**Referee: 1**

General comments:

This manuscript aims to quantify the impact of changes in green spaces on human health in global cities over a ten-year period. The authors use NDVI, aggregated into five-year intervals, to assess city-level changes, which serve as input for an exposure-response function linking NDVI to all-cause mortality. While no major global trends in NDVI or associated health impacts are observed, significant changes are noted in specific cities and years. The paper is well-written and addresses a topic relevant to ERL readers. However, the following issues need to be addressed, as outlined below:

Thank you for your careful read of our paper and for your thoughtful suggestions.

Specific comments:

Health impact assessment

1. It would be helpful to explicitly discuss how increased greenspaces positively impact human health. Which specific diseases see a reduced risk, and through what mechanisms? For example, line 100 mentions that NDVI over a five-year period is used in the health assessment, implying that the exposure-response functions focus on short-term health impacts. Clarifying the pathways through which greenspaces influence health outcomes, along with their respective time scales, would strengthen the analysis. Additionally, how are these effects isolated from other contributing factors? For instance, increased urban greenspaces can lower near-surface temperatures, potentially mitigating extreme events like heat waves. Given that heat-related health impacts can be quantified separately, how do they compare to and differ from the health benefits attributed to NDVI? More importantly, how can these influences be disentangled in the analysis?

Thank you for your comment. In response to the first part of your comment, we’ve added text to the introduction:

“Systematic reviews support an association between increased residential greenspace and decreased risk of depression and anxiety,5 low birth weight,6 cardiovascular events,9 lung and prostate cancer mortality,7 and all-cause mortality8.”

“Three main pathways have been hypothesized to link greenspace with health: reduced environmental harm (i.e. less heat, noise, and air pollution), restoration capacities (i.e. reduced stress), and building capacities (i.e. increased physical activity and social gathering).15 Mediation studies have found evidence that greenspace is associated with improved health through better air quality, increased physical activity, and reduced stress.16”

In terms of the second part of your comment, about the timescale of the influence of greenspace on health, the evidence is less clear. In the studies included in the meta-analysis of exposure to greenspace on all-cause mortality, the exposure is measured in various ways over different time periods. In one study, just one day representing greenest observed NDVI (Orioli (2019)) was used, while Villeneuve (2012) used the greenest cloud-free image from each year of the study period (8 years). In both the James (2016) and Ji (2019) papers, NDVI was measured both contemporaneously (current season) and cumulatively (annual seasonal average over the study period, which consisted of 8 and 14 years, respectively). Vienneau (2017) used the greenest season average from one year, Ziljema (2019) used the greenest season from four different years, while the remaining three studies averaged the greenest season or greenest month across the study period. The follow-up period of the nine studies ranged from 4 years to 18 years (median 8 years). To address your comment, we’ve added text to the methods section:

“We used Rojas-Rueda et al. (2019)’s meta-analysis to define the epidemiologic relationship between increased NDVI and reductions in all-cause mortality. The nine longitudinal studies included in this meta-analysis had follow-up periods ranging from four to 18 years and measured urban greenspace using NDVI. Three studies defined greenspace using the average NDVI value from the greenest season of each year within the study period, while four others uses the greenest day or greenest month from a representative year or years.23–25 To align with the most commonly used exposure metric by the studies included in this meta-analysis, we therefore calculated the population-weighted greenest season NDVI.”

In terms of the last part of your question, the studies included in the meta-analysis from which we derive our exposure-response function control for various covariates. All adjust for age, socio-economic status, and some set of individual and area level factors. Added greenspace is hypothesized to improve health through mitigating other environmental hazards such as air pollution and the urban heat island effect. While five of the nine studies adjusted for some measure of air pollution and two of the nine studies included variables related to climate to disentangle these factors, most did not to allow for reduced environmental harms to be one of the pathways through which greenspace improves health. In our study we rely on the exposure response function from this meta-analysis and do not attempt to isolate the effect of greenspace from reduced heat or air pollution. To address your comment, we have added text to the discussion section:

“Roughly half of the nine studies included in the meta-analysis adjusted for air pollution and two of them controlled for some aspect of climate or temperature. Because of the heterogeneity in confounders across studies, the estimated exposure-response function captures some amount of the benefits from reduced environmental harms such as the urban heat island effect and air pollution. The results presented here likely underestimate the total health benefits from added greenspace and overestimate those provided by greenspace independent of its impact on other environmental harms.”

1. Since the analysis focuses on the peak season, the contribution of individual years may be highly relevant. Wouldn't it be more appropriate to analyze individual years rather than aggregated periods? Additionally, at what scales have the epidemiological functions been developed?

