality of greenness or accessibility (0). They did not measure the time spent in greenspace (0).
Confounding	**Probably high**	The study accounted for one potential confounder, individual level socioeconomic status, age, gender, race, community type, household income and región.
Incomplete outcome data	**Low**	No evidence of missing outcome data.
Selective reporting	**Low**	All outcomes were reported.
Conflict of interest	**Low**	The study did not receive support from a company, study author, or other entity having a financial interest in the outcome of the study as authors make a claim denying conflicts of interest.
Other bias	**Low**	Appears free of other sources of bias.

21. Narrative justification for (Jackson et al., 2021b).

Bias	Rating	Support for judgment
Recruitment	**Low**	Data was collected through a online survey from 30 April 2020 and closed 15 June 2020 in the USA. Targeted participants were adolescents (18 years and older). Participants were recruited from same population at same time frame. Eligibility criteria was nicely explained.
Blinding	**Low**	There was no information about blinding for exposure-outcome, but this does not influence the outcome and exposure measurement.
Exposure assessment	**Probably low**	They assessed multiple greenspace elements (1). Exposure metrics differentiated between types of greenness (1). The exposure assessed more than once over time (1). No buffer distances were considered for exposure measurements (0). They did not assess the quality of greenness or accessibility (0).

		They did not measure the time spent in greenspace (0).
Confounding	**Probably high**	The study accounted for one potential confounder, individual level socioeconomic status, age, gender, race, community type, household income and región.
Incomplete outcome data	**Low**	No evidence of missing outcome data.
Selective reporting	**Low**	All outcomes were reported.
Conflict of interest	**Low**	The study did not receive support from a company, study author, or other entity having a financial interest in the outcome of the study as authors make a claim denying conflicts of interest.
Other bias	**Low**	Appears free of other sources of bias.

22. Narrative justification for (Jato-Espino et al., 2021).

Bias	**Rating**	**Support for judgment**
Recruitment	**Probably low**	Data was collected through an online survey from 2020-04-21 to 2020-05-05 in two áreas of Spain. Targeted participants were adolescents (16 years and older). However, there is insufficient information about participant selection to permit a judgment of 'YES', but there is indirect evidence that suggests that participant recruitment and inclusion/exclusion criteria was consistent, as described by the criteria for a judgment of 'YES'.
Blinding	**Low**	There was no information about blinding for exposure-outcome, but this does not influence the outcome and exposure measurement.
Exposure assessment	**Low**	They assessed multiple greenspace elements (1). Exposure metrics differentiated between types of greenness (1). The exposure assessed more than once over time (1). Buffer distances were considered for exposure measurements (1). They assessed the quality of greenness or accessibility (1). They did not measure the time spent in greenspace (0).
Confounding	**Probably high**	The study accounted for one potential confounder, individual level socioeconomic status, age, gender, and people per household.
Incomplete outcome data	**Low**	No evidence of missing outcome data.
Selective reporting	**Low**	All outcomes were reported.
Conflict of interest	**Low**	The study did not receive support from a company, study author, or other entity having a financial interest in the outcome of the study as authors make a claim denying conflicts of interest.
Other bias	**Low**	Appears free of other sources of bias.

23. Narrative justification for (Kontsevaya et al., 2020)

Bias	Rating	Support for Judgement
Recruitment	**Low**	Data were collected from a cross-sectional online study conducted during the lockdown period between 26 April 2020 and 6 June 2020 in 62 region of Russia. A total of 2432 participants were included in the study. Eligible participants included all Russian adults aged 18 years and over. Participants were recruited from the same population at the same time frame. Eligibility criteria were explained clearly.
Blinding	**Low**	There was no information about blinding for exposure-outcome, but this does not influence the outcome and exposure measurement.
Exposure assessment	**Probably high**	They assessed multiple greenspace elements, including access to outdoor space and access to greenspace (1). Exposure metrics did not differentiate between types of greenness (0). The exposure did not assess more than once over time (0). No buffer distances were considered for exposure measurements (0). They assessed the quality of greenness or accessibility (1). They did not measure the time spent in greenspace (0).
Confounding	**Probably low**	The study accounted for one potential confounder, individual level socioeconomic status, age, sex.
Incomplete outcome data	**Low**	No evidence of missing outcome data.
Selective reporting	**Low**	All outcomes were reported.
Conflict of interest	**High**	The study acknowledged company staff and ministry of health staff for their involvement in the survey development.
Other bias	**Low**	Appears free of other sources of bias.

24. Narrative justification for (Kou et al., 2021).

Bias	Rating	Support for Judgement
Recruitment	**Probably low**	There is insufficient information about participant selection to permit a judgment of 'Yes', but here is indirect evidence that suggests that participant recruitment and inclusion/exclusion criteria was consistent, as described by the criteria for a judgment 'Yes'.
Blinding	**Low**	There was no information about blinding for exposure-outcome, but this does not influence the outcome and exposure measurement.
Exposure assessment	**High**	They assessed a single greenspace element (0). Exposure metrics did not differentiate between types of greenness (0). The exposure did not assess more than once over time (0). No buffer distances were considered for exposure measurements (0). They did not assess the quality of greenness or accessibility (0). They did not measure the time spent in greenspace (0).

Confounding	**High**	The study did not account for any potential confounder.
Incomplete outcome data	**Low**	No evidence of missing outcome data.
Selective reporting	**Low**	All outcomes were reported.
Conflict of interest	**High**	The author declares no conflict of interest. This study was funded by the government organization of China. However, the study acknowledged that two external authors who were involved in the language editing and manuscript reviewing.
Other bias	**Low**	Appears free of other sources of bias.

25. Narrative justification for (Lades et al., 2020).

Bias	**Rating**	**Support for Judgement**
Recruitment	**Low**	Data were collected from a cross-sectional online study conducted on March 25, 2020 public of Ireland. A total of 604 participants were included in the study. Eligible participants included adults aged 18 years and over. Participants were recruited from the same population at the same time frame. Eligibility criteria were explained clearly.
Blinding	**Low**	There was no information about blinding for exposure-outcome, but this does not influence the outcome and exposure measurement.
Exposure assessment	**High**	They assessed single greenspace element (0). Exposure metrics did not differentiate between types of greenness (0). The exposure did not assess more than once over time (0). No buffer distances were considered for exposure measurements (0). They did not assess the quality of greenness or accessibility (0). They did not measure the time spent in greenspace (0).
Confounding	**Probably Low**	The study accounted for only one potential confounder, individual level socioeconomic statuses such as age, sex, education, region, employment status and household income.
Incomplete outcome data	**Low**	No evidence of missing outcome data.
Selective reporting	**Low**	All outcomes were reported.
Conflict of interest	**Low**	Funding source is limited to government, non-profit organizations, or academic grants funded by government, foundations and/or non-profit organizations. Funders had no role in study design.
Other bias	**Low**	Appears free of other sources of bias.

26. Narrative justification for (Larson et al., 2021).

Bias	Rating	Support for Judgement
Recruitment	**Low**	In spring 2020, 10,195 participants were recruited cross-sectionally from representative and targeted samples at four large, state universities. A total of 1280 responses had received. Participants were recruited from same population at same time frame. Inclusion and exclusion criteria was clearly explained.
Blinding	**Low**	There was no information about blinding for exposure-outcome, but this does not influence the outcome and exposure measurement.
Exposure assessment	**Probably low**	They assessed multiple greenspace elements, including (1). Exposure metrics differentiated between types of greenness (1). The exposure did not assess more than once over time (0). Buffer distances or areas were considered for exposure measurements (1). They assessed the quality of greenness or accessibility (1). They did not measure the time spent in greenspace (0).
Confounding	**Low**	The study accounted for both potential confounders, individual and neighborhood level socioeconomic status, person-level (relative in-come, race, sex, and age, general health, BMI, knowing people infected by COVID-19, and worrying about COVID-19) and area-level correlates (urbanity, area of local parks per capita, area of state/national parks per capita, NDVI, university where each student was enrolled, COVID-19 death rates, and COVID-19-related lockdown levels).
Incomplete outcome data	**Low**	No evidence of missing outcome data.
Selective reporting	**Low**	All outcomes were reported.
Conflict of interest	**Low**	The study did not receive support from a company, study author, or other entity having a financial interest in the outcome of the study as authors make a claim denying conflicts of interest.
Other bias	**Low**	Appears free of other sources of bias.

27. Narrative justification for (Lee et al., 2021).

Bias	Rating	Support for Judgement
Recruitment	**Low**	An onsite cross-sectional study was conducted among adults aged (19 years or older) who visited the forest in South Korea from 1 May to 15 July 2020. Participants were recruited from same population at same time frame. Inclusion and exclusion criteria was clearly explained.
Blinding	**Low**	There was no information about blinding for exposure-outcome, but this does not influence the outcome and exposure measurement.
Exposure assessment	**Probably low**	They assessed multiple greenspace elements, including (1). Exposure metrics did not differentiate between types of greenness (0). The exposure did not assess more than once over time (0). No buffer distances or areas were considered for exposure measurements (0). They assessed the quality of greenness or accessibility (1).

		They measured the time spent in greenspace (1).
Confounding	**Probably low**	The study accounted for one potential confounder, individual level socioeconomic status, age, sex.
Incomplete outcome data	**Low**	No evidence of missing outcome data.
Selective reporting	**Low**	All outcomes were reported.
Conflict of interest	**Low**	The study did not receive support from a company, study author, or other entity having a financial interest in the outcome of the study as authors make a claim denying conflicts of interest.
Other bias	**Low**	Appears free of other sources of bias.

28. Narrative justification for (Lehberger et al., 2021).

Bias	**Rating**	**Support for Judgement**
Recruitment	**Low**	An online cross-sectional study was conducted among adults aged (18-65) in Germany during May 14 and 24, 2020. Participants were recruited from same population at same time frame. Inclusion and exclusion criteria was clearly explained.
Blinding	**Low**	There was no information about blinding for exposure-outcome, but this does not influence the outcome and exposure measurement.
Exposure assessment	**Probably low**	They assessed multiple greenspace elements (1). Exposure metrics differentiated between types of greenness (1). The exposure assessed more than once over time (1). No buffer distances or area were considered for exposure measurements (0). They did not assess the quality of greenness or accessibility (0). They measured the time spent in greenspace (1).
Confounding	**Probably low**	The study accounted for one potential confounder, individual level socioeconomic status, age, gender, income, employment, urbanicity, living alone, fear of job loss, garden, change in time spent outdoors.
Incomplete outcome data	**Low**	No evidence of missing outcome data.
Selective reporting	**Low**	All outcomes were reported.
Conflict of interest	**Low**	Funding source is limited to government, non-profit organizations, or academic grants funded by government, foundations and/or non-profit organizations. Funders had no role in study design.
Other bias	**Low**	Appears free of other sources of bias.

29. Narrative justification for (Lenaerts et al., 2021).

Bias	Rating	Support for Judgement
Recruitment	**Probably low**	There is insufficient information about participant selection to permit a judgment of 'Yes', but here is indirect evidence that suggests that participant recruitment and inclusion/exclusion criteria were consistent, as described by the criteria for a judgment 'Yes'.
Blinding	**Low**	There was no information about blinding for exposure-outcome, but this does not influence the outcome and exposure measurement.
Exposure assessment	**Probably low**	They assessed multiple greenspace elements (1). Exposure metrics differentiated between types of greenness (1). The exposure did not assess more than once over time (0). No buffer distances or area were considered for exposure measurements (0). They assessed the quality of greenness or accessibility (1), They did not measure the time spent in greenspace (0).
Confounding	**Probably low**	The study accounted for one potential confounder, individual level socioeconomic status, gender, age, educational attainment, mental and physical health, private garden, satisfaction with home and living environment, and sufficient green space in the living environment.
Incomplete outcome data	**Low**	No evidence of missing outcome data.
Selective reporting	**Low**	All outcomes were reported.
Conflict of interest	**Low**	Funding source is limited to government, non-profit organizations, or academic grants funded by government, foundations and/or non-profit organizations. Funders had no role in study design.
Other bias	**Low**	Appears free of other sources of bias.

