



HAL
open science

Exposure to greenspace and cancer incidence, prevalence, and mortality: A systematic review and meta-analyses

Mohammad Javad Zare Sakhvidi, Jun Yang, Amir Houshang Mehrparvar, Angel M Dzhambov, A Ebrahimi, Payam Dadvand, Bénédicte Jacquemin

► To cite this version:

Mohammad Javad Zare Sakhvidi, Jun Yang, Amir Houshang Mehrparvar, Angel M Dzhambov, A Ebrahimi, et al.. Exposure to greenspace and cancer incidence, prevalence, and mortality: A systematic review and meta-analyses. *Science of the Total Environment*, 2022, 838 (Pt 2), pp.156180. 10.1016/j.scitotenv.2022.156180 . hal-03719629

HAL Id: hal-03719629

<https://hal.science/hal-03719629v1>

Submitted on 21 Jul 2022

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

Exposure to greenspace and cancer incidence, prevalence, and mortality: A systematic review and meta-analyses

Mohammad Javad Zare Sakhvidi^{1,2}, Jun Yang³, Amir Houshang Mehrparvar⁴, Angel M. Dzhambov⁵, AliAsghar Ebrahimi⁶, Payam Dadvand⁷⁻⁹, Bénédicte Jacquemin^{2*}

1- Occupational health department, School of Public Health, Shahid Sadoughi University of Medical Sciences, Yazd, Iran

2- Univ Rennes, Inserm, EHESP, Irset (Institut de recherche en santé, environnement et travail) – UMR_S 1085, F-35000 Rennes, France

3- School of Public Health, Guangzhou Medical University, Guangzhou, 511436, China

4- Industrial diseases research center, Shahid Sadoughi University of Medical Sciences, Yazd, Iran

5- Department of Hygiene, Faculty of Public Health, Medical University of Plovdiv, Plovdiv, Bulgaria

6- Environmental health department, School of Public Health, Shahid Sadoughi University of Medical Sciences, Yazd, Iran

7-ISGlobal, Doctor Aiguader 88, 08003 Barcelona, Spain

8-Universitat Pompeu Fabra (UPF), Doctor Aiguader 88, 08003 Barcelona, Spain

9-CIBER Epidemiología y Salud Pública (CIBERESP), Melchor Fernández Almagro, 3-5, 28029 Madrid, Spain

Corresponding author: Bénédicte Jacquemin; benedicte.jacquemin@inserm.fr

- Findings on the association between greenspace and cancer are not conclusive.
- The overall quality of evidence of all of the exposure-outcome pairs are very low.
- Greenspace is potentially protective for breast cancer.
- For skin cancer the findings are not suggestive for protective role of greenspace.

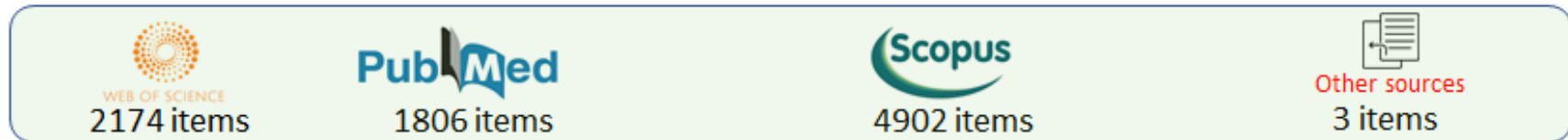
Abstract

We conducted a systematic review and meta-analysis of the available literature on the association between greenspace exposure and all-sites and site-specific cancer incidence, prevalence, and mortality in adults. We searched PubMed, Scopus, and Web of Science for original articles published, without language restriction until September 2021. We assessed the risk of bias in each study and the overall quality of evidence for exposure-outcome pairs that were reported in two or more studies. Out of the 18 included studies, cross-sectional studies were the most common study design (n=8), and most of the studies were conducted in Europe (n=8). In terms of risk of bias, the majority of cohorts (four out of six) and case-control studies (three out of four) were of good or very good quality, and cross-sectional studies were mostly (five out of eight) of poor quality. Outcomes (incidence, prevalence, mortality) on different cancer sites were reported: lung cancer (n=9), prostate cancer (n=4), breast cancer (n=4), skin cancer (n=3), colorectal cancer (n=2), all-sites cancer (n=2), brain cancer (n=1), mouth and throat cancer (n=1), and esophageal cancer (n=1). The meta-analyses for the breast, lung, and prostate cancer incidence did not show statistically significant associations (for example for breast cancer: hazard ratio= 0.83; 95% confidence interval: 0.47-1.48). For skin cancer, the available evidence suggests that greenspace could be a potential risk factor. For the other cancers, the evidence was non-conclusive. The overall quality of evidence of all of the exposure-outcome pairs was very low. Given the wide confidence interval of the pooled estimates and very low quality of evidence, the findings should be interpreted with caution. Future large and longitudinal studies are needed to assess the potential association of greenspace exposure with cancers, considering

types and quality of greenspace, evaluation of cancer sub-types, and adjustment for a sufficient set of covariates.

Keywords: Natural environment; Neoplasms; Non-communicable diseases; Urbanity; Nature

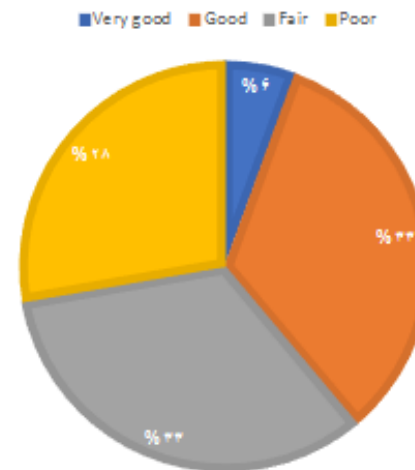
Graphical abstract



18 articles selected: 8 Cross sectional; 6 Longitudinal; 4 Case control

Lung cancer (n=9)
Prostate cancer (n=4)
Breast cancer (n=4)
Skin cancer (n=3)
Colorectal cancer (n=2)
All-sites cancer (n=2)
Brain cancer (n=1)
Mouth and throat cancer (n=1)
Esophageal cancer (n=1)

QUALITY OF STUDIES



Potentially protective for breast cancer.



Potentially risk factor for skin cancer.

Main finding:

Greenspace is potentially protective for breast cancer. The quality of studies needs improvement.

Introduction

Urbanization has diminished human contact with the natural environment and increased exposure to anthropogenic pollutants [1, 2]. Emerging evidence suggests a variety of health beneficial effects of greenspace (which *has been defined by the U.S. Environmental Protection Agency as “land that is partly or completely covered with grass, trees, shrubs, or other vegetation”* [3]). The evidence on the association between exposure to greenspace and cancer is still emerging compared with other health outcomes [4, 5]. The available literature on this association is also inconsistent, while some studies have found a beneficial association for greenspace [6-8], others have reported an increased risk of cancers associated with this exposure [9, 10], or have not observed any associations [11, 12]. Stress reduction [13], mitigation of urban-related pollutants (*e.g.* air pollution and noise) [14, 15], and improved social interaction and physical activity [16, 17] have been suggested to underlie the reported beneficial associations observed between greenspace and cancer. On the other hand, studies that have reported a detrimental association of greenspace exposure and cancer have justified their findings by increased exposure to biogenic volatile organic compounds and ozone, pesticides, or ultraviolet radiation [9, 10, 18]. Such apparent inconsistent results could be attributed to heterogeneity across studies in terms of cancer site and type, greenspace exposure definition, geographical location of the study, quality and design of the study, as well as other contextual factors such as deprivation, and urbanity.

Up to now, two scoping reviews are available on the association between greenspace/natural environment exposure and cancer outcomes [4, 5]. None of these reviews assessed the risk of bias in the literature, and in both reviews some of the current evidence with important contribution to the fields was not included, given the fact that the evidence on this association is

accumulating quickly. Moreover, they did not quantitatively synthesize the available evidence, for example, by conducting meta-analysis. Given rapid increase in the studies on the association between greenspace exposure and cancer, and no available systematic review on the evidence, we aimed to systematically assess the available evidence on the association between exposure to greenspace, and cancer incidence, prevalence, and mortality in adults, and where possible, conduct a meta-analysis of the reported association estimates.

Methods

We conducted this review according to the Preferred Reporting Items for Systematic reviews and Meta-Analyses (PRISMA) statement (2020) [19], and registered the protocol of the review in the PROSPERO (registration ID: CRD42021277074). The study has been approved by the institutional review board of the Shahid Sadoughi University of medical sciences.

Eligibility criteria

The review research was: “Is exposure to greenspace associated with cancer outcomes: incidence prevalence or mortality?”, We formulated the search strategy and PECO (population, exposure, comparison, and outcome) to answer the review research question. The population of interest in this review was human adults. We excluded animal studies (if any). The exposure of interest was defined as greenspace exposure in terms of greenspace availability (*e.g.*, surrounding greenspace, percent of land cover with greenspace), greenspace accessibility (*e.g.*, proximity to greenspaces, visual access to greenspace), or greenspace use (*e.g.*, time attending parks). Only studies with objective exposure assessment (*e.g.*, by quantitative description of greenspace exposure with indicators such as surrounding greenspace within a specific buffer size, or percentage of land

cover with a specific type of greenspace) were entered into the review. Subjective measures of greenspace exposure and self-report scales were excluded (*e.g.*, yes/no answer to the question about availability, access, or use of greenspace without reporting the details on buffer size, duration, or time-scale). The outcome was defined as all-site and site-specific cancer incidence, prevalence, and mortality. We excluded records on secondary tumors and cancer survival (*e.g.*, clearly stated that the population of interest was with cancer at the start of the study). Regarding study design, we included prospective and retrospective cohorts, case-control, and cross-sectional studies (including ecologic studies), without language restriction. We excluded case reports, reviews, non-original reports, and toxicological studies (if any). Conference abstracts were included if sufficient information was reported, including population, exposure, outcome type, and effect size (**Table S1**).

Information sources and search strategy

We used three databases including Medline (searched through PubMed), Scopus, and Web of Science, and systematically searched the available literature based on population, exposure, and outcome of interest up to September 03, 2021. Combinations of relevant search terms were used in each database (**Table S2**). We also conducted a manual search of the reference lists of retained original studies or relevant reviews to identify additional documents.

Study selection and data collection process

After the completion of the searching of the databases and references of the selected reviews and papers, we merged the results and then excluded the duplicate items. Duplicate removal and study selection process were performed using the Rayyan platform [20]. The unit of analysis in

this review was “the study” and in the case of multiple studies per one record (*e.g.*, reporting the results of multiple studies in one paper), we used the data of each study and outcome separately in the review. In the case of finding more than one record per study (*e.g.*, results of a specific study had published in more than one record), we used all of the related records for the review, but for the meta-analysis, we only included the one with the lowest risk of bias. Two independent reviewers (M.Z. and J.Y.) screened the retrieved unique records according to the titles and abstracts, considering the study's inclusion and exclusion criteria. The full-text assessment was performed on the selected records based on the titles and abstracts, and then all studies that reported at least one relevant exposure-outcome association were retained at this step. In the case of disagreement between two reviewers, a third reviewer (B.J.) helped to resolve the conflict.

Data collection process and data items

All relevant data from each retrieved study in seven categories including general study characteristics, population selection and characteristics, exposure assessment and location, outcome measurement and types, details of statistical methods, findings and effect sizes, and funding and conflict of interests were extracted. The first author's name (hereafter as the study ID), publication date, location of study, study design, and the number of participants (cancer cases / total population or controls) were extracted as general characteristics of the study. Source of health data, duration of follow-up for cohort studies, percentage of participation, data coverage duration, and specific population subgroup (sex, age, region of residence) were extracted as population characteristics. Data on exposure location (home, work), type of exposure indicator, source of exposure data (*e.g.* satellite name, land cover maps, etc.), the spatial resolution of exposure data (classification accuracy for land covers), buffer sizes, time point used for the construction of exposure measure (*e.g.* greenest time of the year, summer,

etc.), exclusion of specific type of greenspace or data from exposure maps, and exposure level (in term of descriptive values of greenspace measures) were used for the description of the exposure. Type of cancer, type of outcome (incidence, prevalence, mortality), and its corresponding International Classification of Disease codes (ICD) were extracted in the outcome section. In the statistical section, we extracted data on the type of statistical model, linearity and trend analyses, methods of mediation analysis, effect modification, availability of sensitivity analysis, levels of adjustment, and main findings (including effect estimates and 95% confidence intervals). We also extracted the reported effect estimates including odds ratio (OR), relative risk (RR), hazard ratio (HR), and correlation coefficients and corresponding 95% confidence intervals or other precision estimates. For those studies that reported their findings only in figures, we used Web Plot Digitizer software to extract the values from the figures [21]. In the funding and conflict of interests domain, we extracted declarations of the funding and conflict of interests. For all of the abovementioned items, we considered it as "unclear" if it was not reported or reported vaguely in the record.

Risk of bias assessment

The risk of bias of the selected studies was assessed by the Newcastle–Ottawa Scale (NOS) [22]. Briefly, the NOS is composed of three domains, including: (i) the selection of study groups/ participants, (ii) the comparability of the cases and controls (or of exposed and non-exposed), and (iii) ascertainment of exposure/outcome. The NOS scoring is based on giving a star to each of the "items" in the predefined "domains". For cohort and case-control studies all items except the comparability can earn one star as a scoring mark (the comparability item can earn a maximum of two stars according to the level of adjustment). Summing up the earned stars (maximum of nine) will construct the global score of each study (maximum attainable grand

score is 9). We used a modified version of the NOS for cohort studies to assess the risk of bias of cross-sectional and ecological studies, as is used previously [23, 24]. It was comprised of seven items (four items at selection with a maximum of four stars; one item at comparability with a maximum of two stars; and two items at outcome section with a maximum of three stars). **Table S3** provides the details about the criteria we used to rate the studies in each risk of bias domain.

In this study, we modified the NOS items slightly to be applicable in the field of environmental epidemiology, especially for exposure assessment and ascertainment. Among the different aspects of greenspace exposure assessment (spatial and temporal resolution, life-time coverage, location of exposure allocation (*eg.*, home, working address), etc.) we used “the location of exposure allocation” as a selected aspect for quality assessment. We considered two parameters for defining the highest exposure assessment quality in the risk of bias assessment section. These factors include a) exposure assessment by the addresses geocoded at the residential location (not the workplace, etc.), with the address either recorded during a field survey, or obtained from an administrative register, etc., and b) exposure allocation at a personal level (compared to city or community level exposure allocation). All studies fulfilling these two parameters ranked as the highest quality in exposure assessment. We did not find a reason to consider the different methods of individual-level residential location geocoding differently (for example addresses come from national registries, online databases, face-to-face interviews, or geocoding by approaching directly to the residential addresses). We also assumed that dynamic exposure assessment, such as GPS-tracking, could not yet be considered a realistic benchmark in cancer epidemiology, therefore home-based exposures received high-quality ratings.

We defined age, sex (except for sex-specific cancers such as prostate and breast cancer), personal level or area-level socioeconomic status (such as education, income, or area-level deprivation),

smoking, and urbanity (regardless of the level of assessment) as the minimum cut-set of variables necessary for adjustment according to the literature and expert discussion. Studies with the inclusion of the possible mediators in the models (such as air pollution, mental health, stress, physical activity, body mass index, type-2 diabetes, cardiovascular diseases), without reporting the results of the model without adjustment for the mediators, or adjusted for ancestors of the mediators were penalized because of a possible risk of overadjustment. Those studies that reported the results of models additionally adjusted with a possible mediator in addition to the main model gained a second star (in the case of reporting associations for both modes of adjustment with and without mediators, we used the estimates from the models without mediators).

We reported the NOS score (summing up the attained stars) as a general measure of the risk of bias of each study. Different cut-offs are reported to categorize the studies based on the NOS score. While definition of cut-offs is not stated in the NOS guideline, we defined five categories of risk of bias according to the NOS score which also is reported in other studies (including: "very good " for NOS score ≥ 8 , "good " for NOS score six or seven; "fair " for NOS score four or five, and "poor " for NOS score < 4) based on the previously published similar studies [24].

Synthesis methods

Given the heterogeneity in the reported effect sizes and paucity of studies for each exposure-outcome association, we could not do a meta-analysis for most of the exposure-outcome pairs. Therefore, for those pairs with enough papers (reported in at least two studies *e.g.* all studies on breast cancer incidence reporting the exposure based on the Normalized Difference in Vegetation Index (NDVI), and reported the effect size which were possible to be converted to each other)

we did a meta-analysis, and for those without enough studies, we just described their characteristics and findings. As there is a fundamental difference between effect sizes such as OR and HR measures, in the main meta-analyses we did not merged these effect sizes together. Due to the small number of studies for each exposure-outcome pair, our tests for heterogeneity could have been underpowered to detect heterogeneity. As such, we took a more conservative approach and used the inverse variance heterogeneity meta-analysis to produce the combined effect size for all the evaluated exposure-outcome pairs [25]. Overall heterogeneity in the reported associations was assessed using Cochrane's Q and the proportion of total variation in the point estimates that is attributable to between-study heterogeneity was quantified by applying the I^2 statistic (and categorized as not important, moderate, substantial, and considerable heterogeneity) [26, 27]. The meta-analyses were conducted using the “*metafor*” package in R [28].