In general, most studies included in the meta-analysis measured exposure to greenspace using a multi-year greenest-season average NDVI (see details from response to question 1). When we looked at yearly trends in NDVI, it did appear that the NDVI values for individual years were highly variable in many cities. We average across 5-year periods to minimize some of this year-to-year variation and instead capture trends in longer term exposure to greenspace. The exact timescale on which greenspace exposure impacts all-cause mortality is unknown. To address your comment, we’ve added text to the methods and discussion sections:

Methods:

“We chose five-year time periods to minimize the effect of year-to-year extremes and capture longer-term trends in urban greenspace exposure.”

Discussion:

“Furthermore, the timescale on which exposure to higher levels of NDVI improves health is unknown. The studies included in the meta-analysis range in follow-up time from four to 18 years. If the changes in NDVI across the two time periods do not reflect true trends but rather temporary increases or decreases, our results will not be applicable to future heath projections.”

1. Starting from line 413, the health impact assessment appears to consider different population groups. However, there is no explanation or equation demonstrating how this dependence is incorporated into the exposure-response function. More details should be provided in the methods section.

Thank you for your comment. We’ve added text to the methods section to clarify our population:

“While the Rojas-Rueda et al. meta-analysis restricted to adults aged 18 and over, we used the total population because that was the gridded population data available from JRC at the 100m pixel resolution. Though children were not included in the Rojas-Rueda et al. study, systematic reviews have linked increased NDVI to higher birth weights6 and increased physical activity among children and adolescents36, and a large national study found that higher NDVI was associated with decreased risk of infant and under-5 mortality.37"

We’ve also added text to the discussion section to expand upon how the differences in age groups across these studies might affect our comparison:

“We estimated that NDVI changes were associated with an average of 2.67 more deaths per 100,000 across the entire set of North American cities. Our results include the total population rather than those 65 and older and are inclusive of 57 cities including 8 Canadian cities. For these reasons, the magnitude of the results is not directly comparable. Furthermore, we found that NDVI decreased over our study period, explaining the difference in sign of our results.”

“Though we included the total population rather than restricting to adults, European cities experienced both positive and negative changes in NDVI over the study period, resulting in health estimates that were smaller in magnitude than those found by Barboza et al. Our use of total population may overestimate the health benefits of increased greenspace and health losses from decreases.”

Methods

1. Line 149 and later: The paper attributes the large interannual variability in NDVI to meteorological factors. If this is the case, to what extent can NDVI reliably represent land-use changes and be used to infer health impacts? Would it be possible to quantify how much NDVI variability is driven by temperature fluctuations versus land-use changes due to urbanization?

In this work, we provide estimates of how health has been impacted by changes in NDVI more broadly and not inferring the magnitude of the impact of these factors, beyond saying it’s a mix of weather, climate change, and urbanization. To attribute changes to these various factors we would need to incorporate many additional datasets beyond the scope of this project.

To address your comment, we’ve added text to the discussion section about how this will be a direction for future work:

“We aim to disentangle the impact of different drivers of changes in NDVI in future work to provide a better understanding of the impact of efforts to expand urban greenspace amidst climate change, urbanization, and meteorologic fluctuations.”

1. Existing literature or simple correlation analyses could help address this question. Additionally, could a land-use dataset that includes both land-use type and urban fraction help identify whether surface property changes have occurred? More evidence is needed to explain and attribute the observed interannual variability in NDVI.

Thank you for your comment. In response to your comment, we have added an additional sensitivity analysis to the appendix comparing the yearly urban fraction of each city using the MODIS landcover product and NDVI:

“To explore the contribution of urbanization to changes in NDVI over time, we looked at the correlation between the proportion of a city that is urban or built up area and the corresponding NDVI value for each year in our study period (2014-2023) (Fig. S7). We derived the urban fraction from NASA’s MODIS landcover dataset, available at a 500m resolution and accessed through Google Earth Engine. We used the University of Maryland’s classification system (‘LC\_Type2’) and used pixels designated as “Urban and Built-up Lands: at least 30% impervious surface area including building materials, asphalt and vehicles” to define the urban fraction. We found a weak negative correlation between NDVI and the urban fraction (-0.312), indicating that greener cities generally have smaller fractions of urban or built-up land. While NDVI and urban fraction were weakly correlated, we found no evidence of correlation between urban fraction and year (0.025), indicating that urbanization may not be a large contributor to changes in NDVI over the study period across cities as a whole (Fig. S8). There is a mismatch in the spatial scale between our measurement of NDVI (100m) and the urban fraction (500m), which could also be contributing to the weak correlation that we observed.”