30. Narrative justification for (Lõhmus et al., 2021).

Bias	Rating	Support for judgment
Recruitment	**Low**	An online study was carried out in Stockholm, Sweden, focused on adult residents (aged 20 and above). The survey was conducted from 5th June 2020 to 1st August 2020 and all the study samples (2060) were recruited from same population as well as same time frame. Eligibility criteria were clearly explained.
Blinding	**Low**	There was no information about blinding for exposure-outcome, but this does not influence the outcome and exposure measurement.
Exposure assessment	**Low**	Assessed multiple exposure metrics (1). Exposure metrics did not differentiate between types of greenness (0). The study measured the exposures more than once (1). Multiple buffer distances or areas were considered for exposure measurements (1). They assessed the quality of greenness or accessibility (1). They measured the use of nature (1).

Confounding	**Probably High**	The study accounted for both potential confounders, individual-level and neighborhood-level socioeconomic status (e.g., age, sex, income, and the neighborhood walkability index), but also included potential mediators, alcohol consumption and physical inactivity, in the model. Overadjustment of the model may have introduced substantial bias.
Incomplete outcome data	**Low**	A decrease was noticed in the total number of data that may refer to missing outcome data. However, the missing data did not affect the outcome assessment, which might be negligible.
Selective reporting	**Low**	All outcomes were reported.
Conflict of interest	**Low**	This study was financed by the Centrum for Occupational and Environmental Medicine, Region Stockholm as well as by the Swedish Research Council for Sustainable Development, but they have no influence on this study as the authors make a claim denying conflicts of interest.
Other bias	**Low**	Appears free of other sources of bias.

31. Narrative justification for (Marques et al., 2021).

Bias	**Rating**	**Support for judgment**
Recruitment	**Low**	An online study was carried out in Rio de Janeiro, Brazil, focused on 173 adults (respondents were of legal age 18+) who were living in the city of Rio de Janeiro. The survey was conducted in between November 6th 2020 and January 24th, 2021 and all the study samples were recruited from same time frame. Eligibility criteria were explained clearly.
Blinding	**Low**	There was no information about blinding for exposure-outcome, but this does not influence the outcome and exposure measurement.
Exposure assessment	**Probably Low**	They assessed multiple green elements (1). Exposure metrics differentiated between types of greenness (1). The exposure was assessed for once only (0). No buffer distances or areas were considered for exposure measurements (0). They assessed accessibility of greenness or accessibility (1). They measured the use of green space (1).
Confounding	**Probably Low**	The studied model adjusted for only one potential confounder factor, individual-level socioeconomic variables, such as gender, distance from parks, non-essential activities, and plants taken care of.
Incomplete outcome data	**Low**	No missing outcome data
Selective reporting	**Low**	All outcomes were reported.

Conflict of interest	**Low**	The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.
Other bias	**Low**	Appears free of other sources of bias.

32. Narrative justification for (Mead et al., 2021).

Bias	**Rating**	**Support for judgment**
Recruitment	**Low**	All participants (138 responses) were recruited via advertisements on social media platforms and an internal departmental advertisement site. All are UK residents. The online survey was conducted in same time period, 8th April to 23rd May 2020. Participants were recruited from the same population at the same time frame. Eligibility criteria were explained clearly.
Blinding	**Low**	There was no mention of blinding, however, this does not influence the outcome and exposure measurement.
Exposure assessment	**High**	They assessed a single greenspace element (0). Exposure metrics did not differentiate between types of greenness (0). The exposure was not assessed more than once over time (0). No buffer distances or areas were considered for exposure measurements (0). They did not assess the quality of greenness or accessibility (0). They did not measure the use of green space or exposure time (0).
Confounding	**Probably Low**	The study accounted for individual-level socioeconomic status. Models adjusted for socioeconomic status (SES), age, and gender.
Incomplete outcome data	**Low**	No missing outcome data.
Selective reporting	**Low**	All outcomes were reported.
Conflict of interest	**Probably High**	There is insufficient information to permit a judgment of high risk of bias, as the authors did not clarify their conflict of interest. But there is indirect evidence that suggests the study was not free of support from a company. The study acknowledged one of the corresponding authors, due to getting financial support from industry, might have a financial interest in the outcome of the study, as described by the criteria for a judgment of high risk of bias.
Other bias	**Low**	Appears free of other sources of bias.

33. Narrative justification for (Millán-Jiménez et al., 2021).

Bias	Rating	Support for judgment
Recruitment	**Low**	This study includes 188 young final-year medicine and architecture undergraduate students from the University of Seville, Spain, and all the participants were recruited from same population and same time frame. Eligibility criteria were explained clearly.
Blinding	**Low**	There was no mention of blinding, however, this does not influence the outcome and exposure measurement.
Exposure assessment	**Probably Low**	They assessed multiple greenspace elements (1). Those exposure metrics differentiate between types of greenness (1). The exposure was not assessed more than once over time (0). No buffer distances or areas were considered for exposure measurements (0). They assessed the greenness or accessibility (1). They did not measure the use of green space or exposure time (0).
Confounding	**High**	The study did not account for neither confounders nor mediators.
Incomplete outcome data	**Low**	No missing outcome data.
Selective reporting	**Low**	All outcomes were reported.
Conflict of interest	**Low**	The authors declared no conflict of interest.
Other bias	**Low**	Appears free of other sources of bias.

34. Narrative justification for (Mintz et al., 2021).

Bias	Rating	Support for judgment
Recruitment	**Low**	The data collection for the study was carried out at the end of April 2020 and and limited to people (aged 18 or above) who were the members of a panel company located in Israel. A total number of 776 respondents were recruited for this survey from same population with same time frame. Eligibility criteria were explained clearly.
Blinding	**Low**	There was no mention of blinding, however, this does not influence the outcome and exposure measurement.
Exposure assessment	**Probably Low**	They assessed multiple greenspace elements (1). The metrics differentiated between types of greenness (1). Exposure was not assessed more than once over time (0). No buffer distances or areas were considered for exposure measurements (0). They assessed the quality of greenness or accessibility (1). They did not measure the use of green space or exposure time (0).

Confounding	**Probably Low**	The studied accounted for one of the potential confounders, individual level socioeconomic indicators. Model adjusted for age, gender, cultural and ethnic group (Jewish/Arab), and the extent the pandemic affected participants' income.
Incomplete outcome data	**Low**	No missing outcome data.
Selective reporting	**Low**	All outcomes were reported.
Conflict of interest	**Low**	The study was funded by Israel Ministry of Science and Technology; however, the authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.
Other bias	**Low**	Appears free of other sources of bias.

35. Narrative justification for (Okely et al., 2021).

Bias	**Rating**	**Support for judgment**
Recruitment	**Low**	A longitudinal study by interviewing the parents of children aged 3–5 years, from 14 countries (8 low- and middle-income countries, LMICs) to assess children's changes in movement behaviors and associated with the COVID-19 pandemic. The survey was conducted in two time periods (12 months up to March 2020 and again between May and June 2020) and all the study samples (about 948) were considered from same population. For 86% of the sample, the same primary parent-reported data at both time points.
Blinding	**Low**	There was no information about blinding for exposure-outcome, but this does not influence the outcome and exposure measurement.
Exposure assessment	**Probably Low**	They assessed multiple greenspace elements (1). Exposure metrics did not differentiate between types of greenness (0). The exposure was not assessed for two times in different two time periods (0). No buffer distances or areas were considered for exposure measurements (0). They assessed the quality of greenness or accessibility (1). They measured the exposure time (1).
Confounding	**High**	The study did not account for any potential confounders.
Incomplete outcome data	**Low**	No missing outcome data.
Selective reporting	**Low**	All outcomes were reported.
Conflict of interest	**Low**	The first author of study was supported by NHMRC Investigator Gran, a government research funding organization, but they have no influence on this study as the authors make a claim denying conflicts of interest.
Other bias	**Low**	Appear free of other sources of bias.

36. Narrative justification for (Olszewska-Guizzo et al., 2021).

Bias	Rating	Support for judgment
Recruitment	**Low**	An online and on-site experimental study were carried out in Singapore, focused on 25 adults (aged between 21 and 74 years) who attended two electroencephalography (EEG) lab sessions. The survey was conducted in two time periods (prior to the pandemic (T1) by January 20, 2020, and again T2 by June 1, 2020) and all the study samples were recruited from same population.
Blinding	**Low**	There was no information about blinding for exposure-outcome, but this does not influence the outcome and exposure measurement.
Exposure assessment	**Probably High**	They assessed multiple greenspace elements (1). Exposure metrics did not differentiate between types of greenness (0). The exposure did not assess more than once (0). No buffer distances or areas were considered for exposure measurements (0). They did not assess the quality of greenness or accessibility (0). They measured the use of green space or exposure time (1).
Confounding	**High**	The study did not account for any potential confounders. However, the study accounted for two moderators: degree of nature exposure and landscape types.
Incomplete outcome data	**Low**	No missing outcome data
Selective reporting	**Low**	All outcomes were reported.
Conflict of interest	**Probably Low**	Funding source was limited to the Singapore MND grant and one member was acknowledged for consolidating the supplementary materials but not influencing the design, data analysis and writing sections. There was indirect evidence on no conflict of financial interest that revealed probably low risk of bias.
Other bias	**Low**	Appears free of other sources of bias.

37. Narrative justification for (Oswald et al., 2021).

Bias	Rating	Support for judgment
Recruitment	**Low**	All participants (1004 responses) were recruited based on several inclusion criteria: young Australians living in metropolitan areas, aged between 18 and 24 years, and proficient in English. The online survey was conducted in same time period, 17th of November 2020 to the 9th of January 2021.
Blinding	**Low**	There was no mention of blinding, however, this does not influence the outcome and exposure measurement. The survey link was disseminated by Qualtrics, where no chance to know the residence address prior to the survey.

Exposure assessment	**Probably Low**	They assessed multiple greenspace elements (1). Exposure metrics did not differentiate between types of greenness (0). The exposure was not assessed more than once over time (0). No buffer distances or areas were considered for exposure measurements (0). They assessed the quality of greenness or accessibility (1). They measured the use of green space or exposure time (1).
Confounding	**Probably low**	The study accounted for one potential confounder, individual-level socioeconomic status. Models adjusted for sex, age, education, residence, types of residential dwelling, and Area-Level Socioeconomic Status Quintile. In additional models, they also accounted for other nature variables as mediators.
Incomplete outcome data	**Low**	No missing outcome data
Selective reporting	**Low**	All outcomes were reported.
Conflict of interest	**Low**	The study was supported by Australian Commonwealth Research Training Program Scholarship and the Ian Wilson Liberal Research Scholarship, however, authors declared that there was no conflict of interest.
Other bias	**Low**	Appears free of other sources of bias.

38. Narrative justification for (Pearson et al., 2021).

Bias	**Rating**	**Support for judgment**
Recruitment	**Low**	This study includes 56 participants (18 years or older) diagnosed with breast cancer in US, accessed via Research Match. Participants were recruited from the same population at the same time frame.
Blinding	**Low**	There was no mention of blinding, however, this does not influence the outcome and exposure measurement.
Exposure assessment	**Low**	They assessed multiple green spaces (1). Those exposure metrics differentiate between types of greenness (1). The exposure was assessed for two-time points (1). No buffer distances or areas were considered for exposure measurements (0). They assessed the quality of greenness or accessibility (1). They measured the use of green space (1).
Confounding	**Probably Low**	The study accounted for one potential confounder, individual-level socioeconomic status. Models were adjusted for age, marital status, income, and stage at diagnosis.
Incomplete	**Low**	No missing outcome data.

outcome data		
Selective reporting	**Low**	All outcomes were reported.
Conflict of interest	**Low**	The study did not receive support from a company, study author, or other entity having a financial interest in the outcome of the study as authors make a claim denying conflicts of interest.
Other bias	**Low**	Appears free of other sources of bias.