Effect sizes across the studies were reported according to the different levels of exposure. For example, the effect sizes in the studies with exposure to NDVIs, were reported based on different values of increase in NDVI such as 0.1, 0.2, or an interquartile range (IQR). Therefore, we transformed the effect sizes to the homogenous unit before entering into the meta-analysis (*e.g.* all of the NDVI effect sizes were converted per 0.1 unit increase of NDVI) (**Supplementary method**). In the studies that reported the effect estimates based on the “per quartile” (or tertile, etc.), we considered estimations in each group (tertile, quartile, etc.) as a separate population and entered them separately into the meta-analyses. In the case of reporting several effect estimates based on different sub-groups in a study and not for the whole population (for example separately for men and women, or at different quartiles of exposure), we at first linearized (according to the procedure described in the **Supplementary method**), then combined them under fixed-effect meta-analyses to reach study level estimate. Finally, we used that study-

specific estimate in the random effects meta-analyses. If the outcomes were reported at different buffer sizes, we preferred a 300 m buffer size in the meta-analysis.

Overall quality of evidence assessment

We evaluated the overall quality of evidence for each outcome using the Grading of Recommendations Assessment, Development, and Evaluation system (GRADE). The GRADE evaluates the overall quality of available evidence in four categories: “high”, “moderate”, “low”, and “very low” [29, 30]. It starts the evaluation by attributing a score from the study design and then uses eight domains to modify this score. In its original format for clinical practice recommendations, observational studies were considered of low-quality. As all studies included in our review were observational (which is quite common in environmental epidemiology), we modified the original GRADE methodology, and as a starting point we considered the cohort and case-control studies as the sources with high-quality evidence, and cross-sectional and ecologic studies as the sources with low-quality evidence. The original score can be upgraded / downgraded according to five downgrading (risk of bias, inconsistency, imprecision, indirectness, and publication bias) and three upgrading (dose-response trend, the magnitude of associations, and residual confounding) domains. For the risk of bias, representativeness of population, the origin of controls, and inadequate control of confounders were considered [31]. Heterogeneity in the effect sizes and non-overlapping of reported confidence intervals were considered as the measures of inconsistency [32]. Imprecision was considered as a small number of studies (less than three) or studies showing associations in opposite directions for the same pair of exposure-outcomes (as a measure of wide 95% confidence interval) [33]. The accordance of the population, exposure, and outcome of the studies to the targeted population, exposure, and outcome of this review was considered as a measure of indirectness [34]. In this review, deciding

about publication bias was difficult due to the limited number of available studies for most of the outcomes and hence the infeasibility of formally testing the publication bias using for example Egger's test and funnel plots. Therefore, we could just consider the omission of reporting certain results in the included papers as a possible source of publication bias [35]. Large magnitude effect size, observation of dose-response trend, and small likelihood of residual confounding were considered for upgrading [30]. In the original GRADE protocol for the magnitude of the effect sizes of greater than two and five were considered for one and two levels of upgrade in the quality level respectively. But this magnitude of the effect sizes is not applicable in the field of environmental epidemiology, where the effect of most of the exposures is below these values. Therefore, borrowing from the estimates reported in the studies on air pollution and cancer, reporting of effect size above 40% increase or decrease (*e.g.*, OR of greater than 1.4 or below than 0.6) considered for one level upgrading [24].

Results

Studies selection

Through systematic searches, we found a total of 8904 unique records (**Figure 1**). We also found three relevant documents through reference searching (two journal articles [36, 37], and one thesis [38]). After duplicates removal and titles and abstracts screening, we selected 40 articles for full-text assessment, of which, 22 articles were excluded because they did not meet the inclusion criteria (**Table S4**). Finally, 18 studies, including six cohort studies [7, 12, 18, 26, 39, 40], four case-control studies [6, 10, 41, 42], and eight cross-sectional studies [8, 9, 11, 43-47], were entered into this review (**Figure 1**).

Study characteristics

Table 1 presents the characteristics of the selected studies. All included studies were conducted from 2008 [44, 47] onwards, and 72% (n=13) since 2015 [6-8, 10-12, 18, 38-43]. All of the records were based on unique data sets (study) and no overlap of the datasets across studies was present. Eight studies were conducted in Europe [10, 12, 18, 39, 41, 43, 44, 46], four in North America [6, 7, 38, 40], four in Asia [8, 11, 42, 47], and two in Oceania [9, 45]. Four studies performed their analyses on area-level aggregated health data (area level cancer mortality rate) [8, 11, 46, 47]. In the studies using individual-level cancer data, the number of cancer cases ranged from 122 in a cohort study on breast cancer [38], to 67463 in a cross-sectional survey for skin cancer prevalence [9].

Risk of bias

The overall quality of the studies (in terms of risk of bias according to the NOS) ranged from 1 [47] to 8 [40]. One study (6%) earned a very good quality score [40], six studies (33%) had good quality scores [6, 7, 10, 18, 38, 41], six studies (33%) had fair quality scores [9, 12, 39, 42, 44, 45], and five studies (28%) had poor quality scores [8, 11, 43, 46, 47] (**Table 2**). Four out of six cohort studies had good or very good quality [7, 18, 38, 40], and two other cohort studies had fair quality [12, 39]. Three out of four case-control studies had very good or good quality [6, 10, 41], and one remaining study had fair quality [42]. In contrast, five out of eight cross-sectional studies had poor quality [8, 11, 43, 46, 47], and three remaining had fair to good quality [9, 44, 45].

In the cohort studies, three items including sample size, exposure assessment, and outcome characterization had the highest average quality score (attained star in all of the cohort studies;

100%), whereas the length of follow-up received the lowest quality score (attained star in 66% of the cohort studies). In the case-control studies, four items (including case definition, representativeness, exposure assessment, and outcome characterization) had the highest average score (100%). Reporting of non-respondent, and comparability were the weakest items in the case-control studies (with zero and 37% average scores, respectively). In cross-sectional studies, the strongest items were representativeness of the population and reporting of statistical tests (both with an average 75% score). The average score for the rest of the items in the cross-sectional studies was lower than 50%, with the reporting of non-respondents and assessment of exposure having the lowest average scores (0 and 12.5%, respectively).

Statistical aspects

The statistical aspects of the included studies are presented in **Table 3**. Different levels of adjustment were employed across the studies from not adjusting at all [47] up to adjustment for more than 15 variables [18], for different groups of covariates including demographic, lifestyle, comorbidities, contextual variables, and environmental factors (**Table S5**). Age, area-level deprivation, individual-level socioeconomic status (e.g. income, education, socioeconomic, or occupational status), smoking status, and sex were the most commonly adjusted variables (respectively in 14, 13, 10, 10, and 8 studies). Of nine studies on lung cancer outcomes [8, 11, 12, 18, 42-46], five adjusted for smoking status as a well-known and strong risk factor for lung cancer (four at the individual level [12, 18, 42, 46], and one at area level [45]). All studies on prostate cancer [6, 18, 39, 40] adjusted for age as the most important known risk factor of prostate cancer. For this outcome, only one study adjusted the models for passive smoking and smoking intensity in addition to the smoking status [18]. Adjustment for skin color was performed in one study [9], out of three, which reported associations for skin cancer [9, 18, 39].

Among the retrieved studies, mediation analyses were performed in five studies [6, 7, 9, 40, 41], mostly for/with physical activity (n=5) [6, 7, 9, 40, 41], followed by air pollution (n=3) [6, 7, 41], social engagement [7], body mass index [6], time spent outdoor [9], and mental health [7]. Only two of these studies conducted formal casual mediation analysis reporting direct and indirect associations [7, 40]. Demoury *et al.* [6] did sequential adjustment by introducing mediators in a separate model and based on the negligible change in the estimates inferred lack of mediation of the association. One study used the product of coefficients approach [9], and another study just performed additional adjustment for potential mediators, and because of the absence of a change in the associations in the models with additional adjustment, they did not continue the mediation analyses [41]. Except for James *et al.* [7], none of the studies found a significant mediating effect of the studied variables. James *et al.* [7] found a significant mediating effect of air pollution, physical activity, social cohesion, and depression (as a marker of mental health). The strongest mediating role in James *et al.* [7] was for mental health situations (19.9%), followed by social cohesion (9.7%).

Exposure-response linearity was checked in seven studies [6-8, 12, 38, 40, 43] by applying exposure quantiles and trend analysis (n=3) [6, 7, 38, 40] and spline function for exposure and subsequent likelihood ratio test to compare the models with and without the spline function (n=4) [7, 8, 12, 40]. One study assumed a nonlinear association from the beginning and reported the association based on quintiles [43]. Four studies showed no deviation from the linearity assumption [6-8, 12]. One study found a nonlinear association between exposure and outcome based on the tertiles of exposure (found reverse U-shape association [38]). Another study reported heterogeneous findings in different subgroups or according to type of exposure (linear association for those whose were living in high density areas, and no significant trend in the total

population; no significant trend in the exposure-response association for the NDVI at home, whereas a significant trend for NDVI at work) [40].

Exposure assessment

Different greenspace exposure indicators with different definitions and quantification methods were used across the included studies (**Table 4 and S6**). We classified the reported exposure indicators in three groups – greenspace availability (which was reported in all studies as residential surrounding greenspace and was measured based on the average of NDVI or the percent of greenspace land cover at different buffer sizes around the residential address), greenspace accessibility (such as home or work address proximity to different types of green space), and greenspace use. Twelve studies reported results of surrounding greenness (all based on NDVI) [6-8, 11, 12, 18, 38-41, 47]. In four studies, the NDVI values were assigned to the participants at the area level (i.e. city or ward) [8, 11, 39, 47]. One study reported the results for NDVI at both home and work addresses [40]. In addition to using NDVI at different overlapping buffer sizes, one study also reported the results based on NDVI at nested buffers (for example “nested buffers of 100 m and 500 m” means buffer zone between 100 m and 500 m) [41]. Five studies used the Moderate Resolution Imaging Spectroradiometer (MODIS) instrument images (in four studies from the Terra satellite [7, 8, 11, 42] with 250-m spatial resolution, for one study [39] it was not clear), five applied different Landsat’s images (with 30-m spatial resolution) [6, 12, 18, 38, 41], and two other used the Advanced Very High-Resolution Radiometer (AVHRR) from the National Oceanic and Atmospheric Administration (NOAA) satellite images (with 1000-m spatial resolution) [40, 47]. Eight studies used NDVI in a specific period of the year (mostly spring and/or summer as the greenest seasons of the year) [6, 7, 11, 12, 18, 41, 42, 47]; eight other studies used NDVI averaged over different seasons or time points in a year [9, 10, 39,

40, 43-46]. In two studies, the time of the year for satellite image acquisition was not clear [8, 38].

In studies with measurements at the individual level, NDVI was measured in different buffer sizes ranging from 100-m [18, 41] up to 1250-m [7]. Two studies did not report the NDVI buffer size [8, 38]. NDVI at 300-m buffer size was the most commonly reported [6, 12, 18, 41]. Mean NDVI values (regardless of reporting at individual or area level) in the studies ranged from around 0.20 [40, 41] to 0.50 [8, 12, 18, 39].

Percentage of land cover with greenspace (e.g. total, urban, agricultural green spaces) was reported in nine studies [8-10, 12, 41, 43-46], using maps with different minimum mapping units for the greenspaces from 5m² [12] up to 25 hectare [10, 18, 41]. The minimum mapping unit in three of the studies was not clearly indicated [8, 9, 12]. In addition to the percentage of land cover, one study reported the results based on the presence or absence of different types of green space in different buffers [41]. Four studies reported the percent of different greenspace land cover types in specific buffers around the home addresses [9, 10, 12, 41]. Considering all individual-level and city-level reports on the percentage of green space coverage, it ranged from 19% for the percent of green space at 300-m buffers in the Netherlands [12] to 46.2% at the city-level in the UK [46]. The percentage of green space coverage was not reported in two studies [10, 44].

Proximity (inverse distance) to green space was also reported in two studies [10, 18]. One study reported the proximity associations separately for urban, agricultural, and forest green spaces based on the Coordination of Information on the Environment (CORINE) land cover map [18]. Another study calculated a lifelong cumulated exposure score for each type of greenspace (open

field crops, vineyards, orchards) by multiplying inverse distance by residential duration at each address [10].

Outcomes

Reported outcomes

Four studies reported incidence [18, 38-40], five studies reported prevalence [6, 9, 10, 41, 42], and nine others reported cancer mortality [7, 8, 11, 12, 43-47]. Hereafter, we summarize the evidence on each cancer site ordered by the number of available studies for each outcome.

Lung cancer

Nine studies (six cross-sectional [8, 11, 43-46], two cohort [12, 18], and one case-control [42]) reported association between exposure to greenspace and lung cancer incidence [18], prevalence [42], or mortality [8, 11, 12, 43-46]. The total number of lung cancer cases that were reported in seven studies on individual-level data was 63,617, ranging from 237 [18] up to 30,110 cases [46] (two of the studies did not report the exact number of lung cancer cases [8, 11]). Exposure-outcome associations were reported based on the percent of land cover [8, 12, 43-46], surrounding greenspace [8, 11, 12, 18, 42], and distance to greenspace (including forests, agricultural lands, and urban green spaces) [18]. The outcome definition in five studies was based on both ICD-10 C33 and C34 together (malignant neoplasm of trachea, and malignant neoplasm of bronchus and lung, respectively) [8, 11, 42, 43, 46]. However, three other studies just included ICD-10 C34 [12, 18, 44]. In another study, the ICD code was not reported [45]. Among seven studies on lung cancer mortality, two studies found a significant beneficial association between greenspace exposure and lung cancer mortality [8, 44], whereas one study

reported greenspace exposure as a significant risk factor [46]. Two studies on lung cancer incidence and prevalence found no significant associations [18, 42] (**Table S7**). We conducted three separate meta-analyses for studies reporting the effect estimates based on greenspace availability (all reported based on NDVI), and a) incident rate ratio or risk ratio, b) HR, and c) OR/RR. In all three meta-analyses we found no association in the combined estimates for 0.1 unit increase in NDVI (combined ratios: 0.99 [95%CI: 0.85-1.15; I^2 : 0%; Cochrane's Q test $p = 1$]; combined OR/RR: 1.01 [95%CI: 0.84-1.20; I^2 : 0%; Cochrane's Q test $p = 0.96$]; combined HR: 1.00 [95%CI: 0.87-1.28; I^2 : 0%; Cochrane's Q test $p = 0.75$], combined OR: = 1.01; 95%CI: 0.80-1.27; I^2 : 0%; Cochrane's Q test $p = 0.49$) (**Figure 2**).

Prostate cancer

Four studies including three cohort studies (from the USA [40], France [18], and Germany [39]) and one case-control study (from Canada [6]) reported the associations based on a total of 13843 cases of prostate cancer (ranging from 898 [40] to 9611 [39] cases) (three studies reported incidence [18, 39, 40], and one study reported prevalence [6]). All studies used the NDVI. Proximity to different green spaces (agricultural, urban, and forest) was only reported in one study [18]. The outcome in two of the studies was defined as ICD-10 C61 [18, 39]. However, for two other studies, the exact system of outcome coding was not clear, and one study limited the outcome to "lethal prostate cancer" instead of total prostate cancer cases [40]. The reported associations for surrounding greenspace went in opposite direction, from beneficial [6, 39, 40] to detrimental [18, 40]. Three out of four studies found reduced risks of prostate cancer with increasing NDVI (by 16% [6], 10% [40], and 5% [39] reduction in risk) in their full sample analysis, whereas another cohort study found a non-significant increase of the risk with an increase in surrounding greenspace or decrease in the distance to different types of greenspaces

[18]. In addition to the heterogeneity in these findings, differential associations were also reported across different subgroups within the same studies. For example, the “Health Professionals Follow-Up Study (HPFS)” found a significant beneficial association for the surrounding greenspace in high-density areas (HR=0.90; 95%CI: 0.82–0.99) and non-significant, but a point estimate above one for the low-density areas (HR=1.11; 95%CI: 0.95–1.29) [40] (**Table S8**). We conducted two separate sets of meta-analyses for a) studies reporting the effect size based on OR/RR, and b) studies reporting HR as their effect size estimates. In three of the studies, effect sizes for NDVI at 1000 m buffer size were reported and therefore we selected this buffer for the meta-analyses. The buffer size was not reported in one of the included effect estimates. In both of the meta-analyses, the combined estimates (for 0.1 unit increase in NDVI) were not suggestive for beneficial or detrimental role of exposure to greenspace (combined HR= 0.98; 95%CI: 0.69-1.40; I^2 : 0%; Cochrane’s Q test $p = 0.81$, and combined OR = 0.91; 95%CI: 0.69-1.20; I^2 : 0%; Cochrane’s Q test $p = 0.68$) (**Figure 3**).