1. Over the 10-year period analyzed, many cities have undergone changes in size. How is this accounted for? If city boundary changes are measured, how are they incorporated into the health impact assessment? It would be helpful to interpret the observed changes for specific cities in relation to their growth, urbanization levels, and other relevant factors. Additionally, is there variability within cities at the sub-city level? If so, how is it considered when computing an aggregate metric for individual cities?

Thank you for your points. We have used the GHS-SMOD shapefiles to define city boundaries, which are based on built-up area and population density from 2019. We use the same urban boundary for all the included years. You are correct that urbanization could be another factor that is driving the changes we observed in city-level NDVI over time. We have added text to the discussion section to reflect your important point:

“Urbanization in the past decade could also contribute to these changes, as we used a consistent urban boundary definition across the ten-year period, however cities may have grown and morphed over this time.”

To your last point, yes there is variability within cities in NDVI and population. We performed the health impact assessment at the 100m x 100m grid cell level to account for this. Population data were only updated every 5-years, so we have used 2015 gridded population data for years 2014-2018 and 2020 data for years 2019-2024. We have reorganized the methods section to make clearer the scale of the analysis:

“We calculated this value for each 100m pixel (i) (Equation 2).”

Technical comments

* 1. Line 33: why is NDVI higher and more stable in European and North American cities? The sentence continues by talking about epidemiological studies performed in those areas, however the two aspects are unrelated. Please clarify.

We found that NDVI was higher and more stable in European and North American cities. Because most epidemiological studies were conducted in these regions, there is uncertainty in the applicability of the exposure-response curve outside these regions where urban greenness levels and trends are different. This was stated to highlight the need for future epidemiologic studies in more diverse geographic settings. In response to your comment, we have reworded this section of the abstract to be clearer:

“Health impact assessments of NDVI and all-cause mortality have largely been conducted in European and North American cities, where we found NDVI was generally higher and more stable. Our results highlight large heterogeneity in urban greenspace extent and variability across global cities and the importance of characterizing the relationship between health and NDVI in more diverse contexts.”

* 1. Line 62: is this true everywhere? I believe cities in the developing world have very different emission regimes than the ones in developed countries. Please revise.

The study we cite found that this was the general trend in developed countries and the case for many developing nations. The current phrasing likely overstates the evidence in developing countries, and we have revised to reflect your comment:

“While cities are responsible for over 80% of global greenhouse gas emissions,2 emissions per capita in developed nations tend to be lower in cities than in less dense communities due to more efficient transportation, energy production, and land use.3”

* 1. Line 145: what is “i” in Equation 2? Does it refer to the pixels? How is the HR defined? An explicit expression for that should be provided. Also, it is not clear which spatial resolution NDVI and population data have and how they are homogenized in the present analysis.

Thank you for your comment. We have added language to the methods section to add clarity to these points:

“To calculate the PAF, we used the hazard ratio (*HR*) from a meta-analysis of the protective effect of NDVI on all-cause mortality, which found a pooled hazard ratio of 0.96 (95% confidence interval (CI): 0.94, 0.97) for each 0.1 increase in NDVI within 500m of a person’s home.”

“We used the difference between the average 2014-2018 and 2019-2023 population-weighted greenest season NDVI to define changes in urban greenspace at the 100m pixel (i) level to align with the resolution of our population dataset ().”

“We then downscaled the NDVI dataset to the 100m resolution to align with our population dataset.”

* 1. Figure 3: what do the colored dots represent and the box plots? More details are needed in the caption.

The dots represent cities. We’ve added text to the captions of Figure 3 and Figure 4 for clarity:

“*Each dot represents a city, colored by geographic region.”*

“*Each dot represents a city, colored by climate classification.”*

* 1. Line 309: is the IQR: 0.13 – 8.5?

Because this was reporting fewer deaths, we had arranged the IQR from 8.5-0.13. In response to your comment, we have rearranged from 0.13-8.5,

“Eastern Asia had a median reduction of 4.72 (IQR: 0.13, 8.52) annual premature deaths per 100,000 population…”

* 1. Line 319: this section is confusing. What does a median change in mortality of 0.01 fewer means? There is mention of a change in NDVI in that sentence, but how much change? Also, it would be useful to contextualize changes in NDVI due to various factors. E.g. the mentioned 0.19 change in NDVI how can that be interpreted? Is it typical of a greener season due to more favorable weather conditions or is it the change expected when land use changes from urban/concrete to forest? Some references are needed through the text.