39. Narrative justification for (P´erez-Urrestarazu et al., 2020).

Bias	**Rating**	**Support for judgment**
Recruitment	**Low**	The data collection for the study was carried out between 25th April 2020 and 5th May 2020 and limited to people aged 18 or above. A total number of 4205 respondents were recruited for this survey from 46 different countries. However, the proportion of response was enough to be statistically significant considering the population in those countries, with a confidence level of 95% and a margin of error ≤5%.
Blinding	**Low**	There was no mention of blinding, however, this does not influence the outcome and exposure measurement.
Exposure assessment	**Probably High**	They assessed multiple greenspace elements. (1) The metrics did not differentiate between types of greenness (0). Exposure was not assessed more than once over time (0). No buffer distances or areas were considered for exposure measurements (0). They did not assess the quality of greenness or accessibility (0). They measured the use of green space (1).
Confounding	**High**	The study did not account for any potential confounders.
Incomplete outcome data	**Low**	No missing outcome data.
Selective reporting	**Low**	All outcomes were reported.
Conflict of interest	**Low**	The study was supported by Capes PrInt Brazil; however, the authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.
Other bias	**Low**	Appears free of other sources of bias.

40. Narrative justification for (Poortinga et al., 2021).

Bias	Rating	Support for judgment
Recruitment	**Low**	The mixed-method study recruited 5566 participants aged 18 or above through mostly Health Wise Wales (HWW) and online. This longitudinal study was carried out in two tome periods (13 March and 14 April 2020; 20 June and 20 July 2020). All the participants were recruited during mentioned time frame from the UK.
Blinding	**Low**	There was no mention of blinding, however, this does not influence the outcome and exposure measurement.
Exposure assessment	**Probably Low**	They assessed multiple green space elements (1). Those exposure metrics differentiate between types of greenness (1). The exposure was assessed for more than once (1). No buffer distances or areas were considered for exposure measurements (0). They assessed the accessibility to green space (1). They did not assess the use of green space and exposure time (0).
Confounding	**Probably Low**	The study accounted for only one potential confounder, individual level socioeconomic variables. Models were adjusted for age, sex, employment status, and marital status.
Incomplete outcome data	**Low**	No missing outcome data.
Selective reporting	**Low**	All outcomes were reported.
Conflict of interest	**Probably Low**	The study was supported by a grant from Sˆer Cymru and facilitated by Health Wise Wales, the Heath and Care Research Wales initiative, which is led by Cardiff University in collaboration with SAIL, Swansea University. There was no clear evidence of financial interest in the outcome of the study involved in the design, conduct, analysis, or reporting of the study, and authors had complete access to the data.
Other bias	**Low**	Appears free of other sources of bias.

41. Narrative justification for (Pouso et al., 2020).

Bias	Rating	Support for judgment
Recruitment	**Probably Low**	Study participants were recruited from different countries (77 countries) via online self-reported survey, though the researchers considered the countries presenting at least 100 responses to avoid heterogeneity. Survey responses (5218) were retrieved from 17th April to 8th May 2020, a same time frame.

Blinding	**Low**	There was no mention of blinding, however, this does not influence the outcome and exposure measurement.
Exposure assessment	**Probably Low**	They assessed multiple green space elements (1). Those exposure metrics were differentiated (1). Exposure was not assessed more than once over time (0). No buffer distances or areas were considered for exposure measurements (0). They assessed the accessibility of greenspace (1). They measured the exposure time (1).
Confounding	**Low**	The study accounted for both individual-level and neighborhood-level socioeconomic status as they adjusted models for the following covariates identified a priori: residence, age, gender, education, whether employment situation changed after the coronavirus outbreak, income, whether respondents owned a pet that needed walking outside, residence size and characteristics of the company during the lockdown, such as whether respondent was alone, with kids, and with people with special care needs.
Incomplete outcome data	**Low**	No missing outcome data.
Selective reporting	**Low**	All outcomes were reported.
Conflict of interest	**Low**	Funding source is limited to government, non-profit organizations, or academic grants, however, the authors declared that they had no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.
Other bias	**Low**	Appears free of other sources of bias.

42. Narrative justification for (Riberio et al., 2021).

Bias	**Rating**	**Support for judgment**
Recruitment	**Probably Low**	The cross-sectional online study was conducted between March 27 and May 6, 2020, in two different countries (Spain and Portugal). However, the COVID condition of both countries was same as during this period; both countries were under a strict general lockdown. A total of 3157 participants were recruited from same types of population (aged 18 or above and permanent residents of respected countries) (1).
Blinding	**Low**	There was no mention of blinding, however, this does not influence the outcome and exposure measurement.
Exposure assessment	**Low**	They assessed multiple green space elements (1). Those exposure metrics differentiate between types of greenness (1). The exposure was assessed more than once (1). They did not consider buffer distances for the exposure measurement (0). They assessed quality of greenness or accessibility (1). They assessed the time spent in green spaces (1).
Confounding	**Low**	The study accounted for both individual-level and neighborhood-level socioeconomic status as they adjusted models for the following covariates

		identified a priori: gender, age, residence, education, dependents in households, type of housing stayed during the lockdown, duration of lockdown, outside activities during the lockdown, and frequency with which respondents left their homes during the lockdown.
Incomplete outcome data	**Low**	No evidence of missing outcome data.
Selective reporting	**Low**	All outcomes were reported.
Conflict of interest	**Low**	The authors declared that they had no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.
Other bias	**Low**	Appears free of other sources of bias.

43. Narrative justification for (Robinson et al., 2021).

Bias	**Rating**	**Support for judgment**
Recruitment	**Probably Low**	The non-random sampling was carried out between April and July 2020 and focused on participants aged 18 years or above across the world. Despite having same time frame, insufficient information on the reason for the distribution of participants was confusing. However, authors claimed that recipients were in the UK due to global distribution limitations as well as time limitations that permit a judgment of 'YES', but there is indirect evidence that suggests that participant recruitment and inclusion/exclusion criteria were consistent.
Blinding	**Low**	There was no mention of blinding, however, this does not influence the outcome and exposure measurement.
Exposure assessment	**Probably Low**	They assessed multiple green space elements (1). The exposure metric did not differentiate between types of greenness (0). The exposure was assessed more than once (1). There used different buffer distances for the exposure measurement (1). They did not assess the quality of greenness and accessibility (0). They assessed the time spent visiting green spaces (1).
Confounding	**Probably low**	The study accounted for only one potential confounder, individual-level socioeconomic status. The covariates were: sex, age, socioeconomic status, living/working situation, and level of education.
Incomplete outcome data	**Low**	All subjects had complete outcome data.
Selective reporting	**Low**	All outcomes were reported.
Conflict of interest	**Probably High**	The corresponding author of the study got fund from ESRC for conducting this study. Though the author claimed no conflict of interest, he was an Assistant Guest Editor on this Special Issue that provided indirect evidence suggesting the study design was influenced by funding author.
Other bias	**Low**	Appears free of other sources of bias.

44. Narrative justification for (Samuelsson et al., 2021).

Bias	Rating	Support for judgment
Recruitment	**Low**	Data were collected through an online PPGIS survey of the general population of Sweden. Participants (about 684) were recruited from the same population during the spring of 2020. The inclusion and exclusion criteria were clearly defined.
Blinding	**Low**	There was no mention of blinding, however, this does not influence the outcome and exposure measurement.
Exposure assessment	**Probably Low**	The study assessed multiple green space elements (1). Those exposure metrics differentiated between types of greenness (1). The exposure was not assessed more than once over time (0). Buffer distances or areas were considered for the exposure measurement: 50m (1). They assessed greenspace quality or accessibility (1). They did not assess the time spent in green spaces (0).
Confounding	**Probably Low**	The study accounted for only one potential confounder, individual-level socioeconomic status. The selected confounders included area median income, gender, age, occupation, and whether they were living alone.
Incomplete outcome data	**Low**	No evidence of missing outcome data.
Selective reporting	**Low**	All outcomes were reported.
Conflict of interest	**Probably Low**	There is insufficient information to permit a judgment of low risk of bias, but there is indirect evidence that suggests the study funding did not influence the entire design and outcome of the study.
Other bias	**Low**	Appears free of other sources of bias.

45. Narrative justification for (Samus et al., 2021).

Bias	Rating	Support for judgment

Recruitment	**Probably High**	The data were collected from two different cities in two countries at different time frames in different settings (one situation was the closure of restaurants, retailers, and schools, and the requirement for all except essential workers to remain at home, lifted and the other situation was few days after restaurants reopened, meetings of two households in public were permitted, and a limited number of students were allowed to attend school again). However, participants in both places were recruited only those who had access to a private garden. The entire evidence about participant selection permits a judgment of 'NO', but there is indirect evidence that suggests that participant recruitment or inclusion/exclusion criteria were inconsistent, as described by the criteria for a judgment of 'NO'.
Blinding	**Low**	There was no mention of blinding, however, this does not influence the outcome and exposure measurement.
Exposure assessment	**Probably Low**	The study assessed multiple green space elements (1). Those exposure metrics differentiated between types of greenness (1). The exposure was not assessed more than once over time (0). There were not different buffer distances or areas (e.g., neighborhood, school, home, etc.) considered for the exposure measurement (0). They did not assess the proximity to a major green space (0). They assessed the time spent in green spaces (1).
Confounding	**Probably low**	The study accounted for only one potential confounder, individual-level socioeconomic status, including age, gender, and level of education.
Incomplete outcome data	**Low**	No missing outcome data.
Selective reporting	**Low**	All outcomes were reported.
Conflict of interest	**Low**	The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper. The study did not receive support from a company, study author, or other entity having a financial interest in the outcome of the study.
Other bias	**Low**	Appears free of other sources of bias.

46. Narrative justification for (Sansal et al., 2021).

Bias	**Rating**	**Support for judgment**
Recruitment	**Low**	Snowball sampling method was used on a total of 98 community-dwelling elders residing in ten different cities in Turkey. Among them, nine volunteers failed to meet the inclusion criteria. Participants were recruited from the same population at the same time frame. Eligibility criteria were explained clearly.
Blinding	**Low**	There was no mention of blinding, however, this does not seem to influence the outcome and exposure measurement
Exposure assessment	**High**	They assessed single greenspace variables, including time spent in outdoor space (0). Exposure metrics did not differentiate between types of greenness (e.g., forests, urban parks, gardens, trees, grassland, bushes, etc.) (0). The exposure was not assessed more than once (0).

		Different buffer distances or areas were not considered (0) The quality of greenness was not measured (0). They measured the exposure time of greenspace use (1).
Confounding	**High**	The study did not account for any potential confounders.
Incomplete outcome data	**Probably low**	There is insufficient information about incomplete outcome data to permit a judgment of low risk of bias, but there is indirect evidence that suggests incomplete outcome data was adequately addressed, as they described that to strengthen their analysis, they used the inverse probability weighting methodology to account for the potential sources of bias due to the percentage of children lost to follow-up.
Selective reporting	**Low**	All outcomes were reported.
Conflict of interest	**Low**	The study did not receive support from a company, study author, or other entity having a financial interest in the outcome of the study as authors make a claim denying conflicts of interest.
Other bias	**Low**	Appears free of other sources of bias.

47. Narrative justification for (Sia et al., 2021).

Bias	**Rating**	**Support for judgment**
Recruitment	**Probably Low**	There is insufficient information about participant selection to permit a judgment of 'NO', but there is indirect evidence that suggests that participant recruitment or inclusion/exclusion criteria were inconsistent, as described by the criteria for a judgment of 'NO'.
Blinding	**Low**	There was no mention of blinding, however, this does not seem to influence the outcome and exposure measurement.
Exposure assessment	**High**	One exposure was considered (gardening) (0). Exposure metrics did not differentiate between types of greenness (0). The exposure was not assessed more than once (0). Different buffer distances or areas were not considered (0). The quality of greenness was not measured (0). They measured exposure time for greenspace use (1).
Confounding	**Probably low**	The study accounted for only one potential confounder, individual-level socioeconomic status. Model was adjusted for frequency of smoking, alcohol use, age, gender, and annual household.
Incomplete outcome data	**Low**	No missing outcome data.
Selective reporting	**Low**	All outcomes were reported.

Conflict of interest	**Low**	The study did not receive support from a company, study author, or other entity having a financial interest in the outcome of the study as authors make a claim denying conflicts of interest.
Other bias	**Low**	Appears free of other sources of bias.