Breast cancer

Four studies (three cohort [18, 38, 39], and one case-control [41] studies) with a total of 11,206 cases of breast cancer (ranging from 122 [38] to 9577 [39] cases) reported the associations for breast cancer incidence [18, 38, 39], and prevalence [41]. The associations were reported for the surrounding greenspace based on NDVI [18, 38, 39, 41], the percentage of the greenspace land cover [41], and the proximity to different types of green space (agricultural, forest, and urban green spaces) [18]. The outcome definition in three studies was based on the ICD-10 code of breast cancer (C50), but in one study it was unclear [38]. The reported associations were inconsistent within and between the studies, depending on the type of exposure indicators used. Of four studies with surrounding greenspace as an exposure indicator, two found a significant

beneficial association [18, 39], one found no significant association [38], and one found the surrounding greenspace as a significant risk factor [41]. In addition, the use of different exposure indicators in one study resulted in heterogeneous associations from significantly beneficial (*e.g.*, for the presence of the urban greenspace) to significantly detrimental (*e.g.*, for the surrounding greenspace and the presence of agricultural green space) [41] (**Table S9**). We conducted two separate meta-analyses for studies reporting their effect estimates based on a) OR/RR and b) HR (as these two groups of effect sizes were not convertible). The combined estimates (for 0.1 unit increase in NDVI) were heterogeneous according to type of effect size (HR= 0.83; 95%CI: 0.47-1.48; I^2 : 0%; Cochrane's Q test $p = 0.88$, and combined OR: = 1.01; 95%CI: 0.80-1.27; I^2 : 0%; Cochrane's Q test $p = 0.49$) (**Figure 3**).

Skin cancer

Two cohort [18, 39] and one cross-sectional [9] studies (three studies in total) reported their findings based on a total of 109430 cases of skin cancer incidence [18, 39], and prevalence [9] (**Table S10**). Different exposure indicators were used across the studies including the surrounding greenspace based on NDVI [18, 39], the percentage of land cover [9], and the proximity to different green space types [18]. Studies were heterogeneous in terms of their outcome definition with one study defining it by including both melanoma and non-melanoma skin cancer [9], one study including only melanoma skin cancer [18], and one study with no specified malignant cancer of skin [39]. Findings of the cohorts were contradictory, with one cohort study reporting a 16% lower risk of non-melanoma skin cancer (HR=0.84; 95%CI: 0.79-0.90) [39], and another cohort finding a non-significant increase in the risk of melanoma skin cancer with an increase in surrounding greenspace or proximity to different green spaces [18]. A cross-sectional study in Australia, which merged the melanoma and non-melanoma skin cancers,

found a significant 6 to 11% increase in the risk of skin cancer prevalence (depending on the different quantiles of exposure) in those residing in areas with higher greenspace land cover compared to those residing in the area with less than 20% of greenspace land cover [9]. However, conducting separate analyses for melanoma and non-melanoma skin cancers, they found significant results only for non-melanoma skin cancer. The heterogeneity of exposure-outcome pairs and reporting of the effect sizes based on the different type of effect size prevented us from doing a meta-analysis.

Colorectal cancer

Two cohort studies (one in Germany [39], and another in France [18]) reported the associations for colorectal cancer based on 12,343 cases of colorectal cancer incidence (**Table S11**). Exposure was reported based on surrounding greenspace (NDVI) [18, 39], and distance to green spaces [18]. The outcome definition in one of the studies included the cancers of colon, rectosigmoid, and rectum [18], whereas the other one additionally included malignant neoplasm of anus and anal canal as their outcome [39]. Associations for surrounding greenspace varied from significantly beneficial (RR: 0.84; 95%CI: 0.79-0.90) [39] to null associations [18]. The variation in the reported exposures and outcomes and the paucity of data prevented us from conducting a meta-analysis.

Other cancers

Five other outcomes were reported across the studies including two cohort studies on all-site cancers incidence [18] and all-site cancers mortality [7], a case-control study on brain cancer prevalence (meningioma and glioma cases) [10], a cohort study on mouth and throat cancer

incidence [39], and a cross-sectional study for the esophageal cancer mortality [47]. The findings of two studies on all-sites cancer incidence and mortality were contradictory. While one found an increased risk of cancer incidence with increased surrounding greenness [18], the other one found a significant decrease in the risk of all-site cancers mortality with an increase in surrounding greenness [7]. The results of the study on brain cancer were dependent on the type of brain tumor and type of the greenspace, as a significant increase in the odds of having meningioma was observed for proximity to open field crops, whereas a non-significant increase for proximity to vineyards, and orchards [10]. The study on mouth and throat cancer found that surrounding greenspace had a stronger beneficial association with the mouth and throat cancer compared to colorectal and non-melanoma skin cancer incidence [39]. The study on the esophageal cancer mortality reported the results as a correlation coefficient, and the findings were in opposite direction for different seasons: a positive but non-significant correlation between NDVI measured in January and esophagus cancer mortality, but a negative and non-significant correlation between NDVI measured in the summer and esophagus cancer [47].

Certainty of the available evidence

The GRADE approach was used for the five exposure-outcome pairs that were investigated by two or more studies (lung cancer mortality, prostate cancer incidence, breast cancer incidence, skin cancer incidence) (**Table S12- S15**). In all of these cases, the certainty of the evidence was very low. Indirectness due to the few numbers of studies for each pair, and inconsistency of the findings were among the reasons that downgraded the quality of evidence in reviewed exposure-outcome associations.

Discussion

We conducted a systematic review and meta-analysis of the available evidence on the association between greenspace exposure and all-sites and site-specific cancer incidence, prevalence, and mortality in adults. Among 18 studies included in our review, cross-sectional studies were the most common study design. Heterogeneity was observed in terms of the location of studies with no available studies being conducted in Africa, the Middle East, and South America. Among 15 available unique types of exposure-outcome association, only five of them were reported in two or more than two studies (breast cancer incidence, colorectal cancer incidence, prostate cancer incidence, skin cancer incidence, and lung cancer mortality). Exposure assessment across the studies was done by using greenspace availability and accessibility measures. None of the studies used greenspace use or quality measures. For cohort and case-control studies the average quality score was good, while the average quality score of cross-sectional studies was poor.

The general observed associations between greenspace exposure and breast and prostate cancers in the selected studies in this review - which suggest a potential beneficial association- should be interpreted with caution. Indeed, given the wide confidence interval of the pooled estimates, the small number of studies, and the very low overall quality of evidence, further studies are necessary to draw a conclusion about the beneficial/null/detrimental association. For skin cancer, the available evidence suggested greenspace as a potential risk factor, however the findings were heterogenic according to the melanoma and non-melanoma types of skin cancer. For the rest of the cancers (all-sites cancers, lung cancer, colorectal cancer, brain cancer, mouth and throat cancer, and esophageal cancer) the evidence was not conclusive. The weight of evidence based on GRADE classification for all of the exposure-outcome pairs was very low.

To the best of our knowledge, there are only two reviews available on our evaluated association and none systematically evaluated the evidence [4, 5]. Our review included 18 studies, the previous scoping review [5] included only half of the studies which we included in the review. The limited number of available studies and the heterogeneity of reported associate estimates prevented us to conduct a meta-analysis for most of the exposure-outcome pairs. We could only perform meta-analysis for studies on lung, breast, and prostate cancer. None of the pooled estimates significantly associated with greenspace exposure; however, the pooled HR for breast cancer was suggestive of a possible protective association.

Exposure assessment

Issues such as temporality between exposure and outcome, the type of greenspace evaluated, the type of greenspace indicator, the spatial resolution of the data sources, the applied buffer sizes, and the location of exposure allocation (home or work address) are among the exposure-related factors that could be regarded as a source of heterogeneity and bias in the studies. Temporality is an essential element in establishing causality [48, 49]. Considering that the cancer initiation process occurs even decades before its clinical manifestation, theoretically looking for exposures far before the outcome occurrence is preferable. Most of the studies in this review established the exposure-outcome association based on exposure at a single time point, before or at the time of outcome occurrence, and even after outcome occurrence.

Reported associations based on the type of greenspace were heterogeneous and in some cases were in contrast even within the same study. For example, in a study on breast cancer, the authors found an increased risk of breast cancer for proximity to agricultural lands and increased greenness, but a significant decrease in the risk of breast cancer for the presence of urban

greenspaces around the residential address [41]. Even in the category of “agricultural lands”, different associations were reported according to different agricultural lands which could be interpreted as a proxy for use of different types and quantities of pesticides (for example a significant increase in the risk of brain cancer for open field crops, but not for vineyards or orchards [10]). From this viewpoint, the type and application of greenspace in the urban and rural areas could be different (rural greenspace are mostly agricultural lands, whereas urban greenspace is usually for street side and parks). Therefore, merging urban and rural participants in the same models could be responsible for part of the observed heterogeneity.

The capacity of greenspace in the mitigation of environmental hazards such as air pollutants has been proposed to be associated with vegetation structure, composition, and management. Less managed and complex greenspaces might work better in mitigating air pollution [50]. Therefore, it is advisable to distinguish urban greenspace (usually managed and less complex) from forests (with higher complexity and less managed) in the models [51]. Additionally, considering mitigating air pollution as a potential mediator between greenspace exposure and cancer, it is necessary to consider the role of different types of greenspaces and urban/ rurality in the models. Part of the findings in the included studies in this review such as Iyer *et al.* [40], that found an inverse association between exposure to greenspace and lethal prostate cancer incidence in high-density and low-density areas might be justifiable by this factor.

There is no consensus on preferring one buffer size over another in the studies on greenspace and health, given that this choice of the buffer is very context-specific. However, 300m buffer (as access to public greenspace within 300m of linear distance to home address has been proposed by the World Health Organization (WHO) /Europe [52]) was applied in four studies. Only one study in addition to modeling the associations for exposure assessment at a residential address

reported the results for greenspace exposure at work address [40]. The findings based on greenspace exposure at work address were suggestive of the protective role of greenspace when the results were analyzed for the whole cohort [40]. However, a separate analysis by the greenspace exposure at home or work address showed an increased risk for those who had exposure assessment based on geocoding at the home address.

None of our reviewed studies considered the quality of green spaces. The quality characteristics such as safety, access, biodiversity, aesthetics, facilities, and amenities could influence the potential of these spaces to exert their health benefits [53, 54]. Similarly, none of the available studies looked into other aspects of exposure to greenspace such as the use of green spaces, or visual access to greenspace. Future studies are recommended to also include these aspects of greenspace exposure in their analyses.

Statistical approaches

Our reviewed studies varied in terms of the covariates for which they controlled their analyses. Most of the studies used sequential adjustment procedures based on *a priori* knowledge, and few studies used specific statistical approaches such as directed acyclic graphs (DAGs), suitable selection algorithm, or univariate analyses. The risk of over-adjustment was present in some of the studies for example by the inclusion of both degree of urbanity and population density in the same model [8, 44]. On the other hand, a number of the studies failed to include influential variables on cancer outcomes such as smoking and alcohol consumption in their model [8, 11, 38, 41, 43, 44, 47]. Therefore, the level of adjustment could be regarded as a potential factor in the heterogeneity between the available studies. For example, Richardson *et al.* [46] found a

significantly increased risk of mortality from lung cancer in women living in an area with moderate greenness compared to those in low greenness. The association disappeared after adjustment for smoking. Smoking is a strong correlate of personal and neighborhood deprivation. It is also likely that areas of high deprivation have less greenspace and a higher rate of cancer outcomes, therefore, in the models which included smoking (as it acts as both a cancer risk factor and potential proxy for deprivation level) there is a chance of not seeing an association. Similarly, Astell-Burt *et al.* [9] found attenuation of effect sizes after additional adjustment for socioeconomic status, demographics, and cultural characteristics.

The socioeconomic status could act as a strong confounder in the studies of the health effects of greenspace exposure. Ideally, such studies should control for indicators of socioeconomic status at both individual (e.g. education, occupation type, income) and neighborhood (e.g. census-based indices of deprivation) levels in order to minimize the likelihood of residential socioeconomic confounding. Of our reviewed studies, 10 adjusted for the individual-level indicators of socioeconomic status, 13 adjusted for neighborhood-level indicators, and seven adjusted for both.

Possible mechanisms

The mechanisms behind the association between greenspace exposure and cancer are yet to be established. Here we discuss the observed evidence on the underlined mechanisms through two broad categories of environmental and socio-behavioral benefits /or deleterious impacts of greenspace on health.

a) Sociobehavioral benefits of greenspace on health including increased physical activity [55, 56], enhanced social cohesion [57], and reduced stress [58, 59]. The strongest mediating role in James *et al.* [7] was found for mental health which was able to explain up to 30.6% of their estimated associations followed by social cohesion (19.1%). In this review, from five studies that examined the mediating role of physical activity, only one study found significant mediation of this variable. Higher levels of physical activity have been associated with reduced risk of cancers [60]. Even though the mediating role of physical activity in the association between greenspace and health has been reported more than any other hypothesized mediators in the literature on greenspace and health, it is still inconsistent [61]. It seems that other characteristics of greenspace such as quality and safety are more important in promoting physical activity than just quantity and presence of greenspace [62]. Additionally, it is probable that physical activity may not mediate the association at all or to the same extent in different areas. It could be also difficult to accurately measure long-term physical activity over the entire time window of exposure relevant for cancer development.

b) Possible environmental benefits include reductions in heat [63], noise [64, 65], and air pollution [65, 66], and enriching microbiota diversity [67]. Contrary to the available literature on the mediating role of air pollution in the greenspace-health association (60% of literature reported mediating role of air pollution) [61], only one out of three studies in this review found the mediating role of air pollution [7]. Evidence is available on the association between exposure to the factors used as a mediator in these studies (such as air pollution, physical activity, body mass index, and mental health situation). Additionally, the association between exposure to greenspace and these factors is reported in the literature. Therefore, from a theoretical point of view, considering these variables as a mediator in the association between greenspace exposure and cancer outcomes is reasonable.

There are ongoing debates on a potential link between greenspace exposure and increased exposure to pesticides (especially due to proximity to agricultural fields) [17], allergens, and certain volatile organic compounds (VOCs). The link between pesticides exposure and cancer (for example brain cancer) has been reported in several studies [10, 68]. Increased risk of breast cancer with proximity to agricultural lands in the study of O’Callaghan-Gordo *et al.* [41] might partly be explained by this pathway. Considering the association between higher greenspace and higher levels of outdoor physical activity, the observed detrimental association between greenspace and skin cancer in Astell-Burt *et al.* could be due to increased sun exposure [9].

All in all, although a number of studies have explored the mediation role of some of the potential mediators of the association between greenspace exposure and cancer, establishing the mechanisms underlying this association has remained an open question. Moreover, the exact applied mediation analysis approaches by the available studies were not detailed in most of the studies [6, 9, 41], and formal mediation analyses and reporting of direct and indirect proportions of associations were not reported in those studies.

Strengths and limitations

This review has several strengths including being the first systematic review on the available evidence on cancer, as far as we are aware. Additionally, we tried to explore the methodological issues, especially in terms of exposure assessment and statistical approaches employed in the published works. We included different cancer outcomes including incidence, prevalence, and mortality in this study (however reported the results separately) to reduce the risk of reporting bias in our review. We also assessed the quality and certainty of the evidence for each study and

exposure-outcome pairs respectively, which can give an insight for the rooms for developments of future studies. In this review, the reported associations for exposure-outcome pairs were not sufficient to do a meta-analysis for most of the outcomes. Even in the case of sufficient pairs of exposure-outcome, the use of different and non-convertible effect sizes prevented us to do a meta-analysis. The small number of studies entered in the meta-analyses prevented us to do more in-depth analyses such as assessment of publication bias, and different sub-group analyses and meta-regression to explore the source of heterogeneity. The quality assessment instrument used in this study was not specifically designed for observational studies in environmental epidemiology in which the magnitude of the effect size is smaller than the common effect estimates in clinical epidemiology. Additionally, for cross-sectional studies [6, 10, 41, 42] we used the modified NOS to fit the evaluation of this type of study design.

The outcome definition in some cases of cancer was different across the studies (*e.g.* for lung cancer or colorectal cancer). Linearity of the associations were reported in a small number of the retrieved studies, and the pattern of linearity checking, outcome definition, and exposure definition was not clear in some of the studies. Only one study used the DAG approach for defining variables for adjustment, and the adjustment procedures and variables in the included studies need more elaboration. Adjusting for potential mediators (*e.g.* physical activity), or descending proxy for these mediators (*e.g.* body mass index as a descending proxy of physical activity) can induce over-adjustment bias in the estimation of the total effect [69]. Regarding the distance to the different greenspace, none of the studies considered network distance to the greenspaces. Using conventional Euclidean distance instead of greenspace network distance could overestimate greenspace accessibility even by a factor of three (by representing the distance to access the greenspace closer than the actual distance) [70].