48. Narrative justification for (Soga et al., 2020).

Bias	Rating	Support for judgment
Recruitment	**Low**	An online questionnaire survey delivered through a market research company (Rakuten Insight, Japan) was completed over a 3-d period in early June 2020 by 3,000 adults who are enrolled in their survey database. Participants were recruited from the same population at the same time frame.
Blinding	**Low**	There was no mention of blinding, however, this does not seem to influence the outcome and exposure measurement.
Exposure assessment	**Probably high**	They assessed multiple green space elements (1). Exposure metrics did not differentiate between types of greenness (0). The exposure was not assessed more than once (0). Different buffer distances or areas were not considered (0). The quality of greenness was not measured (0). They measured exposure time for greenspace use (1).
Confounding	**Probably low**	The study accounted for only one potential confounder, individual-level socioeconomic status. Model was adjusted for age, gender, annual household income, housing type, and degree of income changes during COVID-19.
Incomplete outcome data	**Probably low**	No evidence of missing outcome data.
Selective reporting	**Low**	All outcomes were reported.
Conflict of interest	**Low**	They did not mention anything about funding, which may indicate the free of support.
Other bias	**Low**	Appears free of other sources of bias.

49. Narrative justification for (Spano et al., 2021).

Bias	Rating	Support for judgment
Recruitment	**Low**	Online data collection was conducted on a large sample of community dwellers who spent the lockdown period in Italy during the COVID-19 pandemic through the snowball sampling method. The survey was conducted from March 31, 2020, to April 7, 2020. Participants were recruited from the same population at the same time frame, and the inclusion criteria were well described.

Blinding	**Low**	There was no mention of blinding, however, this does not seem to influence the outcome and exposure measurement.
Exposure assessment	**Probably low**	They assessed multiple green space elements (1) Exposure metrics differentiated between types of greenness (1). The exposure was not assessed more than once (0). Different buffer distances or areas were not considered (0). The quality of greenness or accessibility was measured (1). Time spent on green space was not considered (0)
Confounding	**Probably low**	The study accounted for only one potential confounder, individual level socioeconomic statuses such as age, gender, level of education, current place of residence, current working mode, presence of other people living in the home, presence of pets, and estimated decrease in income due to the COVID-19 pandemic lockdown.
Incomplete outcome data	**Low**	No missing data.
Selective reporting	**Low**	All outcomes were reported.
Conflict of interest	**Low**	The study did not receive support from a company, study author, or other entity having a financial interest in the outcome of the study as authors make a claim denying conflicts of interest.
Other bias	**Low**	Appears free of other sources of bias.

50. Narrative justification for (Sundara keeran et al. 2021)

Bias	**Rating**	**Support for judgment**
Recruitment	**Low**	An online study was carried out in Malaysia focused on 30 adults (aged between 18 and 40 years) who believed they were suffering from mental health issues as a direct result of the COVID-19 pandemic or the MCO. The survey was conducted on 19th April 2020 to 23rd April 2020, and all the study samples were recruited from same population. After experiencing nature therapy, further evaluation was carried out on 4th May.
Blinding	**Low**	There was no mention of blinding, however, any kind of personal identifiers such as contact details were only used for direct communication pertaining to this study and was deleted once the study was concluded.
Exposure assessment	**Probably High**	They estimated multiple green space elements (1). This exposure metric differentiated between types of greenness (1). The exposure was assessed one time (0). There were no different buffer distances or areas considered for the exposure measurement (0). They did not assess quality of greenness and accessibility (0). They did not assess the time spent in green spaces (0).

Confounding	**High**	The study did not account for any potential confounders.
Incomplete outcome data	**Low**	No missing outcome data.
Selective reporting	**Low**	All outcomes were reported.
Conflict of interest	**Low**	The study did not receive any financial support and authors declared no conflict of interest.
Other bias	**Low**	Appears free of other sources of bias.

51. Narrative justification for (Theodorou et al., 2021).

Bias	**Rating**	**Support for judgment**
Recruitment	**Low**	The data were collected in between March and May 2020 from Italian territory. The convenient sampling method was used to collect the sample from the participants with 18 or more than 18 years old people. All participants were recruited from the same population at the same time frame.
Blinding	**Low**	There was no mention of blinding, however, this does not seem to influence the outcome and exposure measurement.
Exposure assessment	**High**	Single exposure was considered (0). Exposure metrics did not differentiate between types of greenness (0). The exposure was not assessed more than once (0). Different buffer distances or areas were not considered (0). The quality of greenness or accessibility was not measured (0). They did not measure the use of green space or exposure time (0).
Confounding	**Probably low**	The study accounted for only one potential confounder, individual-level socioeconomic statuses such as gender, age, educational qualification, marital status, smart working, square meters per person, and number of infections per region. In a separate mediation model, Covid-19 related distress was used as a potential mediator.
Incomplete outcome data	**Low**	No missing data.
Selective reporting	**Low**	All outcomes were reported.
Conflict of interest	**Low**	The study received support from REI Cooperative for sponsorship of the Department of Environmental Health's research but the author declares no conflict of interest.

Other bias	**Low**	Appears free of other sources of bias.

52. Narrative justification for (Tomasso et al., 2021).

Bias	Rating	Support for judgment
Recruitment	**Low**	In this study, the individuals originally recruited for focus group interviews to explore formative experiences and origins of attitudes shaping nature-seeking behaviors as adults, which was collected during October 2019. A new recruitment email was sent to all enrolled participants to collect data on nature exposure under COVID-19 during April 2020. Inclusion and exclusión criteria were explained.
Blinding	**Low**	There was no mention of blinding, however, this does not seem to influence the outcome and exposure measurement.
Exposure assessment	**Low**	They assessed multiple green space elements (1). Exposure metrics were differentiated between types of greenness (1). The exposure was not assessed more than once (0). Different buffer zone was used for exposure assessment (1). The quality of greenness was measured (1). They measured the exposure time for greenspace use (1).
Confounding	**Low**	Individual-level and neighborhood-level confounders were assessed in this study, such as age, gender, urbanicity, race, and zip code in which the survey was taken, area-level poverty, and urban population density.
Incomplete outcome data	**Low**	No outcome missing data.
Selective reporting	**Low**	All outcomes were reported.
Conflict of interest	**Low**	The author declares no conflict of interest.
Other bias	**Low**	Appears free of other sources of bias.

53. Narrative justification for (Vos et al., 2021).

Bias	Rating	Support for judgment
Recruitment	**Low**	From December 23rd, 2020 until May 1st, 2021, 1680 participants from the ENVIRONAGE birth cohort (located in Belgium) were con- tacted and were asked to fill out an online survey. Participants were recruited from the same population at the same time frame. Inclusion and exclusión criteria were explained clearly.

Blinding	**Low**	There was no mention of blinding, however, this does not seem to influence the outcome and exposure measurement.
Exposure assessment	**Probably High**	One exposure matrice was used (0). Exposure metrics did not differentiate between types of greenness (0). The exposure was not assessed more than once (0). Multiple buffer zone was considered (1). There assessed the quality of greenness (1). They did not measure the time spent in greenspace (0).
Confounding	**Probably low**	The study accounted for only one potential confounder, individual-level socioeconomic status. All models were adjusted for change in monthly household income, the participant's age, the highest attained degree, and stress related to care for children.
Incomplete outcome data	**Low**	No evidence of missing outcome data.
Selective reporting	**Low**	All outcomes were reported.
Conflict of interest	**Low**	The study did not receive support from a company, study author, or other entity having a financial interest in the outcome of the study.
Other bias	**Low**	Appears free of other sources of bias.

54. Narrative justification for (Wortzel et al., 2021).

Bias	**Rating**	**Support for judgment**
Recruitment	**Low**	Participants for this study were sampled from an online survey (www.covid19resilience.org) launched in April 2020. Participants livingin the U.S. (N=2,089, 83.1% females, mean age 42 years, age range 18–90 years) were mapped to unique ZIP Codes. Participants were recruited from the same population at the same time frame. Inclusion and exclusión criteria were explained clearly.
Blinding	**Low**	There was no mention of blinding, however, this does not seem to influence the outcome and exposure measurement.
Exposure assessment	**Probably high**	One exposure matrice was considered (tree canopy) (0). The exposure metrics did not differentiate between types of greenness (0) They did not assess the exposure more than once over time (0). Different buffer distances were measured (1). Quality of greenness was measured (1). They did not measure the use of green space or exposure time (0).
Confounding	**Probably low**	The study accounted for only one potential confounder, individual-level socioeconomic status. Model were adjusted for gender, age, race, education status, prior mental health diagnosis, recent job loss, living alone, and urbanicity of the participant's ZIP Code of residence.

Incomplete outcome data	**Low**	No evidence of missing outcome data.
Selective reporting	**Low**	All outcomes were reported.
Conflict of interest	**Low**	The study was supported by National Institute of mental health (NIMH) grants and the Lifespan Brain Institute of Children's Hospital of Philadelphia and PennMedicine, University of Pennsylvania. The funding organizations had no role in the design and conduct of the study.
Other bias	**Low**	Appears free of other sources of bias.

55. Narrative justification for (Xie et al., 2020).

Bias	**Rating**	**Support for judgment**
Recruitment	**Low**	They used an anonymous online questionnaire for five days (1–5 April 2020). Participants were recruited from the same population at the same time frame. Inclusion and exclusion criteria were clearly explained.
Blinding	**Low**	There was no mention of blinding, however, this does not seem to influence the outcome and exposure measurement.
Exposure assessment	**Probably high**	One exposure matrice was considered (0). The exposure metrics did not differentiate between types of greenness (0). The exposure was not assessed more than once over time (0). Different buffer distances or áreas were not measured (0). They assessed quality, greenness or accessibility (1). Time spent in greenspace was measured (1).
Confounding	**High**	The study did not account for any potential confounders. However, six mediation model analyses were used separately.
Incomplete outcome data	**Low**	No evidence of outcome missing data.
Selective reporting	**Low**	All outcomes were reported.
Conflict of interest	**Low**	The study did not receive support from a company, study author, or other entity having a financial interest in the outcome of the study
Other bias	**Low**	Appears free of other sources of bias.

56. Narrative justification for (Yao et al., 2021).

Bias	**Rating**	**Support for judgment**

Recruitment	**Probably low**	There is insufficient information about participant selection to permit a judgment of 'Yes', but there is indirect evidence that suggests that participant recruitment or inclusion/exclusion criteria were consistent, as described by the criteria for a judgment of 'Yes'.
Blinding	**Low**	There was no mention of blinding, however, this does not seem to influence the outcome and exposure measurement.
Exposure assessment	**High**	One exposure matrice was considered (0). The exposure metrics did not differentiate between types of greenness (0). The exposure was not assessed more than once over time (0). Different buffer distances or areas were not considered (0). Specifically, the amount of parkland area was measured by the item of "area of parks and green spaces" in the yearbook. (1). They did not measure the use of green space or exposure time (0).
Confounding	**Probably high**	They adjusted one potential confounding factor. Models were adjusted for population, gross domestic product (GDP) per capita and unemployment rate, number of hospital beds, and number of doctors. However, models were also adjusted for environmental factors: average annual temperature (°C), average annual precipitation (mm), and environment stress index, which overadjusted the model.
Incomplete outcome data	**Low**	No evidence of missing outcome data.
Selective reporting	**Low**	All outcomes were reported.
Conflict of interest	**Low**	Funding source is limited to government, non-profit organizations, or academic grants funded by the government, foundations and/or non-profit organizations.
Other bias	**Low**	Appears free of other sources of bias.

57. Narrative justification for (Young et al., 2021).

Bias	**Rating**	**Support for judgment**
Recruitment	**Low**	In April 2020, a series of COVID-19 surveys to 269,363 KPRB participants were launched who were enrolled in a KP health plan as of that date. Approximately every two weeks after completing the baseline survey, a follow-up survey was sent to the 129,385 participants who completed the baseline survey. Participants without common COVID-19 symptoms (fever, cough, shortness of breath, or loss of sense, smell, or taste during the previous 2 weeks) were included. Inclusion and excusión criteria were explained.
Blinding	**Low**	There was no mention of blinding, however, this does not seem to influence the outcome and exposure measurement.
Exposure assessment	**Probably high**	Single exposure matric was considered (0). The exposure metrics did not differentiate between types of greenness (0). Exposure was assessed more than once (1). Different buffer distances or areas were not considered (0). There is no evidence of assessing the quality of greenness (0). They measured exposure time for greenspace use (1).

Confounding	**Low**	The study accounted for both individual and neighborhood-level potential confounders such as demographic data (age, sex, self-reported race/ethnicity), participants' EMR Neighborhood education etc.
Incomplete outcome data	**Low**	No evidence of missing outcome data.
Selective reporting	**Low**	All outcomes were reported.
Conflict of interest	**Low**	The Kaiser Permanente Research Bank provided funding for this work. But the author declares to have no possible, perceived or real financial conflicts of interest.
Other bias	**Low**	Appears free of other sources of bias.

58. Narrative justification for (Zhang et al., 2021).

Bias	**Rating**	**Support for judgment**
Recruitment	**Low**	This study was a cross-sectional study. Five residential areas were randomly selected in each of the three levels, for a total of 15 residential areas selected as research plots. Next, residents were convenience sampled from these 15 residential areas in two waves (May 2019 and May 2020), with different residents sampled in each wave. For both waves, we ensured that the number of valid residents for each residential area reached 30, for a total number of 450 valid residents per wave.
Blinding	**Low**	There was no mention of blinding, however, this does not seem to influence the outcome and exposure measurement.
Exposure assessment	**Low**	They assessed multiple greenspace elements (1). The exposure metrics differentiated between types of greenness (1). The exposure was assessed more than once over time (1). Different buffer distances or areas were considered (1). Quality of greenness or accessibility was assessed (1). They did not measure the use of green space or exposure time (0).
Confounding	**Probably low**	The study accounted for one potential confounder, such as house price, construction year, and number of households.
Incomplete outcome data	**Low**	No evidence of missing outcome data.
Selective reporting	**Low**	All outcomes were reported.
Conflict of interest	**Low**	This work was supported by the National Natural Science Foundation of China, but there is evidence that author does not have a financial interest in

		the outcome of the study.
Other bias	**Low**	Appears free of other sources of bias.

59. Narrative justification for (Zhuo and Zacharias, 2021).

Bias	Rating	Support for judgment
Recruitment	**Low**	An online survey was conducted in China from 20th to 24th February 2020 and collected data from 538 samples. Participants were recruited from the same population at the same time frame. Inclusion and exclusión criteria were explained clearly.
Blinding	**Low**	There was no mention of blinding, however, this does not seem to influence the outcome and exposure measurement.
Exposure assessment	**Probably high**	They assessed multiple greenspace elements (1). The exposure metrics differentiated between types of greenness (1). The exposure did not assess more than once (0). Different buffer distances or areas were not considered (0). The quality of greenness was not measured (0). They did not measure the use of green space or exposure time (0).
Confounding	**Probably Low**	The study accounted for only potential confounding factors, individual-level socioeconomic status such as sex, no income loss, student or not, and education.
Incomplete outcome data	**Low**	The number of analyzed samples is half of the original sample, but they explained the reason for exclusión clearly.
Selective reporting	**Low**	All outcomes were reported.
Conflict of interest	**Low**	The study did not receive support from a company, study author, or other entity having a financial interest in the outcome of the study
Other bias	**Low**	Appears free of other sources of bias.

TABLE S4. QUALITY OF EVIDENCE RATINGS BY OUTCOME

Outcome	Number of studies	Risk of Bias	Indirectness	Inconsistency	Imprecision	Publication Bias	Large Magnitude	Dose-Response	Confounding minimizes effect	Quality of evidence rating
Anxiety	15	Serious limitations[2]	No serious limitations[3]	Serious limitations[4]	No serious limitations[6]	No serious limitations[8]	No upgrade[9]	No upgrade[10]	No upgrade[10]	**VERY LOW**
Depression	15	Serious limitations[2]	No serious limitations[3]	Serious limitations[4]	No serious limitations[6]	No serious limitations[8]	No upgrade[9]	No upgrade[10]	No upgrade[10]	**VERY LOW**
Stress	13	Serious limitations[2]	No serious limitations[3]	Serious limitations[4]	No serious limitations[6]	No serious limitations[8]	No upgrade[9]	No upgrade[10]	No upgrade[10]	**VERY LOW**
General mental health problem	13	Serious limitations[2]	No serious limitations[3]	Serious limitations[4]	No serious limitations[6]	No serious limitations[8]	No upgrade[9]	No upgrade[10]	No upgrade[10]	**VERY LOW**
Sleep disturbances	6	Serious limitations[2]	No serious limitations[3]	Serious limitations[4]	No serious limitations[6]	No serious limitations[8]	No upgrade[9]	No upgrade[10]	No upgrade[10]	**VERY LOW**
Mood	4	Serious limitations[2]	No serious limitations[3]	Serious limitations[5]	No serious limitations[6]	No serious limitations[8]	No upgrade[9]	No upgrade[10]	No upgrade[10]	**VERY LOW**
Positive and negative affect	3	Serious limitations[2]	No serious limitations[3]	No serious limitations[5]	Serious limitations[7]	No serious limitations[8]	No upgrade[9]	No upgrade[10]	No upgrade[10]	**VERY LOW**
Loneliness	2	Serious limitations[2]	No serious limitations[3]	No serious limitations[5]	Serious lmitations[7]	No serious limitations[8]	No upgrade[9]	No upgrade[10]	No upgrade[10]	**VERY LOW**
Emotional distress	2	No serious limitations[1]	No serious limitations[3]	Serious limitations[4]	Serious limitations[7]	No serious limitations[8]	No upgrade[9]	No upgrade[10]	No upgrade[10]	**VERY LOW**
Mental well-being	17	Serious limitations[2]	No serious limitations[3]	Serious limitations[4]	No serious limitations[6]	No serious limitations[8]	No upgrade[9]	No upgrade[10]	No upgrade[10]	**VERY LOW**
Happiness	4	Serious limitations[2]	No serious limitations[3]	Serious limitations[4]	Serious limitations[7]	No serious limitations[8]	No upgrade[9]	No upgrade[10]	No upgrade[10]	**VERY LOW**
Life satisfaction	3	Serious limitations[2]	No serious limitations[3]	Serious limitations[4]	No serious limitations[6]	No serious limitations[8]	No upgrade[9]	No upgrade[10]	No upgrade[10]	**VERY LOW**

Notes:

1. Mostly “low” and “probably low” risk of bias studies with large sample size included.
2. Most studies rated “high” or “probably high” risk of bias.
3. All studies directly measure the exposure and outcome of interest.
4. Substantial heterogeneity that could not be explained by subgrouping, risk of bias, or influence analysis.
5. Low heterogeneity exist, or confidence interval were overlapped.
6. No support for other factors that may lead to a downgrade of quality of evidence.
7. Few studies available to make decisions or imprecision or decision-based on each side of the confidence intervals were associated with different judgments.
8. Inadequate studies to evaluate publication bias. However, given the comprehensive search and size of the sample in the published studies we decided little or even no publication bias.
9. In all reported studies, the magnitude of the effect sizes was not large enough to lead to an upgrade of evidence. No support for other factors that may lead to upgrade of quality of evidence.

TABLE S5. RISK OF BIAS, ASSOCIATION DIRECTION AND QUALITY OF EVIDENCE

	Study	Sample size	Study design	Nature exposure assessment	Buffer	Mental health assessment	Direction	Overall risk of bias
Anxiety (Very Low)								
1	(Asim et al., 2021)	432	Cross-sectional	Plants at home	NA	Anxiety (GAD-7)	↑	High
				Green view from window	NA		↑	
2	(Basu et al., 2021)	408	Cross-sectional	Presence of home garden	NA	Depression, Anxiety, Stress (DASS-21)	↑	Probably high
				Time spent in home garden	NA		↑	
3	(Bourion-Bédès et al., 2021a)	3936	Cross-sectional	Access to private outdoor space	NA	Anxiety (GAD-7)	↑	Low
4	(Bu et al., 2021)	55204	Cross-sectional	Time spent in gardening	NA	Anxiety (GAD-7)	↑	High
5	(Corley et al., 2021)	171	Cross-sectional	Gardening	NA	COVID-19 Anxiety	x	Low
				Garden usage frequency	NA		x	
6	(Dzhambov et al., 2021)	328	Cross-sectional	Plants at home	NA	Anxiety (GAD-7)	x	Low
				Green view from window	NA		↑	
				Presence of a domestic garden	NA		x	
				Perceived neighborhood greenery	NA		↑	
7	(Heo et al., 2021)	322	Cross-sectional	Decreased visit to greenspace	NA	Anxiety (GAD-2)	x	Low
				EVI	25th p		x	
					50th p			
					75th p			
8	(LÃµhmus et al., 2021)	2060	Cross-sectional	NDVI	50-m, 100-m, 300-m, 500-m	Anxiety (SCL90)	x	Probably high
				Frequency of nature visits	NA		x	
9	(Millán-Jiménez et al., 2021)	188	Cross-sectional	Green view from the window	NA	Anxiety	NA	High
				Presence of a livable outdoor space	NA		NA	
10	(Sarai Pouso et al., 2021)	5218	Cross-sectional	Green view from window	NA	Anxiety (GAD-2)	↑	Low
				Presence of a livable outdoor space	NA		↑	
11	(Sansal et al., 2021)	79	Cross-sectional	Time spent in outdoors space	NA	Depression, Anxiety, Stress (DASS-42)	↓	High
12	(Spano et al., 2021)	3886	Cross-sectional	Plants at home	NA	Anxiety	x	Low
				Green view from window	NA		↑	
				Accessibility to private green space	NA		↑	
13	(Sundara Keeren et al., 2021)	30	Cross-sectional	Nature exercise	NA	Depression, Anxiety, Stress (DASS-42)	↑	High
				Nature therapy	NA		↑	
14	(Wortzel et al., 2021)	2089	Cohort	Greenspace	0-m, 100-m, 250-m, 500-m, 100-m	Anxiety (GAD-7)	x	Probably high
15	(Young et al., 2021)	20012	Longitudinal	Time spent in outdoor nature	NA	Anxiety (GAD-2)	↑	Probably high
Depression (Very Low)								
1	(Amerio et al., 2020)	8177	Cross-sectional	Presence of a livable outdoor space	NA	Depression (PHQ-9)	↑	High
				Green view from window	NA		↑	
				Poor quality of views	NA		↓	
2	(Asim et al., 2021)	432	Cross-sectional	Plants at home	NA	Depression (CES-D)	↑	High
				Green view from window	NA		↑	
3	(Basu et al., 2021)	408	Cross-sectional	Absence of home garden	NA	Depression, Anxiety, Stress (DASS-21)	↑	Probably high
				Time spent in home garden	NA		↑	
4	(Bu et al., 2021)	55204	Cross-sectional	Time spent in gardening	NA	Depression (PHQ-9)	↑	High
5	(Dzhambov et al., 2021)	328	Cross-sectional	Plants at home	NA	Depression (PHQ-9)	↑	Low