In general, the body of evidence on the association between exposure to greenspace and cancer is accumulating; however, the results are quite heterogeneous. Nearly none of the studies were conducted in warm and arid climates. Most of the available evidence are from cross-sectional studies, which are limited to considering the temporality in exposure-outcome, and future longitudinal studies are warranted. Associations are reported based on the different cancer sites and also types of greenspaces. Heterogeneity in the exposure definition (in terms of the type of greenspace, buffer sizes, and type of greenspace indicator), outcome definition and assessment (for example using different ICD codes for defining skin or lung cancer), a different approach in statistical analysis and adjustment (inclusion of mediators in the models or cases of over adjustment) are among the most common pitfalls in the included studies in this review.

Conclusion and future directions

Part of the available literature is suggestive of a possible protective association of greenspace exposure with breast and prostate cancer, while for other cancers the evidence is still very limited and heterogeneous. However, considering controversial findings in the available studies in addition to the null pooled estimates, and overall, very low quality of the evidence it is not possible to draw a conclusion from the literature at this point. Generally, the overall quality of evidence for the association between greenspace and cancer is still very low, highlighting the need for further studies. Future studies should consider more precise exposure assessment and definition by type of greenspace, considering linearity and adjustment for a sufficient set of variables.

Registration and protocol

The systematic review protocol has been registered in the PROSPERO (registration ID: CRD42021277074).

.

Support

The supporting bodies in this study had no role in hypothesis generation, data collection, analyses, interpretation and writing of the manuscript.

Competing interests

The authors declare that they have no conflict of interest.

Availability of data, code and other materials

Data collection forms, and the data extracted from included studies are available upon the request from the corresponding author.

Human and Animal Rights and Informed Consent

This article does not contain any studies with human or animal subjects.

Authors Contribution:

Mohammad Javad Zare Sakhvidi: Conceptualization; Methodology; Data curation; Software;

Formal analysis; Visualization; Writing – original draft; Writing – review & editing

Jun Yang: Conceptualization; Methodology; Data curation; Writing –review & editing

Amir Houshang Mehrparvar: Data curation; Formal analysis; Writing– review & editing

Angel M. Dzhambov: Methodology; Writing –review & editing

AliAsghar Ebrahimi: Writing –review & editing

Payam Dadvand: Methodology; Writing –review & editing

Bénédicte Jacquemin: Conceptualization; Methodology; Data curation; Writing – review & editing

References

1. Du, Y., et al., *How does urbanization influence PM_{2.5} concentrations? Perspective of spillover effect of multi-dimensional urbanization impact*. *Journal of Cleaner Production*, 2019. **220**: p. 974-983.
2. Han, L., et al., *Impact of urbanization level on urban air quality: A case of fine particles (PM_{2.5}) in Chinese cities*. *Environmental Pollution*, 2014. **194**: p. 163-170.
3. Agency, U.S.E.P. *What is Open Space/Green Space?* 2021 [cited 2021 19 December]; Available from: <https://www3.epa.gov/region1/eco/uep/openspace.html>.
4. Namin, S., et al., *Neighborhood Characteristics and Cancer Survivorship: An Overview of the Current Literature on Neighborhood Landscapes and Cancer Care*. *International journal of environmental research and public health*, 2021. **18**(13): p. 7192.
5. Porcherie, M., et al., *Relationship between Urban Green Spaces and Cancer: A Scoping Review*. *International journal of environmental research and public health*, 2021. **18**(4): p. 1751.
6. Demoury, C., et al., *Residential greenness and risk of prostate cancer: A case-control study in Montreal, Canada*. *Environment international*, 2017. **98**: p. 129-136.
7. James, P., et al., *Exposure to Greenness and Mortality in a Nationwide Prospective Cohort Study of Women*. *Environmental Health Perspectives*, 2016. **124**(9): p. 1344-1352.
8. Lee, H.-Y., et al., *Association between Surrounding Greenness and Mortality: An Ecological Study in Taiwan*. *International Journal of Environmental Research and Public Health*, 2020. **17**(12): p. 4525.

9. Astell-Burt, T., X. Feng, and G.S. Kolt, *Neighbourhood green space and the odds of having skin cancer: multilevel evidence of survey data from 267072 Australians*. *Journal of Epidemiology and Community Health*, 2014. **68**(4): p. 370-374.
10. Carles, C., et al., *Residential proximity to agricultural land and risk of brain tumor in the general population*. *Environmental research*, 2017. **159**: p. 321-330.
11. Kim, S., H. Kim, and J.-T. Lee, *Interactions between Ambient Air Particles and Greenness on Cause-specific Mortality in Seven Korean Metropolitan Cities, 2008–2016*. *International Journal of Environmental Research and Public Health*, 2019. **16**(10): p. 1866.
12. Klompaker, J.O., et al., *Surrounding green, air pollution, traffic noise exposure and non-accidental and cause-specific mortality*. *Environment International*, 2020. **134**: p. 105341.
13. Roe, J.J., et al., *Green space and stress: evidence from cortisol measures in deprived urban communities*. *International journal of environmental research and public health*, 2013. **10**(9): p. 4086-4103.
14. Baldauf, R., *Air pollution mitigation through vegetation barriers and green space*, in *Traffic-Related Air Pollution*. 2020, Elsevier. p. 437-453.
15. Margaritis, E. and J. Kang, *Relationship between green space-related morphology and noise pollution*. *Ecological indicators*, 2017. **72**: p. 921-933.
16. Dzhambov, A.M., et al., *Analytical approaches to testing pathways linking greenspace to health: A scoping review of the empirical literature*. *Environmental Research*, 2020. **186**: p. 109613.
17. Markevych, I., et al., *Exploring pathways linking greenspace to health: Theoretical and methodological guidance*. *Environmental research*, 2017. **158**: p. 301-317.

18. Zare Sakhvidi, M.J., et al., *Greenspace exposure and cancer incidence: A 27-year follow-up of the French GAZEL cohort*. *Science of The Total Environment*, 2021. **787**: p. 147553.
19. Page, M.J., et al., *The PRISMA 2020 statement: an updated guideline for reporting systematic reviews*. *Bmj*, 2021. **372**.
20. Ouzzani, M., et al., *Rayyan—a web and mobile app for systematic reviews*. *Systematic reviews*, 2016. **5**(1): p. 1-10.
21. Drevon, D., S.R. Fursa, and A.L. Malcolm, *Intercoder reliability and validity of WebPlotDigitizer in extracting graphed data*. *Behavior modification*, 2017. **41**(2): p. 323-339.
22. Stang, A., *Critical evaluation of the Newcastle-Ottawa scale for the assessment of the quality of nonrandomized studies in meta-analyses*. *European journal of epidemiology*, 2010. **25**(9): p. 603-605.
23. Modesti, P.A., et al., *Panethnic differences in blood pressure in Europe: a systematic review and meta-analysis*. *PloS one*, 2016. **11**(1): p. e0147601.
24. Zare Sakhvidi, M.J., et al., *Air pollution exposure and bladder, kidney and urinary tract cancer risk: A systematic review*. *Environmental Pollution*, 2020. **267**: p. 115328.
25. DerSimonian, R. and R. Kacker, *Random-effects model for meta-analysis of clinical trials: an update*. *Contemporary clinical trials*, 2007. **28**(2): p. 105-114.
26. Higgins, J.P.T., et al., *Cochrane handbook for systematic reviews of interventions*. 2019: John Wiley & Sons.
27. Coory, M.D., *Comment on: Heterogeneity in meta-analysis should be expected and appropriately quantified*. *International journal of epidemiology*, 2010. **39**(3): p. 932-932.
28. Viechtbauer, W., *Conducting Meta-Analyses in R with the metafor Package*. *Journal of Statistical Software*, 2010. **36**(03): p. 1-48.

29. Dijkers, M., *Introducing GRADE: a systematic approach to rating evidence in systematic reviews and to guideline development*. *KT Update*, 2013. **1**(5): p. 1-9.
30. Guyatt, G.H., et al., *GRADE guidelines: 9. Rating up the quality of evidence*. *Journal of clinical epidemiology*, 2011. **64**(12): p. 1311-1316.
31. Guyatt, G.H., et al., *GRADE guidelines: 4. Rating the quality of evidence—study limitations (risk of bias)*. *Journal of clinical epidemiology*, 2011. **64**(4): p. 407-415.
32. Guyatt, G.H., et al., *GRADE guidelines: 7. Rating the quality of evidence—inconsistency*. *Journal of clinical epidemiology*, 2011. **64**(12): p. 1294-1302.
33. Guyatt, G.H., et al., *GRADE guidelines 6. Rating the quality of evidence—imprecision*. *Journal of clinical epidemiology*, 2011. **64**(12): p. 1283-1293.
34. Guyatt, G.H., et al., *GRADE guidelines: 8. Rating the quality of evidence—indirectness*. *Journal of clinical epidemiology*, 2011. **64**(12): p. 1303-1310.
35. Guyatt, G.H., et al., *GRADE guidelines: 5. Rating the quality of evidence—publication bias*. *Journal of clinical epidemiology*, 2011. **64**(12): p. 1277-1282.
36. Keegan, T.H.M., et al., *Neighborhood influences on recreational physical activity and survival after breast cancer*. *Cancer Causes & Control*, 2014. **25**(10): p. 1295-1308.
37. Mears, M., et al., *Greenspace spatial characteristics and human health in an urban environment: An epidemiological study using landscape metrics in Sheffield, UK*. *Ecological Indicators*, 2019. **106**: p. 105464.
38. Harrigan, S.P., *Built environment and post-menopausal breast cancer risk : analysis of a linked British Columbian cohort*. 2021.
39. Datzmann, T., et al., *Outdoor air pollution, green space, and cancer incidence in Saxony: a semi-individual cohort study*. *BMC Public Health*, 2018. **18**(1): p. 715.

40. Iyer, H.S., et al., *The association between neighborhood greenness and incidence of lethal prostate cancer: A prospective cohort study*. Environmental Epidemiology, 2020. **4**(2): p. e091.
41. O'Callaghan-Gordo, C., et al., *Residential proximity to green spaces and breast cancer risk: The multicase-control study in Spain (MCC-Spain)*. International Journal of Hygiene and Environmental Health, 2018. **221**(8): p. 1097-1106.
42. Shao, Y., et al., *Geographical variation in lung cancer risk associated with road traffics in Jiading District, Shanghai*. Science of The Total Environment, 2019. **652**: p. 729-735.
43. Bixby, H., et al., *Associations between Green Space and Health in English Cities: An Ecological, Cross-Sectional Study*. PLOS ONE, 2015. **10**(3): p. e0119495.
44. Mitchell, R. and F. Popham, *Effect of exposure to natural environment on health inequalities: an observational population study*. The Lancet, 2008. **372**(9650): p. 1655-1660.
45. Richardson, E., et al., *The association between green space and cause-specific mortality in urban New Zealand: an ecological analysis of green space utility*. BMC Public Health, 2010. **10**(1): p. 240.
46. Richardson, E.A. and R. Mitchell, *Gender differences in relationships between urban green space and health in the United Kingdom*. Social Science & Medicine, 2010. **71**(3): p. 568-575.
47. Wu, K.-S., X. Huo, and G.-H. Zhu, *Relationships between esophageal cancer and spatial environment factors by using Geographic Information System*. Science of The Total Environment, 2008. **393**(2-3): p. 219-225.
48. Rothman, K.J. and S. Greenland, *Causation and causal inference in epidemiology*. American journal of public health, 2005. **95**(S1): p. S144-S150.

49. Weed, D.L. and L.S. Gorelic, *The practice of causal inference in cancer epidemiology*. Cancer Epidemiology and Prevention Biomarkers, 1996. **5**(4): p. 303-311.
50. Diener, A. and P. Mudu, *How can vegetation protect us from air pollution? A critical review on green spaces' mitigation abilities for air-borne particles from a public health perspective-with implications for urban planning*. Science of The Total Environment, 2021. **796**: p. 148605.
51. Vieira, J., et al., *Green spaces are not all the same for the provision of air purification and climate regulation services: The case of urban parks*. Environmental Research, 2018. **160**: p. 306-313.
52. Hunter R, Cleary A, Cleland C, Braubach M. Urban green space interventions and health: A review of impacts and effectiveness. Full report. 2017
53. Knobel, P., P. Dadvand, and R. Maneja-Zaragoza, *A systematic review of multi-dimensional quality assessment tools for urban green spaces*. Health & place, 2019. **59**: p. 102198.
54. Knobel, P., et al., *Development of the urban green space quality assessment tool (RECITAL)*. Urban Forestry & Urban Greening, 2021. **57**: p. 126895.
55. Richardson, E.A., et al., *Role of physical activity in the relationship between urban green space and health*. Public health, 2013. **127**(4): p. 318-324.
56. Schipperijn, J., et al., *Associations between physical activity and characteristics of urban green space*. Urban forestry & urban greening, 2013. **12**(1): p. 109-116.
57. Jennings, V. and O. Bamkole, *The relationship between social cohesion and urban green space: An avenue for health promotion*. International journal of environmental research and public health, 2019. **16**(3): p. 452.

58. Hedblom, M., et al., *Reduction of physiological stress by urban green space in a multisensory virtual experiment*. Scientific reports, 2019. **9**(1): p. 1-11.
59. Ibes, D., I. Hirama, and C. Schuyler, *Greenspace ecotherapy interventions: The stress-reduction potential of green micro-breaks integrating nature connection and mind-body skills*. Ecopsychology, 2018. **10**(3): p. 137-150.
60. Speck, R.M., et al., *An update of controlled physical activity trials in cancer survivors: a systematic review and meta-analysis*. Journal of Cancer Survivorship, 2010. **4**(2): p. 87-100.
61. Dzhambov, A.M., et al., *Analytical approaches to testing pathways linking greenspace to health: A scoping review of the empirical literature*. Environmental Research, 2020. **186**: p. 109613.
62. Klompaker, J.O., et al., *Green space definition affects associations of green space with overweight and physical activity*. Environmental research, 2018. **160**: p. 531-540.
63. Doick, K.J., A. Peace, and T.R. Hutchings, *The role of one large greenspace in mitigating London's nocturnal urban heat island*. Science of the total environment, 2014. **493**: p. 662-671.
64. Dzhambov, A.M., et al., *Lower noise annoyance associated with GIS-derived greenspace: pathways through perceived greenspace and residential noise*. International journal of environmental research and public health, 2018. **15**(7): p. 1533.
65. Mueller, W., et al., *Urban greenspace and the indoor environment: Pathways to health via indoor particulate matter, noise, and road noise annoyance*. Environmental research, 2020. **180**: p. 108850.

66. Liu, H.-L. and Y.-S. Shen, *The impact of green space changes on air pollution and microclimates: A case study of the Taipei metropolitan area*. Sustainability, 2014. **6**(12): p. 8827-8855.
67. Mills, J.G., et al., *Revegetation of urban green space rewilds soil microbiotas with implications for human health and urban design*. Restoration Ecology, 2020. **28**: p. S322-S334.
68. Provost, D., et al., *Brain tumours and exposure to pesticides: a case–control study in southwestern France*. Occupational and environmental medicine, 2007. **64**(8): p. 509-514.
69. Schisterman, E.F., S.R. Cole, and R.W. Platt, *Overadjustment bias and unnecessary adjustment in epidemiologic studies*. Epidemiology (Cambridge, Mass.), 2009. **20**(4): p. 488.
70. Moseley, D., et al., *Green networks for people: Application of a functional approach to support the planning and management of greenspace*. Landscape and urban planning, 2013. **116**: p. 1-12.

Figures captions:

Figure 1. PRISMA chart for the study selection process.

Figure 2. Meta-analyses for the studies reported the associations between greenspace exposure (as 0.1 increase in NDVI) and lung cancer (N represents the number of cancer cases in each study).

Figure3. Meta-analyses for the studies reported the associations between greenspace exposure (as 0.1 increase in NDVI) and prostate cancer incidence/prevalence (N represents the number of cancer cases in each study).

Figure 4. Meta-analyses for the studies reported the associations between greenspace exposure (as 0.1 increase in NDVI) and breast cancer incidence/prevalence (N represents the number of cancer cases in each study).

Supplementary method

Supplementary method:

Transformation of the effect sizes for 0.1 unit increase in NDVI

To be able to compare the associate estimates across the studies through meta-analysis, we converted all the reported associate estimates and confidence intervals for the 0.1 unit increase in surrounding greenness indices or 1 km proximity for proximity indicators according to the method described in other papers (**Equation 1 and 2**) (24, 25).

$$\beta_i = \ln(RR) \times \left(\frac{0.1}{\Delta \text{ greenspace}(A)} \right) \quad \text{Equation 1}$$

$$\sigma_i = \left(\frac{(\ln RR_{up}) - (\ln RR_{low})}{3.92} \right) \times \left(\frac{0.1}{\Delta \text{ greenspace}(A)} \right) \quad \text{Equation 2}$$

Where β_i is the natural logarithm of the effect size (odds ratio (OR) or relative risk (RR)). The σ_i denotes the corresponding standard error. $\ln RR_{up}$ and $\ln RR_{low}$ are natural logarithms of upper and lower levels of the 95% confidence interval of the effect size, respectively. The “ Δ greenspace (A)” refers to the unit of increase used for reporting the effect size in the papers.

Tables

Table 1. Summary of the characteristics for the studies included in the review by alphabetical order of the first author's name.

First author (Published year); Study location	Design (Cancer cases/ Total population or control); follow-up length	Exposure ^a	Outcome	Main findings
Astell-Burt <i>et al.</i> (2014); Australia	Cross-sectional (76096/ 267000)	Land cover	Skin cancer prevalence	Higher exposure to greenspace was associated with higher skin cancer prevalence.
Bixby <i>et al.</i> (2015); UK	Cross-sectional (12005/ 11000000); urban residents	Land cover	Lung cancer mortality	Lower risk of lung cancer mortality in 5 th and 2 nd quantile of exposure to greenspace (based on land cover) compared to other quantiles. No trend was observed. However, the effect estimates reduced to protective when the models were adjusted for further variables.
Carles <i>et al.</i> (2017); France	Case-control (490/ 980)	Proximity; Land cover	Brain tumor prevalence (meningioma and glioma)	Residential proximity to agricultural lands was suggestive of an increased risk of meningioma, however, it was only significant for the exposure near open field crops. A nonsignificant association was found for the proximity of vineyards and orchards above the 95 th percentile.
Datzman <i>et al.</i> (2018); Germany	Semi-individual Cohort (64341/1918449)	Greenness	Colorectal cancer incidence; Non-melanoma skin cancer incidence; Breast cancer incidence; Mouth and throat cancer incidence; Prostate cancer incidence	Higher exposure to greenspace was significantly associated with a decrease risk of breast cancer, non-melanoma skin cancer (NMSC), and mouth and throat cancers. Subgroup analysis based on sex showed an increased risk of colorectal, mouth and throat, and NMSC in men. Subgroup analysis based on alcohol consumption showed an increased risk factor for colorectal, mouth and throat, and breast cancer in alcohol drinkers. The associations for colorectal and prostate cancer were null.
Demoury <i>et al.</i> (2017); Canada	Case-control (1933/ 1994)	Greenness	Prostate cancer prevalence	A significant protective association was found for NDVI at index date. However, for NDVI 10 years before cancer prevalence, only the association for 5000 m buffer was significantly protective.
Harrigan (2017); Canada	Cohort (122/ 7330); 7 years follow-up	Greenness	Breast cancer incidence	An increase in greenness was associated with a non-significant reduction of breast cancer risk.
Iyer <i>et al.</i> (2020); USA	Cohort (898/ 47958); 28 years follow-up	Greenness	Lethal prostate cancer incidence	Higher NDVI was associated with lower risk in all study population and non-movers respectively. Lower risk was found for high-population density areas whereas an increase in risk for low-density areas.
James <i>et al.</i> (2016); USA	Cohort (3363/ 108630); 8 years follow-up	Greenness	All cancer mortality	Increase in NDVI was associated with a decreased risk of all cancer mortality.

First author (Published year); Study location	Design (Cancer cases/ Total population or control); follow-up length	Exposure ^a	Outcome	Main findings
Kim <i>et al.</i> (2019); South Korea	Time-series analysis (73 districts/ covering 23.2 million populations)	Greenness	Lung cancer mortality	There was no association between greenness and lung cancer mortality.
Klompemaker <i>et al.</i> (2020); the Netherlands	Cohort (2195/ 244814); 4 years follow-up	Greenness; Land cover	Lung cancer mortality	Surrounding greenspace was not associated with lung cancer mortality.
Lee <i>et al.</i> (2020); Taiwan	Cross-sectional (Unclear)	Greenness; Land cover	Lung cancer mortality	Higher levels of NDVI, or proximity to area covered with forests were associated with lower lung cancer mortality.
Mitchel <i>et al.</i> (2018); UK	Cross-sectional (25472/ 8345088)	Percentage of land cover	Lung cancer mortality	A weak and non-significant protective association was observed between exposure to greenspace and lung cancer mortality. Removing rural participants did not change the findings.
O'Callaghan-Gordo <i>et al.</i> (2018); Spain	Case-control (1738/ 1910)	Greenness; Proximity	Breast cancer prevalence	The residential proximity to urban greenspace was associated with a decreased risk of breast cancer. However, the proximity to agricultural lands or higher greenness was associated with an increased risk of breast cancer.
Richardson <i>et al.</i> (2010); New Zealand	Cross-sectional (2603/ 1546405)	Land cover	Lung cancer mortality	An increase in total greenspace was significantly associated with an increased risk of lung cancer mortality. For useable greenspace, no association was found.
Richardson <i>et al.</i> (2010); UK	Cross-sectional (on 6432 wards covering 28600000)	Land cover	Lung cancer mortality	For men, there was no association between greenspace and lung cancer mortality. But in women, intermediate greenspace [percentile 50 and 75] was associated with significantly higher lung cancer mortality compared to below 25% greenspace. For the higher than 75% greenspace, the non-significant higher risk was observed for women.
Shao <i>et al.</i> (2019); China	Case-control (1461/ 954 colorectal cancer as control)	Greenness	Lung cancer prevalence	No significant increase in the risk of lung cancer was observed with an increase in greenness.
Wu <i>et al.</i> (2008); China	Cross-sectional (area-level mortality rate for 273 counties or cities)	Greenness	Esophageal cancer mortality	NDVI in the high-risk area was lower than in the low-risk area. Positive but non-significant correlation between NDVI in January and esophagus cancer mortality, but a negative and nonsignificant association for NDVI at July and esophagus cancer.
Zare Sakhvidi <i>et al.</i> (2020); France	Cohort (4075/ 19408); 27 years follow-up	Greenness; Proximity	All cancer incidence; Prostate cancer; Skin cancer; Colorectal cancer; Breast cancer; Lung	Greenness was linked to an increased risk of all-site cancers Greenspace seemed to have a protective role in breast cancer. No association was found for the other types of cancer.

First author (Published year); Study location	Design (Cancer cases/ Total population or control); follow-up length	Exposure ^a	Outcome	Main findings
cancer				

a) Details on exposure assessment, allocation and characteristics are presented in Table 4.

Table 2. Design specific quality assessment of the included studies (by alphabetical order of the first author's name) based on the Newcastle Ottawa Scale

Cohort studies	Selection				Comparability	Outcome			Total
First author (year)	Representativeness	Non-exposed selection	Ascertainment of exposure	No outcome at the start	Comparability	Outcome assessment	Length of follow-up	Adequacy of follow up	
Datzman <i>et al.</i> (2018)	1	1	0	1	0	1	0	0	4
Harrigan (2017)	1	1	1	1	0	1	1	1	7
Iyer <i>et al.</i> (2020)	0	1	1	1	0	1	1	1	8
James <i>et al.</i> (2016)	0	1	1	1	2	1	0	1	7
Klompaker <i>et al.</i> (2020)	1	1	1	1	0	1	0	0	5
Zare Sakhvidi <i>et al.</i> (2020)	1	1	1	1	0	1	1	1	7
Percent	66.67	100.00	83.33	100.00	17%	100.00	50.00	66.67	5.4

Case-control studies	Selection				Comparability	Exposure			Total
Firstauthor (year)	Representativeness	Case definition	Control selection	Control definition	Comparability	Exposure assessment	Exposure ascertainment	Non-Response rate	
Carles <i>et al.</i> (2017)	1	1	1	1	0	1	1	0	6
Demoury <i>et al.</i> (2017)	1	1	1	1	0	1	1	0	6
OCallaghan-Gordo <i>et al.</i> (2018)	1	1	1	1	0	1	1	0	6
Shao <i>et al.</i> (2019)	1	1	0	0	0	1	1	0	4
Percent	100.00	100.00	75.00	75.00	0.00	100.00	100.00	0.00	5.5

Cross-sectional studies	Selection				Comparability	Outcome		--	Total
First author (year)	Representativeness	Sample size	Non-respondent	Ascertainment of exposure	Comparability	Outcome assessment	Statistical tests	---	
Astell-Burt <i>et al.</i> (2014)	1	1	0	1	0	1	1		5
Bixby <i>et al.</i> (2015)	0	0	0	0	0	1	1		2
Kim <i>et al.</i> (2019)	1	0	0	0	0	1	1		3
Lee <i>et al.</i> (2020)	1	0	0	0	0	1	0		2
Mitchel <i>et al.</i> (2008)	1	1	0	0	0	2	1		5
Richardson <i>et al.</i> (2010) [NZ]	1	1	0	0	0	1	1		4
Richardson <i>et al.</i> (2010) [UK]	0	0	0	0	0	2	1		3
Wu <i>et al.</i> (2008)	1	0	0	0	0	0	0		1
Percent	75.00	37.50	0.00	12.50	0.00%	56.25	75.00		3.1

Table 3. Summary of the statistical approaches used in the studies entered into the review.

Author (year)	Statistical model (effect size)	Linearity test	Mediation analysis	Effect modification/ stratified analysis/ interaction	Sensitivity analysis	Adjustment for air pollution
Astell-Burt <i>et al.</i> (2014)	Multilevel logistic regression (OR)	No	Yes (time spent outdoor)	Yes (Socioeconomic status)	Yes (melanoma and non-melanoma skin cancer)	No
Bixby <i>et al.</i> (2015)	Multilevel- Poisson regression (RR*)	Yes (unclear procedure)	No	No	No	Yes (PM ₁₀)
Carles <i>et al.</i> (2017)	Conditional logistic regression (OR)	No	No	Yes (tumor morphology, type of land cover)	Yes (excluding those with missing in geocoding, different lag periods (2 or 5 years); complete data set analysis; exposure to other types of lands; analysis based on 1000 m buffer)	No
Datzman <i>et al.</i> (2018)	Multilevel-Poisson model (RR*)	No	No	Yes (sex, alcohol-related disorders)	Yes (change of residence address)	No
Demoury <i>et al.</i> (2017)	Unconditional logistic regression models (OR)	Yes (check the scatterplot and doing likelihood ratio test)	Yes (air pollution (NO ₂), physical activity, BMI)	Yes (social and material deprivation)	Yes (exclusion of proxy respondents)	No
Harrigan (2017)	Cox proportional-hazards model (HR)	Yes (p-trend for quantiles).	No	No	No	No
Iyer <i>et al.</i> (2020)	Cox proportional-hazards model (HR)	Yes (p-trend for quantiles; spline)	Yes (physical activity)	Yes (population density; address type; country region; prostate-specific antigen screening; prostate-specific antigen intensity)	Yes (change of residence address)	No
James <i>et al.</i> (2016)	Extended cox regression model (HR)	Yes (p-trend for quintiles; cubic spline, likelihood ratio test)	Yes (air pollution (PM _{2.5}), physical activity, social engagement, mental health)	Yes (Race, smoking, area-level income, home value, overweight, physical activity, air pollution, region, urbanity, change in address)	No	No

Author (year)	Statistical model (effect size)	Linearity test	Mediation analysis	Effect modification/ stratified analysis/ interaction	Sensitivity analysis	Adjustment for air pollution
Kim <i>et al.</i> (2019)	Random effect binominal regression model (% increase)	No	No	Yes (Air pollution as interaction term)	Yes (using a different definition for area-level socioeconomic status, adding gaseous pollutants in the models)	No
Klompmaker <i>et al.</i> (2020)	Cox proportional-hazards model (HR)	Yes (spline, likelihood ratio test)	No	Yes (age; multipollutant interaction)	Yes (change of residence address; subjects younger than 85, and subject younger than 75 at baseline)	No
Lee <i>et al.</i> (2020)	Generalized additive mixed models (GAMM) with Poisson regression (RR)	Yes (spline)	No	Yes (sex ratio, age, annual household income, urban/rural areas)	No	Yes (PM _{2.5} , NOX)
Mitchel <i>et al.</i> (2018)	Negative binominal regression models with robust estimation of variance (IRR)	No	No	Yes (deprivation)	Yes (urbanity)	No
OCallaghan-Gordo <i>et al.</i> (2018)	Mixed effect logistic regression model (OR)	No	Yes (air pollution, physical activity)	Yes (urban/rural areas, familial history of cancer, menopausal status, socioeconomic status)	Yes (change of residence address; urban area participants)	No
Richardson <i>et al.</i> (2010) [UK]	Negative Binominal Poisson regression with robust standard errors (IRR)	No	No	Yes (sex)	No	Yes
Richardson <i>et al.</i> (2010) [NZ]	Negative binominal regression model (IRR)	No	No	Yes (sex, age group, area deprivation, smoking rate, air pollution)	No	Yes (area level air pollution)
Shao <i>et al.</i> (2019)	Multivariate logistic regression model (OR)	No	No	No	No	No
Wu <i>et al.</i> (2008)	Correlation test; Principal component regression (correlation coefficient)	No	No	Yes (sex)	No	No
Zare Sakhvidi <i>et al.</i> (2020)	Extended cox regression model (HR)	No	No	Yes (urban/rural areas)	Yes (occupational exposure to carcinogens; change of residence address; lag time between exposure and outcome, different buffers)	Yes (PM _{2.5})

PM₁₀: Particulate matters with an aerodynamic diameter less than 10 micrometers; PM_{2.5}: Particulate matters with an aerodynamic diameter less than 2.5 micrometers; NOX: Nitrogen oxides; RR*: Relative risk; RR: Risk ratio; IRR: Incidence rate ratios; HR: Hazard ratio

Table 4. Summary of the exposure assessment in the studies entered into the review.

The first author (year)	Exposure Location	Exposure indicator(s)	Method of data acquisition (source name; resolution or classification accuracy)	Buffer size	Timepoint in year used for data acquisition	Temporality between exposure assessment and outcome occurrence	Exclusion of specific data/ lands	Summary statistics of exposure intensity (mean and SD for NDVI; percent for land cover)
Astell-Burt <i>et al.</i> (2014)	Home	Land cover (% of green space)	Land cover (Australian Meshblock classification; unclear)	1000m	All over the year	Single time point before outcome occurrence (for most of the participants)	Private gardens and agricultural lands	Land cover= for most of the participants less than 20%
Bixby <i>et al.</i> (2015)	Residential city level	Land cover (% of city covered by green space)	Land cover (Land cover map 2007; 20-30 m)	City area	All over the year	Mixed (for some participants before, and some other after the outcome occurrence), a single time point	Areas less than 0.5 hectare	Land cover= 32% [17%-61%]
Carles <i>et al.</i> (2017)	Home (some cases at municipality level)	Land cover (% of green space); Proximity (Inverse distance)	Land cover (CORINE land cover; 25 hectares)	500m	All over the year	Before the outcome until the outcome occurrence	Unclear	Unclear
Datzman <i>et al.</i> (2018)	Home (at postal code level)	Greenness (NDVI)	Remote sensing (MODIS, unclear satellite; 250m)	250m	All over the year	Before the outcome occurrence (for most of the participants)	Unclear	NDVI 250m= 0.51 [range:0.38-0.64]
Demoury <i>et al.</i> (2017)	Home	Greenness (NDVI)	Remote sensing (Landsat; 30m)	150m; 300 m; 500m; 1000m	August for 1996, and June for 2005	Time of the outcome occurrence; A decade before the outcome occurrence	Unclear	NDVI 300m= 0.33(0.09) at time of diagnosis; 0.42(0.12) at 10 years before diagnosis
Harrigan (2017)	Participants Residential address	Greenness (NDVI)	Remote sensing (Landsat; 30m)	Unclear	Unclear	Averaged over follow up time; also at baseline	Unclear	0.3 (0.1)
Iyer <i>et al.</i> (2020)	Home; Work	Greenness (NDVI)	Remote sensing (AVHRR on NOAA; 1000m)	1000m	Average of different seasons; Maximum NDVI.	Baseline; Cumulative during follow-up	Negative values	Baseline NDVI: 0.28(0.09); Cumulative NDVI: 0.31 (0.09)
James <i>et al.</i> (2016)	Home	Greenness (NDVI)	Remote sensing (MODIS on Terra, 250m)	250m; 1250m	One month as a representative of each season	Time of the outcome occurrence; All-time during lifetime	Unclear	NDVI 250m= 0.47(0.12)