				Green view from window	NA		↑	
				Presence of a domestic garden	NA		x	
				Perceived neighborhood greenery	NA		↑	
6	(Heo et al., 2021)	322	Cross-sectional	Decreased visit to greenspace	NA	Depression (PHQ-9)	↑	Low
				EVI	25th p		x	
					50th p			
					75th p			
7	(LÃµhmus et al., 2021)	2060	Cross-sectional	NDVI	50-m, 100-m, 300-m, 500-m	Depression (SCL90)	x	Probably high
				Frequency of nature visit	NA		↑	
8	(Millán-Jiménez et al., 2021)	188	Cross-sectional	View from the window	NA	Depression symptoms	NA	High
				Presence of livable outdoor space	NA		NA	
9	(A Olszewska-Guizzo et al., 2021)	25	Longitudinal	Frequency of visits to nature areas	NA	Depression (BDI- II)	x	High
				Time spent in nature areas	NA		x	
10	(Sarai Pouso et al., 2021)	5218	Cross-sectional	Green view from window;	NA	Depression (PHQ-2)	↑	Low
				Presence of available outdoor space	NA		↑	
11	(Samus et al., 2022)	261	Cross-sectional	Time spent in nature spaces	NA	Depression (CES-D)	x	Probably high
				Garden use	NA		x	
				Nature connection	NA		x	
12	(Sansal et al., 2021)	79	Cross-sectional	Time spent in outdoor space	NA	Depression, Anxiety, Stress (DASS-42)	↑	High
13	(Sundara Keeren et al., 2021)	30	Cross-sectional	Nature exercise	NA	Depression, Anxiety, Stress (DASS-21)	↑	High
				Nature therapy	NA		↑	
14	(Wortzel et al., 2021)	2089	Cohort	Greenspace	0-m, 100-m, 250-m, 500-m, 100-m	Depression (PHQ-2)	↑	Probably high
15	(Young et al., 2021)	20012	Longitudinal	Less time spent in outdoor nature	NA	Depression (PHQ-2)	↑	Probably high
Stress (Very Low)								
1	(Basu et al., 2021)	408	Cross-sectional	Presence of home garden	NA	Depression, Anxiety, Stress (DASS-21)	↑	Probably high
				Time spent in home garden	NA		↑	
2	(Bourion-Bédès et al., 2021b)	3764	Cross-sectional	Access to private outdoor space	NA	Perceived Stress Scale (PSS-10)	↑	Low
3	(Jato-Espino et al., 2022)	9883	Cross-sectional	Green area	150-m,	Increase in stress	↓	Probably high
					300-m,		↓	
					500-m		x	
				Urban park	150-m,		NA	
					300-m,		NA	
					500-m		NA	
				Natural area	300-m		NA	
				Vegetative area	500-m		NA	
4	(LÃµhmus et al., 2021)	2060	Cross-sectional	NDVI	50-m, 100-m, 300-m, 500-m	Stress (PSS)	↑	Probably high
				Frequency of nature visits	NA		↑	Low
5	(Mintz et al., 2021)	776	Cross-sectional	Nature near home	NA	Stress (STAI)	↑	
				Green view from window	NA		↑	
				Nature connection	NA		x	
6	(Pearson et al., 2021)	56	Cross-sectional	Changed usage of greenspace	NA	Stress (PSS-10)	↑	Low
7	(Ribeiro et al., 2021)	1638	Cross-sectional	Indoor plants	NA	Stress (1-item)	↑	Low
				Balcony plants	NA		x	
				Presence of home garden	NA		x	
				Green view from window	NA		↑	

				Community private greenspace	NA		x	
				Public natural spaces	NA		↑	
8	(Robinson et al., 2021)	1184	Ecological	NDVI	50-m	Stress	x	Probably high
					100-m		x	
					250-m		↑	
					500-m		x	
				Greenspace presence	50-m		x	
					100-m		x	
					250-m		x	
					500-m		x	
				Greenspace abundance	50-m		x	
					100-m		x	
					250-m		x	
					500-m		x	
				Distance to green space	NA		x	
9	(Sansal et al., 2021)	79	Cross-sectional	Time spent in outdoors space	NA	Depression, Anxiety, Stress (DASS-42)	↑	High
10	(Sundara Keeren et al., 2021)	30	Cross-sectional	Nature exercise	NA	Depression, Anxiety, Stress (DASS-21)	↑	High
				Nature therapy	NA		↑	
11	(Vos et al., 2022)	766	Cohort	Greenspace area	50-m, 100-m, 300-m, 500-m, 1000-m	Stress	↑	Probably high
12	(Wortzel et al., 2021)	2089	Cohort	Greenspace	0-m, 100-m, 250-m, 500-m, 100-m	COVID-19 stress	x	Probably high
13	(Yao et al., 2022)	Urban residents of 296 cities in China	Ecological	Area of parks and green spaces (Parkland)	NA	Mental stress index	↑	High
General mental health problem (Very Low)								
1	(Browning et al., 2021)	2534	Cross-sectional	Time spent in green outdoors	NA	General mental health	↑	Probably high
2	(Chen & Liu, 2021)	937	Cross-sectional	Distance to Park	NA	General mental health (K6 scale)	NA	High
3	(Corley et al., 2021)	171	Cross-sectional	Gardening	NA	General mental health	x	Low
				Garden usage frequency	NA		↑	
4	(Hubbard et al., 2021)	2969	Cross-sectional	Access to garden/balcony	NA	General mental health (PHQ-4)	↑	High
				Frequency of visit to greenspace	NA		↑	
				Time spent in greenspace	NA		x	
5	(Kou et al., 2021)	1154	Participatory action research (PAR)	Community gardening	NA	General Mental health	NA	High
6	(Lenaerts et al., 2021)	11352	Cross-sectional	Frequency of nature visit	NA	General Mental health (6 items)	↑	Low
7	(Marques et al., 2021)	173	Cross-sectional	Frequency of visit to parks and beach	NA	General mental health (DASS21)	↑	Low
				Green view from window	NA		↑	
				Plants at home	NA		x	
				Presence of a garden	NA		↑	
				Distance to greenspace	NA		x	
8	(Oswald et al., 2021)	1004	Cross-sectional	Presence of livable outdoor greenspace	NA	General mental health (MHC-SF, K10)	x	Low
				Walking distance of a greenspace	NA		x	
				Perceived neighborhood greenery	NA		↑	
				Time spent in nature	NA		↑	
9	(Ribeiro et al., 2021)	1638	Cross-sectional	Indoor plats	NA	General mental health (GHQ12)	x	Low
				Balcony plants	NA		x	
				Presence of home garden	NA		x	
				Green view from window	NA		↑	
				Community private greenspace	NA		x	
				Public natural spaces	NA		x	
10	(Soga et al., 2020)	3000	Cross-sectional	NDVI	250-m	General mental health (GHQ-12)	x	Probably high

				Green views from window	NA		↑	
				Frequency to visit greenspace	NA		↑	
11	(Theodorou et al., 2021)	303	Cross-sectional	Gardening activities	NA	General mental health (SCL-K-9)	↑	High
12	(Xie et al., 2020)	386	Cross-sectional	Time spent in greenspace	NA	General mental health (1 item)	↑	High
				Frequency of visits to greenspace	NA		x	
13	(Zhang et al., 2022)	900	Cross-sectional	Tree canopy coverage (TCC)	NA	General mental health (GHQ12)	↑	Low
Sleep disturbances (Very Low)								
1	(Corley et al., 2021)	171	Cross-sectional	Gardening	NA	Sleep Quality (1-item PSQI)	x	Low
				Garden usage frequency	NA		x	
2	(Kontsevaya et al., 2021)	2432	Cross-sectional	Access to greenspace	NA	Sleep quality (3-item)	x	High
3	(Millán-Jiménez et al., 2021)	188	Cross-sectional	Green view from window	NA	Sleep problem (Insomnia),	NA	High
				Presence of a livable outdoor space	NA		NA	
4	(Okely et al., 2021)	948	Longitudinal	Presence of outdoor greenspace	NA	Sleep duration	x	High
5	(Sansal et al., 2021)	79	Cross-sectional	Time spent in outdoor space	NA	Sleep	↑	High
6	(Spano et al., 2021)	3886	Cross-sectional	Plants at home	NA	Sleep disturbance	↑	Low
				Green view from window	NA		↑	
				Accessibility to private green space	NA		↑	
Mood (Very Low)								
1	(Gola et al., 2021)	77	Cross-sectional	Green Space in Healthcare facility	NA	Profile of Mood States-34 Items)	↑	High
				Public & private green space	NA		↑	
2	(Jato-Espino et al., 2022)	9883	Cross-sectional	Green area	150-m,	Mood	NA	Probably high
					300-m,		NA	
					500-m		NA	
				Urban park	150-m,		NA	
					300-m,		NA	
					500-m		↓	
				Natural area	300-m		NA	
				Vegetated area	500-m		NA	
3	(Millán-Jiménez et al., 2021)	188	Cross-sectional	Green view from window	NA	Irritability or bad modos	NA	High
				Presence of outdoor space	NA		NA	
4	(A Olszewska-Guizzo et al., 2021)	25	Longitudinal	Frequency of visits to nature areas	NA	Profile of Mood Scale (POMS)	x	High
				Time spent in nature areas	NA		x	
Positive and Negative Affect (Very Low)								
1	(Lades et al., 2020)	604	Cross-sectional	Gardening	NA	Positive and Negative Affect (6-item)	↑	High
2	(Mintz et al., 2021)	776	Cross-sectional	Nature near home	NA	Positive and Negative Affect (PANAS)	↑	Low
				Green view from window	NA		↑	
				Nature connection	NA		x	
3	(Samus et al., 2022)	261	Cross-sectional	Time spent in nature spaces	NA	Positive and Negative Affect (PANAS)	↑	Probably high
				Garden use	NA		x	
				Nature connection	NA		↑	
Loneliness (Very low)								
1	(Groot et al., 2022)	7445	Cohort	Access to outdoor greenspace	NA	Loneliness (1-item)	↑	High
2	(Soga et al., 2020)	3000	Cross-sectional	NDVI	250-m	Loneliness (UCLA Loneliness);	x	Probably high
				Green views from windows from home	NA		↑	
				Frequency to visit greenspace	NA		↑	
Emotional distress (Very Low)								
1	(Friedman et al., 2022)	376	Cross-sectional	Nature connection	NA	Emotional problem	↑	High
2	(Larson et al., 2022)	1280	Cross-sectional	Area of local park/10000 residents	NA	Emotional problem (5-item)	x	Low
				Area of national park/10000 residents	NA		↑	

				NDVI	Residential Zip code level		x	
				Reducing park usage	NA		↑	
Mental well-being (Very Low)								
1	(Garrido-Cumbrera et al., 2021)	2464	Cross-sectional	Nature sound	NA	Mental Well-being (WHO-5)	↑	High
2	(Groot et al., 2022)	7445	Cohort	Access to outdoor greenspace	NA	Mental well-being (7-item)	↑	High
3	(Hansmann et al., 2021)	5932	Cross-sectional	Access to outdoor greenspace	NA	Mental well-being (4-item)	↑	High
4	(Huerta & Utomo, 2021)	1954	Cross-sectional	Quality of urban greenspace (UGS)	NA	Mental Well-being Scale (WEMWS)	x	Low
				Frequency of UGS use	NA		↑	
				Distance to the nearest green space	NA		x	
5	(Jackson et al., 2021a)	624	Cross-sectional	Frequency of nature-based activities	NA	Mental Well-being change (WHO-5)	↑	Probably high
				Nature experiences	NA		x	
6	(Jackson et al., 2021b)	624	Cross-sectional	Nature experience activities	NA	Subjective well-being (SWB) (4-item)	↑	Probably high
				Connectedness to nature (6-items)	NA		x	
7	(Lee et al., 2021)	1196	Cross-sectional	Frequency of visits to the forest site	NA	Mental Well-being (PWI-SF)	x	Low
				Time spent at the forest site	NA		↑	
				Walking time distance to the forest site	NA		x	
8	(Lehberger et al., 2021)	495	Cross-sectional	Presence of a garden	NA	Mental Well-being (WEMWBS)	x	Low
				Time spent in gardening weekly	NA		↑	
				Time spent in the public green	NA		↑	
9	(LÃµhmus et al., 2021)	2060	Cross-sectional	NDVI	50-m, 100-m, 300-m, 500-m	Mental Well-being (RAND-36)	↑	Probably high
				Frequency of nature visits	NA		↑	
10	(Mead et al., 2021)	138	Cross-sectional	Connectedness to nature	NA	Mental Well-being (WEMWBS)	x	High
11	(Pérez-Urrestarazu et al., 2021)	4205	Cross-sectional	Frequency of visits to greenspace	NA	Emotional Well-being	x	High
				Plants at home	NA		↑	
12	(Poortinga et al., 2021)	5566	Longitudinal	Distance to nearest greenspace	NA	Mental Well-being (SF36 items)	↑	Low
				Access to private garden	NA		↑	
13	(Robinson et al., 2021)	1184	Ecological	NDVI	50-m	Mental Well-being (WEMWBS)	x	Probably high
					100-m		x	
					250-m		↑	
					500-m		↑	
				Greenspace presence	50-m		x	
					100-m		x	
					250-m		x	
					500-m		x	
				Greenspace abundance	50-m		x	
					100-m		x	
					250-m		x	
					500-m		↑	
				Distance to green space	NA			
14	(Samuelsson et al., 2021)	684	Cross-sectional	Distance to greenspace	NA	Subjective well-being (1-item)	x	Low
				Frequency of visit to natural land cover	NA		↑	
15	(Sia et al., 2022)	5661	Cross-sectional	Time spent in gardening weekly	NA	Mental resilience (SYRSS-10)	↑	High
16	(Tomasso et al., 2021)	529	Cross-sectional	Connectedness to nature (INS scale)	NA	Mental Well-being (flourishing HFI)	x	Low
				Nature deprivation perceived	NA		x	
17	(Zhuo & Zacharias, 2021)	284	Cross-sectional	Presence of yard, balcony or terrace	NA	Mental well-being (6-item)	↓	Probably high
Happiness (Very Low)								
1	(Cheng et al., 2021)	1305	Ecological	NDVI	City scale	Happiness	↑	Probably high
2	(Millán-Jiménez et al.,	188	Cross-sectional	Green view from window	NA	Happiness	NA	High