The first author (year)	Exposure Location	Exposure indicator(s)	Method of data acquisition (source name; resolution or classification accuracy)	Buffer size	Timepoint in year used for data acquisition	Temporality between exposure assessment and outcome occurrence	Exclusion of specific data/ lands	Summary statistics of exposure intensity (mean and SD for NDVI; percent for land cover)
Kim <i>et al.</i> (2019)	District level	Greenness (NDVI)	Remote sensing (MODIS on Terra; 250m)	City area	Median greenness during summer (May-October)	Time of outcome occurrence	Negative values	NDVI 250m= 0.48 (0.13)
Klompmaker <i>et al.</i> (2020)	Home	Greenness (NDVI); Land cover (% of green space)	Remote sensing (Landsat; 30m); Land cover (land-use database of the Netherlands; unclear)	300m; 500m	Summer	3-7 years before outcome occurrence	Negative values	NDVI 300m= 0.52(0.13)** Land cover at 300m= 19% (24)**
Lee <i>et al.</i> (2020)	Town level	Greenness (NDVI); Land cover (% of green space)	Remote sensing (MODIS on Terra, 250m); Land cover (Taiwan Land-use Investigation; unclear)	250m (unclear for land cover)	Unclear	Greenness at the time of outcome occurrence; Land cover (after outcome occurrence, and at 2014)	Unclear	NDVI 250m= 0.51 (0.17); Forest area= 30.1% (30.8); Park area =1.17% (2.99)
Mitchel <i>et al.</i> (2018)	Residential city level (LSOA)	Land cover (% of green space)	Land cover (Generalized land use database in the UK; 10 m ²)	Unclear	All over the year	Before the outcome occurrence (for most of the participants)	Unclear	Unclear
O'Callaghan-Gordo <i>et al.</i> (2018)	Home	Greenness (NDVI); Land cover (Presence or absence of urban and agricultural green space)	Remote sensing (Landsat; 30m); Land cover (Urban atlas 0.25 hectares; CORINE land cover; 25 hectares)	300m; nested buffers in 100m; 300m; and 500m	Spring (maximum vegetation of year)	Before the outcome occurrence (nearly a decade)	Negative values	NDVI 300m=0.22(0.14) **; Presence of urban green area: 80.5%; Presence of agricultural lands: 24.1%
Richardson <i>et al.</i> (2010) [New Zealand]	Residential city level (CAU)	Land cover (% total; % useable, and %non-useable green space)	The Land Cover Data Base (LCDB2; unclear); Department of Conservation (DOC; unclear) database; Land Information New Zealand's (LINZ; unclear)	City area	All over the year	After the occurrence of the outcome (for most of the participants)	Aquatic areas	Total green space land cover=42%; Useable green space land cover=17%

The first author (year)	Exposure Location	Exposure indicator(s)	Method of data acquisition (source name; resolution or classification accuracy)	Buffer size	Timepoint in year used for data acquisition	Temporality between exposure assessment and outcome occurrence	Exclusion of specific data/ lands	Summary statistics of exposure intensity (mean and SD for NDVI; percent for land cover)
Richardson <i>et al.</i> (2010) [UK]	Ward level	Land cover (% of green spaces)	Land cover (GLUD; 5 m ² , CORINE; 25 hectares)	Ward area	All over the year	Unclear	Domestic gardens	Unclear
Shao <i>et al.</i> (2019)	Participants residential address	Greenness (NDVI)	Remote sensing (MODIS on Terra; 250m)	Unclear	Averaged NDVI values in September (high vegetation coverage) and January (low vegetation coverage)	Single time point; 3-5 years before the incidence of the outcome	Unclear	Unclear
Wu <i>et al.</i> (2008)	Residential city level	Greenness (NDVI)	Remote sensing (AVHRR on NOAA, 1000 m)	City area	January and July separately	Single time point; before the occurrence of the outcome (for most of the participants)	Unclear	Unclear
Zare Sakhvidi <i>et al.</i> (2020)	Participants Residential address	Greenness (NDVI); Proximity (Inverse distance)	Remote sensing (Landsat; 30m); Land cover (CORINE land cover; 25 hectares)	100m; 300m; 500m; 1000m	Averaged NDVI in the greenness time of the years (may-July)	Averaged over follow up time with and without a 10-year delay	Negative values	NDVI 300m=0.49 (0.16); Proximity to the forest: 1.62 km (1.99); Proximity to agricultural lands: 0.90 km (1.53), proximity to urban greenspace: 4.09 km (5.00)

Abbreviations:

NOAA: National Oceanic and Atmospheric Administration

AVHRR: The Advanced Very High-Resolution Radiometer

MODIS: Moderate Resolution Imaging Spectroradiometer

GLUD: Greenspace category comprises parks, other open spaces, and agriculture lands but not domestic gardens

LSOA: Geographical unit used for reporting small area statistics in the UK

*: The inverse distance and area ratios (percentage of crop surface) were then multiplied by residential duration at the address, and the addresses' scores were summed for each participant to create a lifelong cumulated exposure score for each crop. The resulting scores were named CLC and UAA scores respectively. Values of NDIV are mean and standard deviation, except than specified otherwise.

** : median and IQR in cases

Table S1. Inclusion and exclusion criteria based on PECOS (population, exposure, comparison, outcome, and study type) for a systematic review on the association between greenspace and cancers.

Decision	Population	Exposure	Comparison	Outcome	Type of study
Inclusion	Human; Adult	Greenspace including a) Greenspace availability b) Greenspace accessibility c) Greenspace use	Not applicable	All-site and site-specific cancer prevalence, incidence and mortality (self-report of medically diagnosed; linkage with registrations)	Prospective and retrospective cohort; case-control; cross-sectional; ecologic studies
Exclusion	Children; Animal; workers		Not applicable	Childhood cancer; secondary tumors; cancer survival; hospital admission	Case-report; reviews; in-vitro studies

Table S2. Sample search strategy in Scopus, PubMed, and Web of Science database (until September 04; 2021).

Order	Search query	Number of retrieved objects
Scopus		
1	(TITLE-ABS-KEY (cancer) OR TITLE-ABS-KEY (neoplasms) OR TITLE-ABS-KEY (neoplasm) OR TITLE-ABS-KEY (cancers) OR TITLE-ABS-KEY (tumor) OR TITLE-ABS-KEY (tumors) OR TITLE-ABS-KEY ("cause-specific") OR TITLE-ABS-KEY ("cause specific"))	5,278,871
2	(TITLE-ABS-KEY (greenspace) OR TITLE-ABS-KEY (greenness) OR TITLE-ABS-KEY ("green spaces") OR TITLE-ABS-KEY (greenspaces) OR TITLE-ABS-KEY ("land cover") OR TITLE-ABS-KEY (agriculture) OR TITLE-ABS-KEY (parkland) OR TITLE-ABS-KEY (ndvi) OR TITLE-ABS-KEY (vegetation) OR TITLE-ABS-KEY ("normalized difference vegetation index") OR TITLE-ABS-KEY (msavi) OR TITLE-ABS-KEY (savi) OR TITLE-ABS-KEY ("exposure to nature") OR TITLE-ABS-KEY ("nature exposure") OR TITLE-ABS-KEY ("proximity to nature") OR TITLE-ABS-KEY ("natural environment") OR TITLE-ABS-KEY ("nature contact"))	717,956
3	1 AND 2	4902
PubMed		
1	((((((((((((((greenspace[Title/Abstract]) OR (greenness[Title/Abstract])) OR ("green spaces"[Title/Abstract])) OR (greenspaces[Title/Abstract])) OR ("land cover"[Title/Abstract])) OR (agriculture[Title/Abstract])) OR (parkland[Title/Abstract])) OR (ndvi[Title/Abstract])) OR (vegetation[Title/Abstract])) OR ("normalized difference vegetation index"[Title/Abstract])) OR (MSAVI[Title/Abstract])) OR (SAVI[Title/Abstract])) OR ("exposure to nature"[Title/Abstract])) OR ("nature exposure"[Title/Abstract])) OR ("proximity to nature"[Title/Abstract])) OR ("natural environment"[Title/Abstract])) OR ("nature contact"[Title/Abstract])) OR ("Nature"[Mesh])	82,834
2	((("Carcinogens, Environmental"[Mesh]) OR "Neoplasms"[Mesh]) OR (((((((cancer[Title/Abstract]) OR (cancers[Title/Abstract])) OR (neoplasm[Title/Abstract])) OR (neoplasms[Title/Abstract])) OR (tumor[Title/Abstract])) OR (tumors[Title/Abstract])) OR ("cause-specific"[Title/Abstract])) OR ("cause specific"[Title/Abstract]))	4,406,487
3	1 AND 2	1806
Web of Science		
1	TS=(greenspace) OR TS=(greenness) OR TS=(green?space) OR TS=(green?spaces) OR TS=(greenspaces) OR TS=(land?cover) OR TS=(agriculture) OR TS=(parkland) OR TS=(ndvi) OR TS=(vegetation) OR TS=(normalized?difference?vegetation?index) OR TS=(MSAVI) OR TS=(SAVI) OR TS=(exposure?to?nature) OR TS=(nature?exposure) OR TS=(proximity?to?nature) OR TS=(natural?environment) OR TS=(nature?contact)	440,859

2	TI=(neoplasms) OR TI=(neoplasm) OR TI=("cancer") OR TI=("cancers") OR TI=(tumor) OR TI=("tumors") OR TI=("cause specific") OR TI=("cause-specific") OR AB=(neoplasms) OR AB=(neoplasm) OR AB=("cancer") OR AB=("cancers") OR AB=(tumor) OR AB=("tumors") OR AB=("cause specific") OR AB=("cause-specific")	3,760,487
3	1 AND 2	2174

Table S3. Modified Newcastle-Ottawa Scale (NOS) for assessing the risk of bias in the cohort, case-control and cross-sectional studies

Domain	Item	Description (for gaining the star)	Attainable score
Cohort studies			
Population selection	Representativeness	Participants form general population with random sampling or close to general population (not on specific regions such as only urban participants or specific age group)	One star
	Non-exposed selection	Drawn from the same community as the exposed cohort	One star
	Ascertainment of exposure	Using residential adress at individual level	One star
	No outcome at start	demonstrated that outcome of interest was not present at start of study	One star
Comparability	Comparability	A) Adjustment for at least minimal cat set variables determined by literature and expert discussion, and not inclusion of mediators or over adjustment (age, sex, personal level or area-level socioeconomic status variables [including education, income, area-level deprivation], smoking and urbanity). AND B) Additional model with inclusion of possible mediators (such as air pollution, mental health, stress, physical activity, body mass index, type-2 diabetes, cardiovascular diseases).	A) One star AND B) One star
Outcome	Outcome assessment	Independent blind assessment or record linkage (national or local registrations)	One star
	Length of follow-up	10 years of follow-up	One star
	Adequacy of follow up	Less than 20% loss to follow-up and that the loss to follow-up is not associated with bias in the estimates.	One star
Case-control studies			
Population selection	Representativeness	Representative of the disease of interest	One star
	Case definition	With independent validation (record linkage, hospital registration)	One star
	Control selection	Community control	One star
	Control definition	No history of cancer	One star
Comparability	Comparability	A) Adjustment for at least minimal cat set variables determined by the DAG (and not inclusion of mediators or over adjustment) AND B) Additional model with inclusion of possible mediators	A) One star AND B) One star
Exposure	Exposure assessment	Personal level interview	One star
	Exposure ascertainment	Using residential adress at individual level	One star
	Non-Response rate	Below 20%	One star
Cross-sectional studies			
Population selection	Representativeness	Truly or somewhat representative of the source population (random sampling), or not from specific group, region or ethnicity	One star

	Sample size	Unit of study be at individual level and not area level outcome data.	One star
	Non-respondent	Comparability between respondent and non-respondents, and given participation rate	One star
	Ascertainment of exposure	Individual level at home adress	One star
Comparability	Comparability	A) Adjustment for at least minimal cat set variables determined by the DAG (and not inclusion of mediators or over adjustment) AND B) Additional model with inclusion of possible mediators	A) One star AND B) One star
Outcome	Outcome assessment	National or local registration, medically diagnosed (not self-report) OR Self-report	Two stars OR One star
	Statistical tests	Give data on point estimate and confidence interval	One star

Table S4. Excluded studies after full texts assessment (with exclusion reasons)

No	Item	Reason(s)
1	Crouse, Dan L., et al. "Urban greenness and mortality in Canada's largest cities: a national cohort study." <i>The Lancet Planetary Health</i> 1.7 (2017): e289-e297.	No cancer outcome
2	DeRouen, Mindy C., et al. "Impact of individual and neighborhood factors on disparities in prostate cancer survival." <i>Cancer epidemiology</i> 53 (2018): 1-11.	No relevant cancer outcome (survival outcome in this paper)
3	DeRouen, Mindy C., et al. "Impact of individual and neighborhood factors on socioeconomic disparities in localized and advanced prostate cancer risk." <i>Cancer Causes & Control</i> 29.10 (2018): 951-966.	Not relevant exposure (no report of greenspace exposure)
4	DuPré, Natalie, Jaime E. Hart, and Peter James. "Spatial Analyses of Environmental Exposures and Breast Cancer: Natural Vegetation, Ambient Air Pollution and Outdoor Light at Night as Examples." <i>Geospatial Approaches to Energy Balance and Breast Cancer</i> . Springer, Cham, 2019. 189-219.	Not original finding
5	Gómez-Barroso, Diana, et al. "Agricultural crop exposure and risk of childhood cancer: new findings from a case–control study in Spain." <i>International journal of health geographics</i> 15.1 (2016): 1-11.	Not relevant population (cancer in childhood)
6	Orioli, Riccardo, et al. "Exposure to residential greenness as a predictor of cause-specific mortality and stroke incidence in the Rome Longitudinal Study." <i>Environmental health perspectives</i> 127.2 (2019): 027002.	No cancer outcome
7	Russell, Emily, et al. "Residential racial composition, spatial access to care, and breast cancer mortality among women in Georgia." <i>Journal of Urban Health</i> 88.6 (2011): 1117-1129.	Not relevant exposure (no report of greenspace exposure)
8	Collin, Lindsay J., et al. "Neighborhood-level redlining and lending bias are associated with breast cancer mortality in a large and diverse metropolitan area." <i>Cancer Epidemiology and Prevention Biomarkers</i> 30.1 (2021): 53-60.	Not relevant exposure (no report of greenspace exposure)
9	Keegan, Theresa HM, et al. "Neighborhood influences on recreational physical activity and survival after breast cancer." <i>Cancer Causes & Control</i> 25.10 (2014): 1295-1308.	No relevant cancer outcome (survival outcome in this paper)
10	Granger, Catherine L., et al. "Physical activity measured using global positioning system tracking in non–small cell lung cancer: an observational study." <i>Integrative cancer therapies</i> 13.6 (2014): 482-492.	No relevant exposure and outcome (Non cancer outcome, no greenspace)

11	Ji, John S., et al. "Residential greenness and mortality in oldest-old women and men in China: a longitudinal cohort study." <i>The Lancet Planetary Health</i> 3.1 (2019): e17-e25.	exposure) No cancer outcome
12	Wang, Lan, et al. "Correlation analysis of lung cancer and urban spatial factor: based on survey in Shanghai." <i>Journal of thoracic disease</i> 8.9 (2016): 2626.	Not objective definition of greenspace (defined based on self-report of yes/ no question or qualitative rating).
13	Markevych, Ianna, et al. "Outdoor air pollution, green space, and cancer incidence in Saxony: a semi-individual cohort study." (2018).	Duplicate report (same as Datzman et al. 2018)
14	Wiese, Daniel, et al. "Measuring Neighborhood Landscapes: Associations between a Neighborhood's Landscape Characteristics and Colon Cancer Survival." <i>International journal of environmental research and public health</i> 18.9 (2021): 4728.	No relevant exposure and cancer outcome (survival outcome in this paper; exposure was not greenspace)
15	Villeneuve, Paul J., et al. "A cohort study relating urban green space with mortality in Ontario, Canada." <i>Environmental research</i> 115 (2012): 51-58.	No cancer outcome
16	Jaafari, Shirkou, et al. "Applying landscape metrics and structural equation modeling to predict the effect of urban green space on air pollution and respiratory mortality in Tehran." <i>Environmental Monitoring and Assessment</i> 192 (2020): 1-15.	Not appropriate greenspace indicator (composition metric)
17	Astell-Burt, Thomas, Xiaoqi Feng, and Gregory S. Kolt. "Are the odds of having skin cancer higher among people in greener neighbourhoods due to more active and outdoor lifestyles?: multilevel evidence of survey data from 267,072 Australians." <i>11th Scientific Conference of the Australasian Society for Behavioural Health and Medicine: Health Behaviours: Connecting Individuals and Communities, 12-14 February 2014, Auckland, New Zealand</i> . 2014.	Duplicate report (same as Astell-Burt et al. 2014)
18	Shen, Yanan, Fengyun Sun, and Yue Che. "Public green spaces and human wellbeing: Mapping the spatial inequity and mismatching status of public green space in the Central City of Shanghai." <i>Urban Forestry & Urban Greening</i> 27 (2017): 59-68.	No cancer outcome
19	Wang, Dan, et al. "Neighbouring green space and mortality in community-dwelling elderly Hong Kong Chinese: a cohort study." <i>BMJ open</i> 7.7 (2017): e015794.	No cancer outcome
20	Mears, Meghann, et al. "Greenspace spatial characteristics and human health in an urban environment: An epidemiological study using landscape metrics in Sheffield, UK." <i>Ecological Indicators</i> 106 (2019): 105464.	Not relevant cancer outcome (hospital admission)

Table S5. Levels of adjustment in the included studies in the review

Study	Variable	Astell-Burt (2014)	Bixby et al. (2015)	Carles et al. (2017)	Datzman et al. (2018)	Demoury et al. (2017)	Harrigan (2017)	Iyer et al. (2020)	James et al. (2016)	Kim et al. (2019)	Klompaker et al. (2020)	Lee et al. (2020)	Mitchel et al. (2018)	O'Gallaghan-Gordo et al. (2018)	Richardson et al. (2010) [NZ]	Richardson et al. (2010) [UK]	Shao et al. (2019)	Wu et al. (2008)	Zare Sakhvidi et al. (2020)	Total number with adjustment	Percent of studies with adjustment
Demographic	Age	*		*	*	*		*	*		*	X	*	*	*	*	*	*	*	14	77.8
	Sex	*		*	*	*					*	X	*	*	*	*	*	*	*	8	44.4
	Income		*	*		*					*		*	*	*	*	*	*	*	6	33.3
	Education					*					*			*					*	4	22.2
	Socioeconomic/ occupational status	*				*			*		*			*	*	*	*	*	*	5	27.8
	Marital status	*				*			*								*	*	*	4	22.2
	Occupational exposure			*															*	2	11.1
Lifestyle	Alcohol			*	*	*					*								*	5	27.8
	Smoking status	*		*		*		*	*		*				X	*	*	*	*	10	55.6
	Passive smoking																		*	1	5.56
	Smoking intensity																		*	1	5.56
	Physical activity	*						*			*								*	3	16.7
	Body mass index							*											*	2	11.1
	Nutrition					*													*	2	11.1
	Stress	*																		1	5.56
	Screening/protective behaviors	*						*												2	11.1
Comorbidities	Type 2 diabetes					*		*												2	11.1
	Family history of cancer					*		*												2	11.1
	Race	*				*		*	*											4	22.2
	Country of origin										*									1	5.56
	Number of children													*						1	5.56
Environmental factors	Personal level air pollution		*														*		*	3	16.7
	Area level air pollution											*		*	*	*			*	2	11.1
	Distance to road											*							*	1	5.56
	Climate variables											*								1	5.56
	Pesticide exposure			*																1	5.56
Area level variables	Deprivation/socioeconomic		*			*	*	*	*	*	*	*	*	*	*	*	*	*	*	13	72.2
	Urbanity											*	*						*	2	11.1
	Population density					*		*				*	*	*	*	*			*	6	33.3
	Other area-level variables				*					*	*									2	11.1
Other variables	*		*				*														

Total number of variables	10	3	8	4	12	1	11	6	2	8	7	5	5	8	5	3	0	15
---------------------------	----	---	---	---	----	---	----	---	---	---	---	---	---	---	---	---	---	----

X: Variable measured at area-level; Red cells means adjustment for intermediate variables.

Table S6. Definitions of greenspace across the included papers in the review.

Author	Measure	Definition
Astell-Burt et al. (2014)	Green space	Percentage of parkland green spaces located within a 1 km catchment area around the residence of each participant. Private domestic gardens and agricultural land were not included as they are not typically available for public use.
Bixby et al. (2015)	Green space	The proportion of city area is covered by 'green' land such as woodland, agricultural land, grassland, and other natural vegetated lands.
Carles et al. (2017)	Land cover	Lifelong cumulated exposure score for each crop including open field crops, orchards, and vineyards (Inverse distance multiplied by residential duration at the address, and the addresses' scores were summed for each group).
	Utilized agricultural areas	Percentage of crop surface at 500 m buffer multiplied by residential duration at the address. The scores were summed to create a lifelong cumulated exposure score for open field crops, orchards, and vineyards separately.
Datzman et al. (2018)	Greenness	Averaged composite NDVI images for the years 2005 to 2009.
Demoury et al. (2017)	Greenness at diagnosis	Average NDVI at June (highest greenness) corrected for atmospheric effects, considering cloud cover.
	Greenness 10 years before diagnosis	Average NDVI at August (highest greenness, considering cloud cover) corrected for atmospheric effects, around 10 years before incidence.
Harrigan (2017)	Greenness	Averaged NDVI values during flow up and at baseline
Iyer et al. (2020)	Baseline greenness	Average of NDVI at different seasons (January, April, July, and September) at the start of follow-up.
	Cumulative greenness	Average of NDVI at different seasons (January, April, July, and September) during the follow-up.
	Maximum baseline greenness	Maximum of NDVI at different seasons (January, April, July, and September) at the start of follow-up.
James et al. (2016)	Short term greenness directly accessible	NDVI at 250-m radius for the season of cancer incidence.
	Short term greenness within a 10- to 15-min walk	NDVI at 1250-m radius for the season of cancer incidence.
	Long term greenness directly accessible	Seasonally time-varying NDVI for a representative month in each season (January, April, July, and October) at 250-m radius.
	Short term greenness within a 10- to 15-min walk	Seasonally time-varying NDVI for a representative month in each season (January, April, July, and October) at 1250-m radius.

Kim et al. (2019)	Greenness in peak bloom	Median of averaged non-negative NDVI within the administrative boundary for each district in summer (May–October).
Klompmaker et al. (2020)	Greenness	NDVI at 300 and 500 m buffers at summer time.
	Green space	The proportion of green space within a buffer around the participant's residential address (300 and 1000m) (it does not include private green property (such as gardens) and street greenery).
Lee et al. (2020); Taiwan	Greenness	NDVI at 250 m resolution
	Land cover	Percent of area for the forest, park, and forest/park in township levels.
Mitchel et al. (2018)	Percentage of greenspace	Percentage of parks, other open spaces, and agricultural land (excludes domestic gardens and areas of green space with coverage of less than 5 m ² e.g., single or small clumps of trees on a street) at lower-level super output areas (LSOA) level.
OCallaghan-Gordo et al. (2018)	Presence of urban greenspaces	Presence of public green for predominantly recreational use such as gardens, zoos, parks, of at least 0.25 Ha and/or “forest” (<i>i.e.</i> “forests with a ground coverage of tree canopy>30%, tree height>5 m, including bushes and shrubs at the fringe of the forest” of at least 1 Ha).
	Presence of agricultural areas Surrounding greenness	Presence of arable land, permanent crops, pastures, and heterogeneous agricultural areas. Cloud-free and negative values removed images during spring (the maximum vegetation period for the study region).
Richardson et al. (2010); New Zealand	Total greenspace	Natural areas (e.g., parks, beaches, and fields); excluded aquatic areas (e.g., lakes and the sea).
	Usable greenspace	Urban parkland/open space, beaches, and any non-commercial forestry that was accessible.
	Non-useable greenspace	All other natural areas (non-useable), including agricultural land, salt marsh, and commercial forestry.
Richardson et al. (2010); UK	Urban greenspace availability	Combined percentage coverage of all green spaces larger than 5 m ² (excluding domestic gardens) for each ward (ranged from transport verges and neighborhood greens to parks, playing fields, and woodlands).
Shao et al. (2019)	Greenness	Averaged NDVI values in September (high vegetation coverage) and January (low vegetation coverage) in 2011 as residential NDVI values for each participant.
Wu et al. (2008)	Greenness at summer	NDVI at January of 1990 (cancer data were for 1990-1992)
	Greenness at winter	NDVI at July (cancer data were for 1990-1992)
Zare Sakhvidi et al. (2020)	Greenness	Averaged NDVI values in May-July (high vegetation coverage) during flow up and 10 years before incidence at different buffer sizes (100-1000m)

Land cover

Proximity to different green spaces (agricultural lands, forests, and urban green spaces) defined by CORINE land cover maps.

Table S7. reported associations between exposure to greenspace and lung cancer incidence, prevalence and mortality.

Study name	Outcome	Exposure	Effect size (95% CI)	Effect size per:	Type of effect size
Lung cancer prevalence					
Shao et al. (2019)	C33-C34 ^a	NDVI ^b	0.97 (0.80, 1.18)	1	OR ^c
Lung cancer incidence					
Zare Sakhvidi et al. (2020)	C34	NDVI	0.870 (0.691, 1.095)	0.216	HR ^d
Zare Sakhvidi et al. (2020)	C34	Distance to urban greenspace	1.189 (0.950, 1.489)	4.07 km	HR
Zare Sakhvidi et al. (2020)	C34	Distance to forest greenspace	1.179 (0.982, 1.416)	1.73 km	HR
Zare Sakhvidi et al. (2020)	C34	Distance to agricultural greenspace	1.102 (0.973, 1.248)	0.90 km	HR
Lung cancer mortality					
Bixby et al. (2015) [Men]	C33-C34	NDVI [q2 vs q1]	0.91 (0.81, 1.04)	0.20	RR ^e
Bixby et al. (2015) [Men]	C33-C34	NDVI [q3 vs q1]	0.96 (0.86, 1.10)	0.20	RR
Bixby et al. (2015) [Men]	C33-C34	NDVI [q4 vs q1]	1.06 (0.91, 1.24)	0.20	RR
Bixby et al. (2015) [Men]	C33-C34	NDVI [q5 vs q1]	0.98 (0.86, 1.12)	0.20	RR
Bixby et al. (2015) [Women]	C33-C34	NDVI [q2 vs q1]	1.01 (0.85, 1.21)	0.20	RR
Bixby et al. (2015) [Women]	C33-C34	NDVI [q3 vs q1]	1.09 (0.91, 1.32)	0.20	RR
Bixby et al. (2015) [Women]	C33-C34	NDVI [q4 vs q1]	1.18 (0.94, 1.49)	0.20	RR
Bixby et al. (2015) [Women]	C33-C34	NDVI [q5 vs q1]	1.03 (0.85, 1.24)	0.20	RR
Kim et al. (2019)	C33-C34	NDVI	1.10 (-1.22, 3.47)	0.20	Percent increase
Klomp maker et al. (2020)	C34	NDVI	1.03 (0.97, 1.09)	0.13	HR
Klomp maker et al. (2020)	C34	Percent of total greenspace	1.03 (0.97, 1.09)	24%	HR
Lee et al. (2020)	C33-C34	NDVI	0.87 (0.73, 1.03)	1	RIR ^f
Lee et al. (2020)	C33-C34	Percent of forest	0.88 (0.86, 0.90)	10	RIR
Lee et al. (2020)	C33-C34	Percent of forest or park	0.88 (0.86, 0.90)	10	RIR
Mitchel et al. (2018)	C34	NDVI [q2 vs q1]	1.00 (0.96, 1.04)	0.20	IRR ^g
Mitchel et al. (2018)	C34	NDVI [q3 vs q1]	0.97 (0.93, 1.01)	0.20	IRR
Mitchel et al. (2018)	C34	NDVI [q4 vs q1]	0.94 (0.90, 0.99)	0.20	IRR
Mitchel et al. (2018)	C34	NDVI [q5 vs q1]	0.96 (0.90, 1.02)	0.20	IRR

Richardson et al. (2010); UK	C33-C34	NDVI [q2 vs q1]	1.02 (0.98, 1.07)	0.25	IRR
Richardson et al. (2010); UK	C33-C34	NDVI [q3 vs q1]	1.01 (0.96, 1.06)	0.25	IRR
Richardson et al. (2010); UK	C33-C34	NDVI [q4 vs q1]	0.96 (0.90, 1.02)	0.25	IRR
Richardson et al. (2010); UK	C33-C34	NDVI [q2 vs q1]	1.09 (1.03, 1.15)	0.25	IRR
Richardson et al. (2010); UK	C33-C34	NDVI [q3 vs q1]	1.09 (1.02, 1.15)	0.25	IRR
Richardson et al. (2010); UK	C33-C34	NDVI [q4 vs q1]	1.02 (0.94, 1.11)	0.25	IRR
Richardson et al. (2010); NZ	C33-C34	Percent of total greenspace [q2 vs q1]	1.11 (0.99, 1.23)	25%	IRR
Richardson et al. (2010); NZ	Unclear	Percent of total greenspace [q3 vs q1]	1.09 (0.97, 1.22)	25%	IRR
Richardson et al. (2010); NZ	Unclear	Percent of total greenspace [q4 vs q1]	1.12 (0.94, 1.32)	25%	IRR
Richardson et al. (2010); NZ	Unclear	Percent of useable greenspace [q2 vs q1]	0.99 (0.88, 1.12)	25%	IRR
Richardson et al. (2010); NZ	Unclear	Percent of useable greenspace [q3 vs q1]	1.00 (0.89, 1.13)	25%	IRR
Richardson et al. (2010); NZ	Unclear	Percent of useable greenspace [q4 vs q1]	1.02 (0.90, 1.15)	25%	IRR

- a) International Classification of Diseases (ICD) codes
- b) Normalized difference vegetation index
- c) Odds ratio
- d) Hazard ratio
- e) Relative risk
- f) Risk ratio
- g) Incidence rate ratios

Table S8. reported associations between exposure to greenspace and prostate cancer incidence and prevalence.

Study name	Outcome	Exposure	Effect size (95% CI)	Effect size per:	Type of effect size
Prostate cancer prevalence					
Demoury et al. (20)	Unclear	NDVI ^a	0.84 (0.75, 0.94)	0.10	OR ^b
Prostate cancer incidence					
Datzman et al. (2018)	C61 ^c	NDVI	0.95 (0.90, 1.01)	0.10	RR ^d
Iyer et al. (2020)	Unclear (lethal Prostate cancer)	NDVI	0.95 (0.88, 1.03)	0.11	HR ^e
Zare Sakhvidi et al. (2020)	C61	NDVI	1.071 (0.956, 1.200)	0.23	HR
Zare Sakhvidi et al. (2020)	C61	Distance to urban greenspace	1.302 (0.987, 1.078)	4.07 km	HR
Zare Sakhvidi et al. (2020)	C61	Distance to forest greenspace	1.020 (0.960, 1.083)	1.73 km	HR
Zare Sakhvidi et al. (2020)	C61	Distance to agricultural greenspace	0.990 (0.938, 1.044)	0.90 km	HR

a) Normalized difference vegetation index

b) Odds ratio

c) International Classification of Diseases (ICD) codes

d) Relative risk

e) Hazard ratio

Table S9. reported associations between exposure to greenspace and breast cancer incidence and prevalence.

Study name	Outcome	Exposure	Effect size (95% CI)	Effect size per:	Type
Breast cancer prevalence					
O'Callaghan-Gordo	C50, D05.1, D05.7 ^a	NDVI ^b	1.17 (1.06, 1.28)	0.11	OR ^c
O'Callaghan-Gordo	C50, D05.1, D05.8	Percent of agricultural greenspace	1.35 (1.08, 1.69)	0.13	OR
O'Callaghan-Gordo	C50, D05.1, D05.9	Percent of urban greenspace	0.71 (0.53, 0.95)	0.42	OR
Breast cancer incidence					
Datzman et al. (2018)	C50	NDVI	0.96 (0.92, 0.99)	0.10	RR ^d
Harrigan (2017)	Unclear	NDVI	0.80(0.62, 1.03)	1	HR ^e
Zare Sakhvidi et al. (2020)	C50	NDVI	0.824 (0.687, 0.989)	0.23	HR
Zare Sakhvidi et al. (2020)	C50	Distance to urban greenspace	0.997 (0.935, 1.064)	4.07 km	HR
Zare Sakhvidi et al. (2020)	C50	Distance to forest greenspace	0.959 (0.880, 1.044)	1.73 km	HR
Zare Sakhvidi et al. (2020)	C50	Distance to agricultural greenspace	1.039 (0.910, 1.186)	0.90 km	HR

a) International Classification of Diseases (ICD) codes

b) Normalized difference vegetation index

c) Odds ratio

d) Relative risk

e) Hazard ratio

Table S10. reported associations between exposure to greenspace and skin cancer incidence and prevalence.

Study name	Outcome	Exposure	Effect size (95% CI)	Effect size per:	Type of effect size
Skin cancer prevalence					
Astell-Burt et al. (2014)	Melanoma and/or non- melanoma skin cancer	Percent of greenspace (q2 vs. q1)	1.06 (1.04, 1.09)	0.20	OR ^a
Astell-Burt et al. (2014)	Melanoma and/or non- melanoma skin cancer	Percent of greenspace (q3 vs. q1)	1.09 (1.05, 1.14)	0.20	OR
Astell-Burt et al. (2014)	Melanoma and/or non- melanoma skin cancer	Percent of greenspace (q4 vs. q1)	1.11 (1.05, 1.17)	0.20	OR
Astell-Burt et al. (2014)	Melanoma and/or non- melanoma skin cancer	Percent of greenspace (q5 vs. q1)	1.09 (1.02, 1.16)	0.20	OR
Skin cancer incidence					
Datzman et al. (2018)	Non- melanoma skin cancer, Actinic keratosis	NDVI ^b	0.84 (0.79, 0.90)	0.10	RR ^c
Zare Sakhvidi et al. (2020)	Malignant melanoma skin cancer	NDVI	1.039 (0.748, 1.445)	0.23	HR ^d
Zare Sakhvidi et al. (2020)	Malignant melanoma skin cancer	Distance to urban greenspace	1.189 (0.950, 1.489)	4.07 km	HR
Zare Sakhvidi et al. (2020)	Malignant melanoma skin cancer	Distance to forest greenspace	1.179 (0.982, 1.416)	1.73 km	HR
Zare Sakhvidi et al. (2020)	Malignant melanoma skin cancer	Distance to agricultural greenspace	1.102 (0.973, 1.248)	0.90 km	HR

- a) Odds ratio
- b) Normalized difference vegetation index
- c) Relative risk
- d) Hazard ratio

Table S11. reported associations between exposure to greenspace and colorectal cancer incidence.