	2021)							
				Presence of a livable outdoor space	NA		NA	
3	(Mintz et al., 2021)	776	Cross-sectional	Nature near home	NA	Happiness (1-item)	X	Low
				Green view from window	NA		X	
				Nature connection	NA		X	
4	(Soga et al., 2020)	3000	Cross-sectional	NDVI	250-m	Happiness (SHS)	X	Probably high
				Green views from window	NA		↑	
				Frequency to visit greenspace	NA		↑	
Life Satisfaction (Very Low)								
1	(Bu et al., 2021)	55204	Cross-sectional	Time spent in gardening	NA	Life satisfaction (1-item)	↑	High
2	(Lehberger et al., 2021)	495	Cross-sectional	Presence of a garden	NA	Life satisfaction (1 item)	x	Low
				Time spent in gardening weekly	NA		x	
				Time spent in the public greens	NA		↑	
3	(Soga et al., 2020)	3000	Cross-sectional	NDVI	250-m	Life Satisfaction (LSI-A)	x	Probably high
				Green view from window	NA		↑	
				Frequency to visit greenspace	NA		↑	

Notes: Measures in which an association was identified as positive (↑) indicating nature exposure improves mental health or negative (↓) indicating nature exposure worsens mental health or no significant association (x) or no statistical association reported (NA) are shown in the table.

TABLE S6. INPUT DATA FOR META-ANALYSES

MA model	Study	Effect size type	Source of extracted data	Transformation	Interpretation
Well-being, greater distance to nature (> 10 min/>250m)	Huerta	OR	Table 4	no	Longer walking distance to greenspace not associated with positive change in well-being
	Lee	OR	Table 5	Calculated from raw number of cases in Table 5, and reversed by dividing 1 by it to have the same interpretation as the other studies, because the outcome here is negative well-being	Longer walking distance (>10 min vs < 10 min) to a forest not associated with better well-being
	Robinson	OR	Table 4	no	Distance to greenspace < 250m not associated with better well-bieng
Well-being, Higher frequency of greenspace visits	Jackson	b standardized	Table 5	95% CI calculated from p-value	Continued participation in nature-based activities associated with increased well-being
	Soga	b standardized	Suppl Table S2	no	More frequent visits to greenspace associated with more happiness
Anxiety, Presence of Garden	Dzhambov	OR	Table 2	no	Having a garden vs not having doesn't change the odds of anxiety
	Pouso	OR	Table 2	no	Having access to a garden/outdoor space associated with lower odds of anxiety
Anxiety, Window view of nature	Dzhambov	OR	Table 2	Model rerun by the author using a dichotomized exposure	Having a greener view is associated with lower odds of anxiety
	Pouso	OR	Table 2	no	Having a nature view associated with lower odds of anxiety
Depression/loneliness, Presence of a Garden	Groot	OR	Figure 6	Reversed OR and 95% ci first to ln scale, then to negative lnOR, then calculated SE from 95% CI, then calculated new 95% CI on ln scale and then reverted back to exp scale	No access to ourdoor nature (garden) associated with higher loneliness
	Dzhambov	OR	Table 2	no	Having a garden vs not having doesn't change the odds of depression
	Pouso	OR	Table 2	no	Having access to a garden/outdoor space associated with lower odds of depression
Depression, Time spent in greenspace	Basu	b standardized	Table 3	95% CI calculated from p-value	Time spent in outdoor nature assocated with lower depression
	Young	b standardized	Table 3	no	Spending less time in outdoors nature associated with higher depression scores
Depression, Window view of nature	Amerio	OR	Table 4	no	Having a green view reduces the odds of moderate-severe depression
	Dzhambov	OR	Table 2	Model rerun by the author using a dichotomized exposure	Having a greener view is associated with lower odds of depression
	Pouso	OR	Table 2	no	Having a nature view associated with lower odds of depression
General mental health, Higher frequency of nature visits	Soga	b standardized	Suppl Table S2	Reversed sign of the estimate	More frequent visits to greenspace associated with better mental health
	Xie	b standardized	Figure 4, panel f	95% CI calculated from SE	More frequent visits to greenspace associated with better mental health
General mental health, Presence of a Garden	Corley	OR	Table 4	Calculated from lnOR	Gardening vs non-gardening not associated with general meantal health
	Oswald	OR	Table 5	no	Having a garden/balcony/yard vs not assocoiated with moderate mental health problems
Poor general mental health, Distance to nature (> 300/500m)	Oswald	OR	Table 5	no	Living >300m from a greenspace not associated with moderate mental health problems
	Zhang-GS	OR	Table 1	Calculated from raw number of cases	Living >500m from a greenspace not associated with poor mental health

FIGURE S1. META-ANALYSIS PLOTS

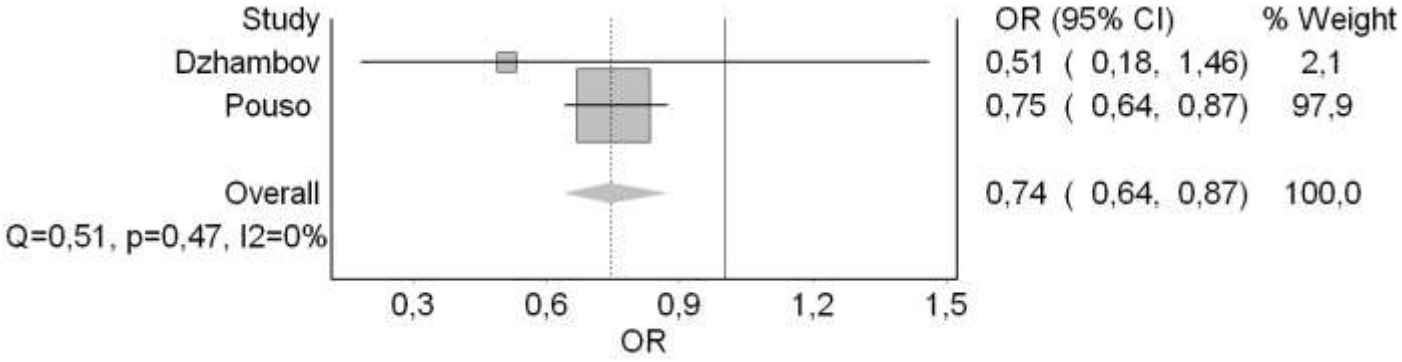

Figure S1A. Anxiety and presence of a garden

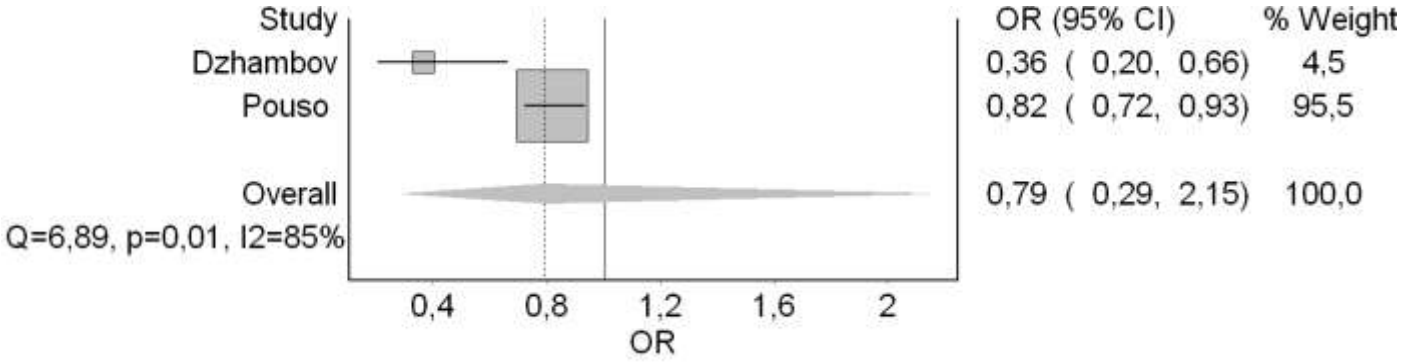

Figure S1B. Anxiety and window view of nature

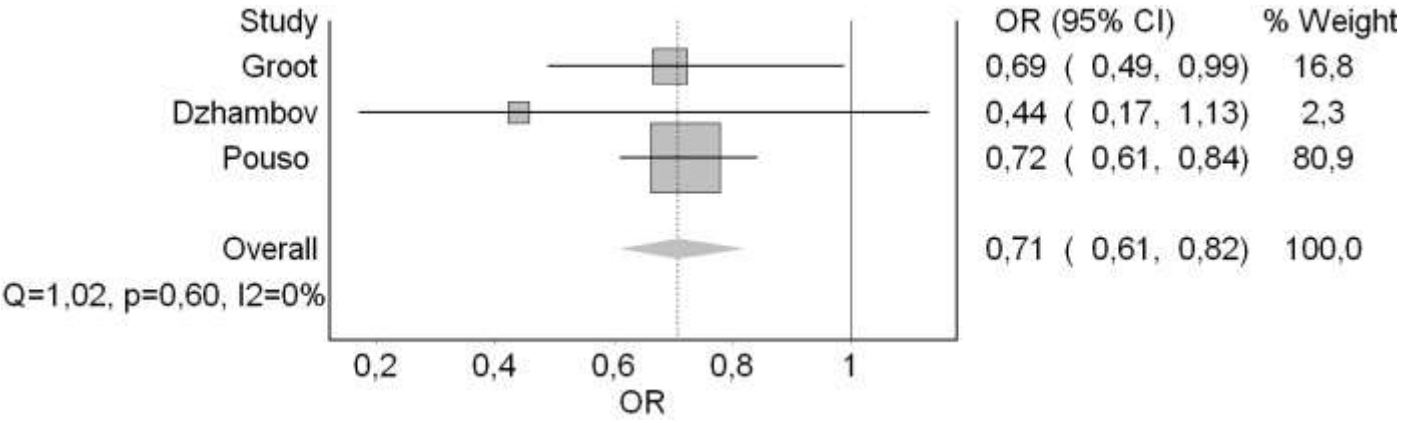

Figure S1C. Depression/loneliness and presence of a garden

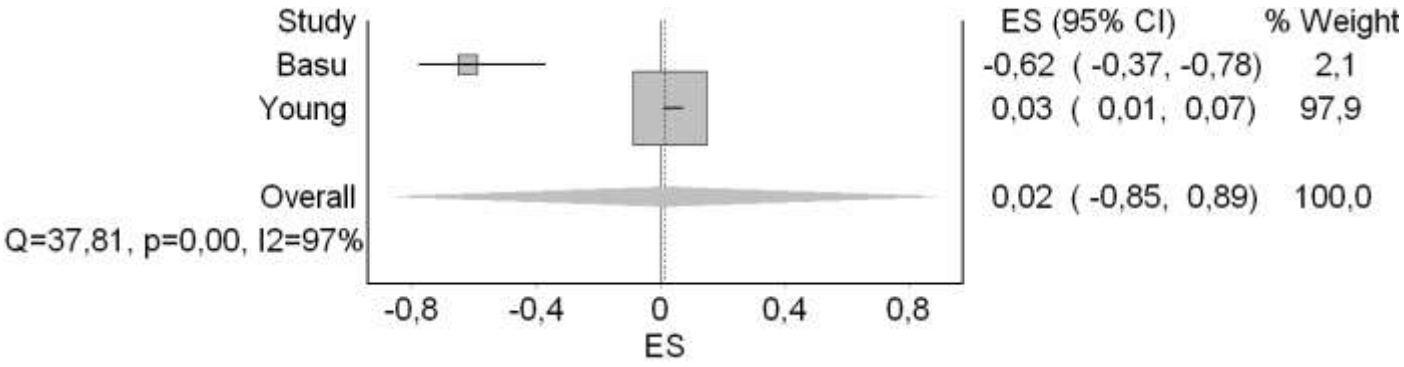

Figure S1D. Depression and time spent in greenspace

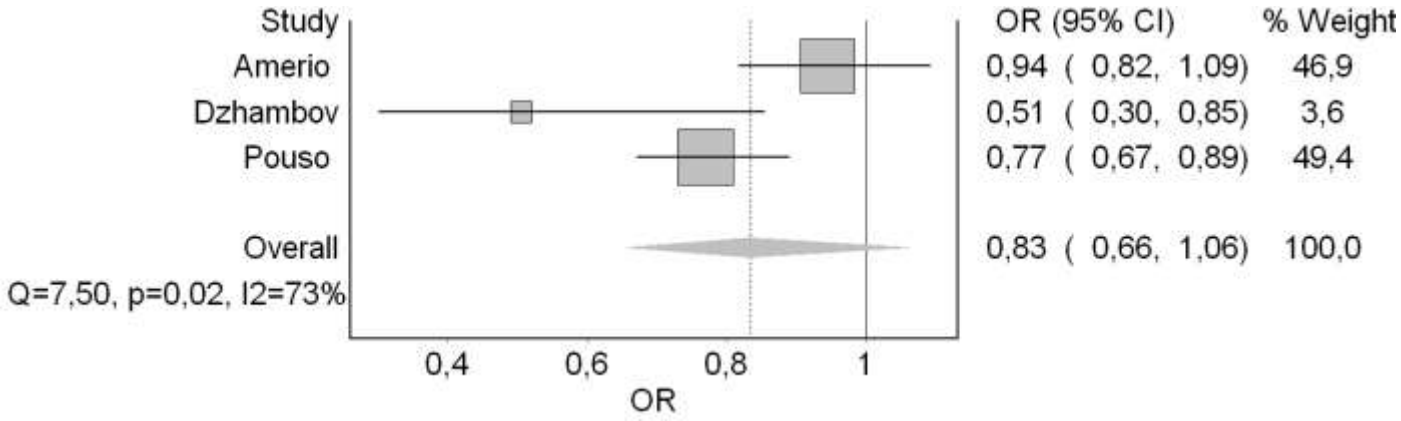

Figure S1E. Depression and window view of nature

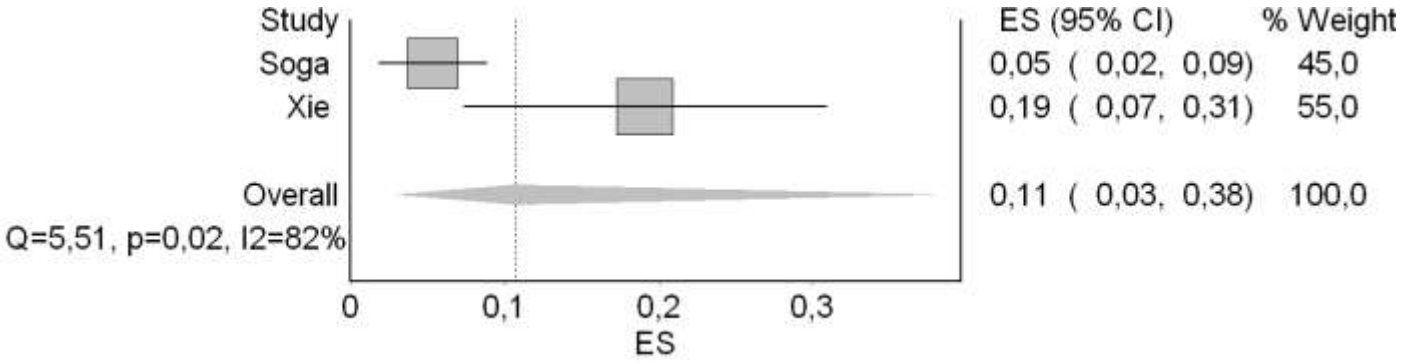

Figure S1F. General mental health and higher frequency of nature visits

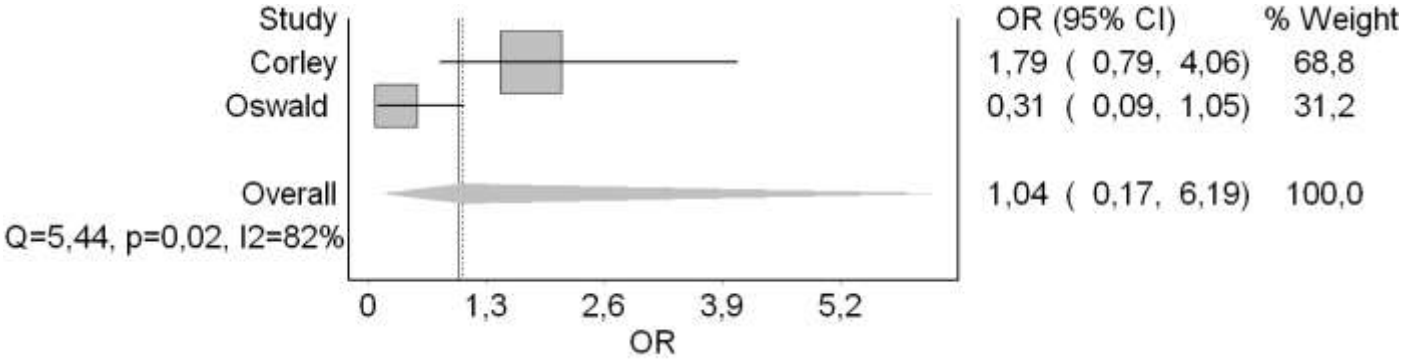

Figure S1G. General mental health and presence of a garden

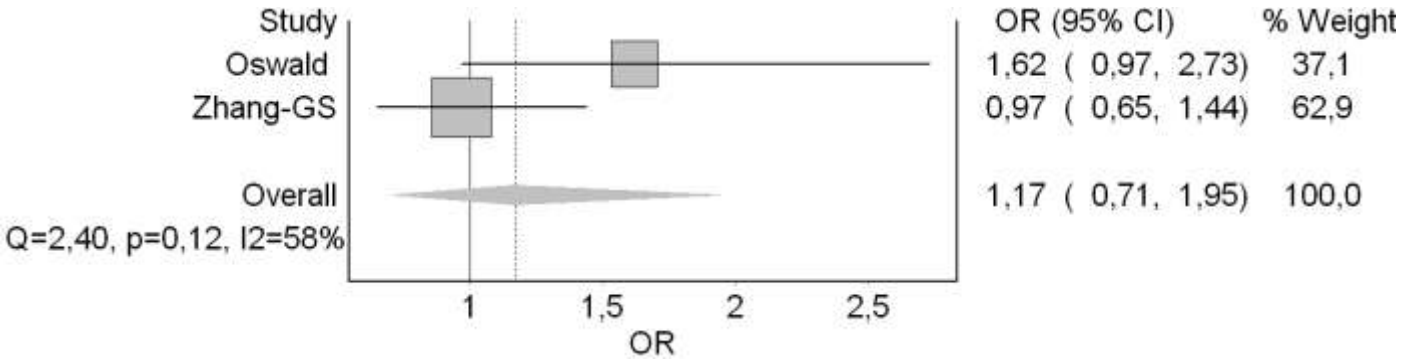

Figure S1H. Poor general mental health and distance to nature (>300/500m)

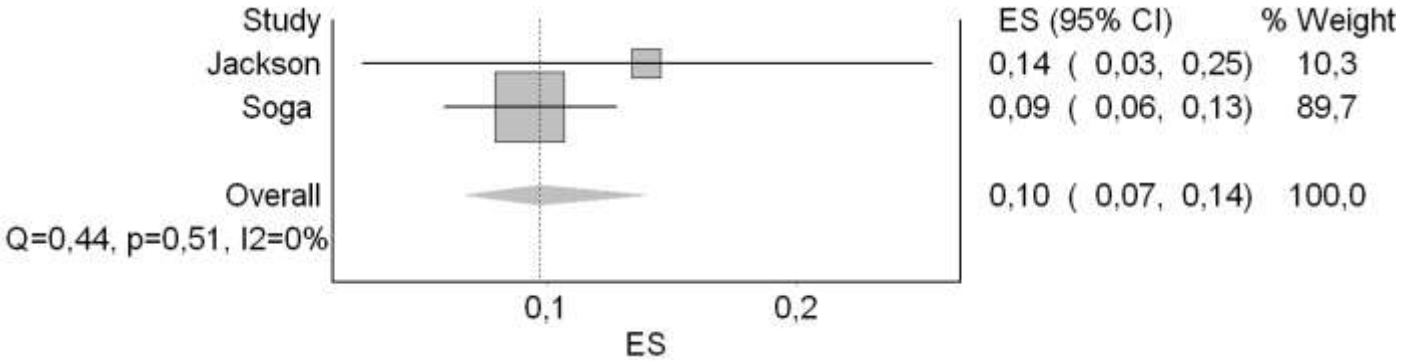

Figure S1I. Well-being and higher frequency of greenspace visits

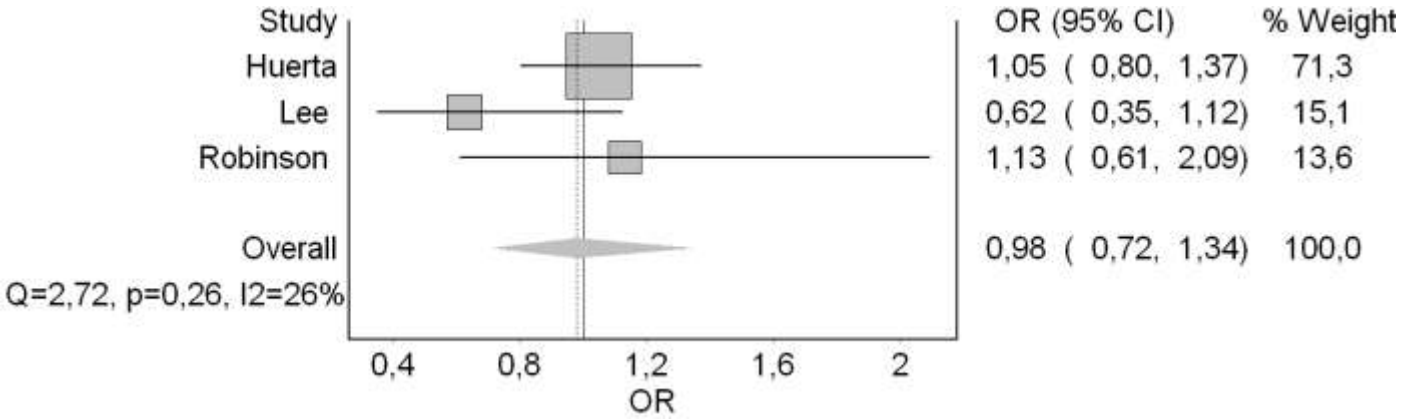

Figure S1J. Well-being and greater distance to nature