Study name	Outcome	Exposure	Effect size (95% CI)	Effect size per:	Type of effect size
Datzman et al. (2018)	C18; C19; C20; C21 ^a	NDVI ^b	1.03 (0.98, 1.07)	0.10	RR ^c
Zare Sakhvidi et al. (2020)	C18; C19; C20	NDVI	1.002 (0.813, 1.235)	0.23	HR ^d
Zare Sakhvidi et al. (2020)	C18; C19; C20	Distance to urban greenspace	0.945 (0.853, 1.047)	4.07 km	HR
Zare Sakhvidi et al. (2020)	C18; C19; C20	Distance to forest greenspace	1.049 (0.936, 1.176)	1.73 km	HR
Zare Sakhvidi et al. (2020)	C18; C19; C20	Distance to agricultural greenspace	0.977 (0.905, 1.054)	0.90 km	HR

a) International Classification of Diseases (ICD) codes

b) Normalized difference vegetation index

c) Relative risk

d) Hazard ratio

Table S12. GRADE assessment for the association between exposure to greenspace and risk of lung cancer mortality

Domains	Assessment	Downgrading/ upgrading
Start level	Seven studies (six cross-sectional and time-series and one cohort)	Low
Risk of bias	In the studies, it was a deficiency in control.	No change
Inconsistency	The point estimates in the studies were in the range of 0.88 to 1.18, however, the confidence intervals were partly overlapped	Downgrade
Indirectness	In general, the population (in three studies), exposure, and outcome were by the PECO. In part of the studies, lung cancer is considered as a negative control	Downgrade
Imprecision	Decision-based on each side of the confidence intervals was associated with different judgments.	Downgrade
Publication bias	Given the comprehensive search and size of the sample in the published studies we decided little even no publication bias.	Downgrade
Dose-response trend	Not reported	No change
Magnitude of associations	In all reported associations, the magnitude of the effect sizes was not large enough to lead to an upgrade of evidence.	No change
Residual confounding	In one out of two studies, the confounding adjustment was completely non-sufficient	No change
Overall judgment		Very low

Table S13. GRADE assessment for the association between exposure to greenspace and risk of prostate cancer incidence

Domains	Assessment	Downgrading/ upgrading
Start level	One cohort, six cross-sectional and ecologic studies	Good
Risk of bias	In the case-control study, the control definition was not acceptable.	Downgrade
Inconsistency	The point estimates in the studies were in the range of 0.84 to 1.12, however, the confidence intervals in most of the studies were not overlapped.	Downgrade
Indirectness	In general, the population, exposure, and exposure were by the PECO.	No change
Imprecision	Decision-based on each side of the confidence intervals was associated with different judgments.	Downgrade
Publication bias	Given the comprehensive search and size of the sample in the published studies we decided little even no publication bias. One study just selected lethal prostate cancer	Unclear
Dose-response trend	Just reported unclearly in one study	No change
Magnitude of associations	In all studies, the magnitude of the effect sizes was not large enough to lead to an upgrade of evidence.	No change
Residual confounding	In two out of four studies the confounding adjustment was not done for air pollution and physical activity. We think adjustment would decrease the observed strength of observed associations.	No change
Overall judgment		Very low

Table S14. GRADE assessment for the association between exposure to greenspace and risk of breast cancer incidence

Domains	Assessment	Downgrading/ upgrading
Start level	Four studies (one case-control and three cohorts)	Good
Risk of bias	One study did acceptable control; whereas in two others it was a deficiency in control. Length of follow-up and representativeness were satisfactory in two of the studies.	Downgrade
Inconsistency	The point estimates in the studies were in the range of 0.8 to 1.35, however, the confidence intervals were not overlapped	Downgrade
Indirectness	In general, the population (in three studies), exposure, and outcome were by the PECO.	No change
Imprecision	Decision based on each side of the confidence intervals was associated with different judgments.	Downgrade
Publication bias	Given the comprehensive search and size of the sample in the published studies we decided little even no publication bias.	No change
Dose-response trend	Not reported	No change
Magnitude of associations	In all reported associations, the magnitude of the effect sizes was not large enough to lead to an upgrade of evidence.	No change
Residual confounding	In one out of two studies, the confounding adjustment was completely non-sufficient	No change
Overall judgment	Very low	

Table S15. GRADE assessment for the association between exposure to Greenspace and risk of skin cancer incidence

Domains	Assessment	Downgrading/ upgrading
Start level	Three studies (one cross-sectional and two cohorts)	Good
Risk of bias	One study did acceptable control; whereas in two others it was a deficiency in control. Length of follow-up and representativeness were satisfactory in two of the studies.	Downgrade
Inconsistency	The point estimates in the studies were in the range of 0.84 to 1.19, however, the confidence intervals were not overlapped	Downgrade
Indirectness	In general, the population (in three studies), exposure, and outcome were by the PECO. Definition of outcome was not homogenous	Downgrade
Imprecision	Decision based on each side of the confidence intervals was associated with similar judgment in two out of three studies.	No change
Publication bias	Given the comprehensive search and size of the sample in the published studies we decided little even no publication bias.	No change
Dose-response trend	One study found a trend. In two other was not reported	No change
Magnitude of associations	In all reported associations, the magnitude of the effect sizes was not large enough to lead to an upgrade of evidence.	No change
Residual confounding	In one out of two studies, the confounding adjustment was not sufficient.	No change
Overall judgment		Very low

Figure 1. PRISMA chart for the study selection process.

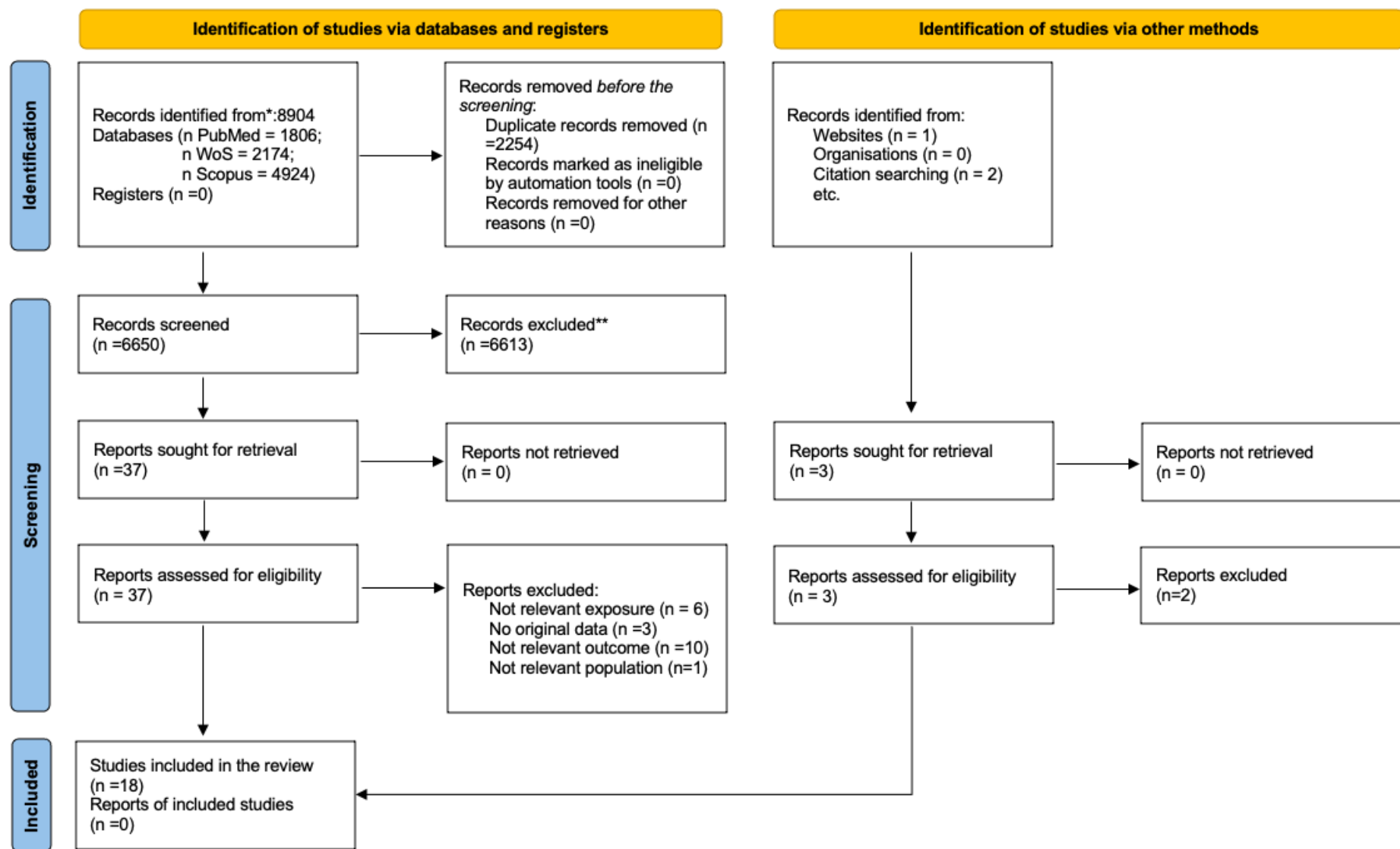


Figure 2. Meta-analyses for the studies reported the associations between greenspace exposure (as 0.1 increase in NDVI) and lung cancer (N represents the number of cancer cases in each study).

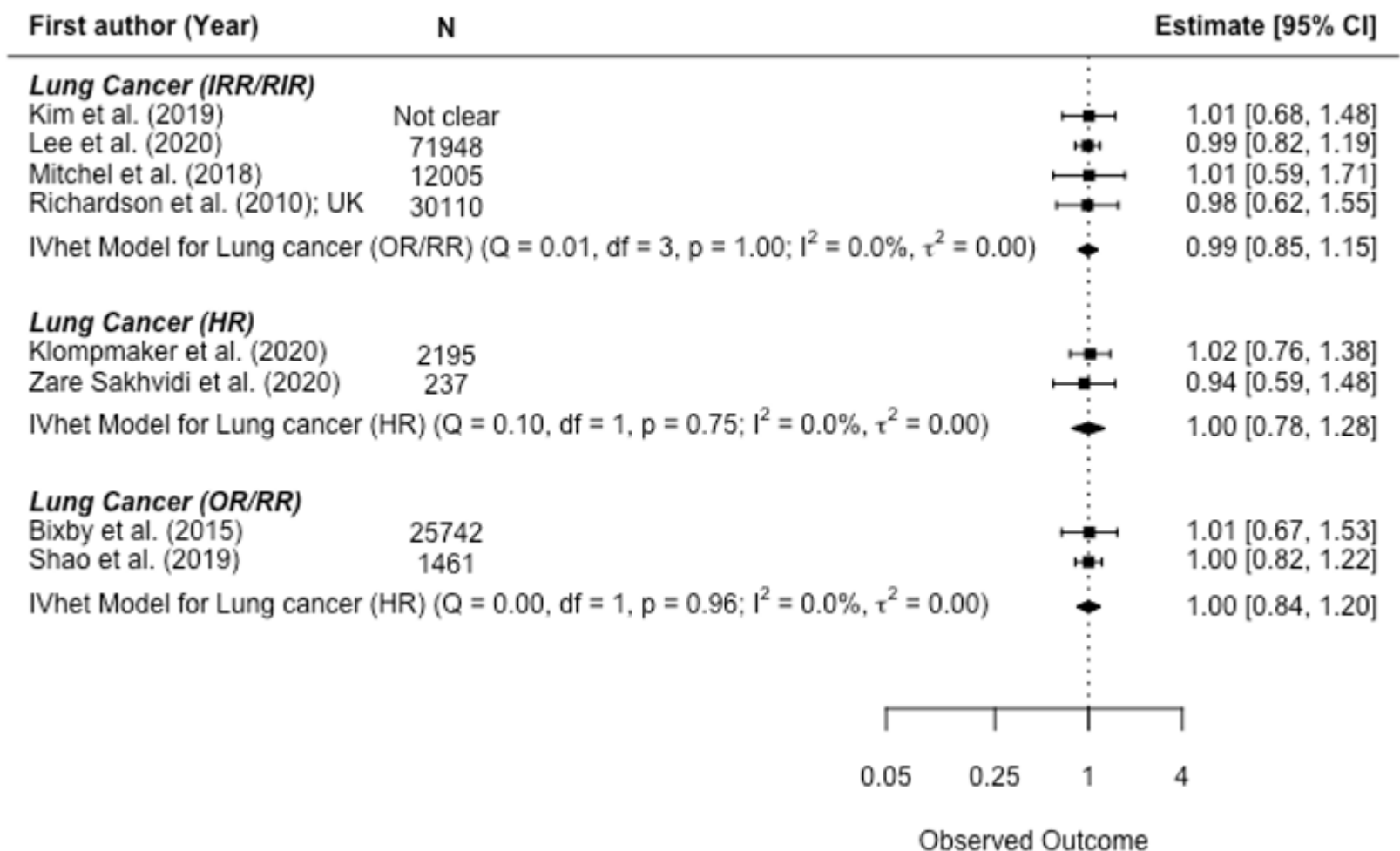


Figure 3. Meta-analyses for the studies reported the associations between greenspace exposure (as 0.1 increase in NDVI) and prostate cancer incidence/prevalence (N represents the number of cancer cases in each study).

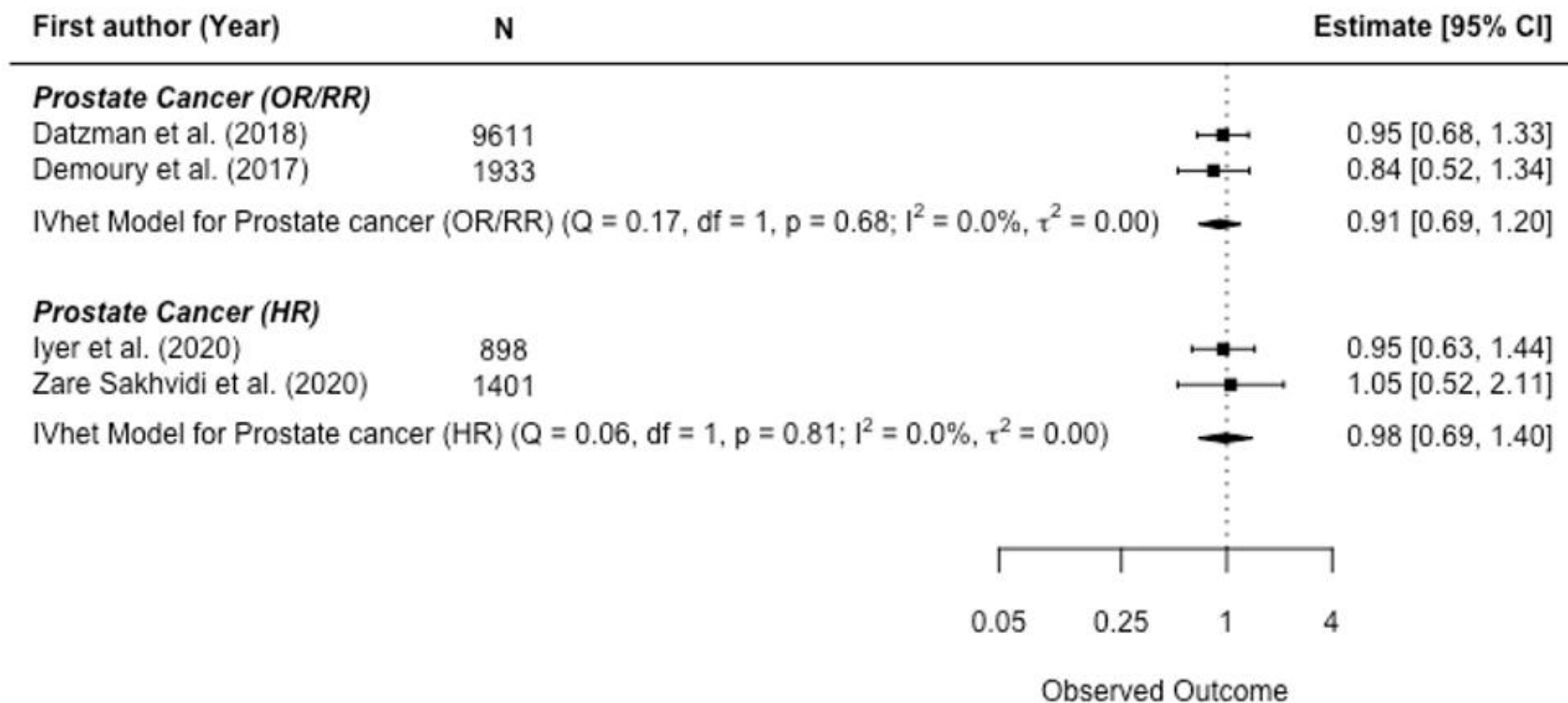


Figure 4. Meta-analyses for the studies reported the associations between greenspace exposure (as 0.1 increase in NDVI) and breast cancer incidence/prevalence (N represents the number of cancer cases in each study).