Thank you for your comment. We have reworded the beginning of this paragraph as well as moved the “deaths per 100,000” before the ranges to make this section easier to read. The changes in NDVI refer to the changes reported earlier in the paper.

“We also considered NDVI-associated mortality changes by climate classification. Arid cities had stable NDVI values over time, and this was reflected in the median associated changes in mortality, which was very close to zero at 0.01 fewer deaths per 100,000 (range: 12.90 fewer to 12.14 more) (Fig. 5B).”

Because we are averaging across all cities within a climate classification or region, it is hard to contextualize the NDVI changes. For example, if they were to increase from 0.0 to 0.1 this would mean that the area went from urban or barren land to sparsely vegetated, while an increase from 0.3 to 0.4 would represent an increase in the density or health (more green leaves) of the vegetation. We are averaging spatially within cities and also across all cities within these groupings so there is no easy interpretation of this increase.

* 1. Line 326: the ranges expressed as “fewer to more” are quite confusing, so it would be clearer if there was a sentence explaining how this wording will be used and interpreted

We have added text to be more explicit about the interpretation of this wording.

“The premature mortality impact from urban greenspace change was not evenly distributed around the world, with fewer associated deaths in areas that experience increases in NDVI across the time periods and more associated deaths in areas where NDVI decreased (Fig. 5A). The range in associated mortality from greenspace changes spanned fewer to more deaths, reflecting that there were cities across all regions that experienced both increases and decreases in NDVI.”

* 1. Figure 5: remove “Associated”.

We have changed the figure caption to:

*“Changes in city-level mortality per 100,000 population associated with changes in average population-weighted peak season Normalized Difference Vegetation Index (NDVI) from 2014-2018 to 2019-2023 to the 2020 population, by geographical region (panel A) and climate classification (panel B).”*

* 1. Line 385: the impact on individual cities would be very relevant to be discussed and possibly compared with cities in the same region that do not experience such changes.

Thank you for your comment. In response to your suggestion, we have added text to the results section, highlighting regional outliers in greenspace change:

“There were many outlier cities across several regions. For example, five Venezuelan cities: Barcelona, Maturin, Barquisimeto, Maracay, and Valencia had increases in NDVI across the two periods despite a general decline in urban greenspace across Latin America and the Caribbean. Buram, Sudan in Northern Africa and Gonda, India in Southern Asia were also positive greenspace outliers. In contrast, many cities were negative greenspace outliers in their regions including Auckland, New Zealand; San Antonio and Providence, United States; Mataram, Indonesia; Lakhimpur, India; Drachevo, Macedonia; and Dortmund and Wuppertal, Germany. There is likely a mix of driving factors contributing to each of these cities’ greenspace changes. Some of the negative outliers such as Auckland, San Antonio, Mataram, Lakhimpu, and Drachevo have experienced urbanization over the past decade that may be contributing to their decline in greenspaces. Other cities situated near one another such as the five cities of northern Venezuela and the two German cities likely have experienced similar temperature and rainfall changes due to weather and climate change.”

* 1. Line 419: most of the paper mentioned that NDVI had remained stable over the period analyzed. However, the authors mention that NDVI has decreased over time. This seems to contradict what discussed before.

This reference is to just North American cities for the comparison with Brochu et al. We have added an explicit reference to North America in this sentence for clarity.

“Furthermore, we found that NDVI decreased in North American cities over our study period, explaining the difference in sign of our results”

* 1. Conclusions are too short. I suggest merging with discussion.

Thank you for your comment. We have followed the ERL article structure guidelines, which includes a separate Conclusion section.

**Referee: 2**  
  
COMMENTS TO THE AUTHOR(S)  
In this study, the authors attempt to quantify the contribution of NDVI to mortality in global cities. The attempt is ambitious and interesting, and I appreciate the authors' efforts. However, the validity of the methods and assumptions used in this study are questionable, and thus, the results and conclusions are not convincing enough. I do not recommend that this article be published in ERL.

Thank you for taking the time to read and review our study.  
  
In this study, temporal changes in NDVI (i.e., the difference between 2014-2018 and 2019-2023) are translated to the differences in mortality. Temporal changes should be discriminated against spatial variations. In the original cohort studies on which the meta-analysis (Rojas-Rueda et al., 2019) is based, in principle, the risks are calculated from the spatial variation of the NDVI. It is also indicated that the temporal change in NDVI is not associated with mortality by a cited study (Ji et al., 2019). There is no justification for attributing the mortality to temporal changes in NDVI, as assumed in this study.

Thank you for your comment. The exposure-response function provided by the Rojas-Rueda et al. meta-analysis does not specify an exposure time scale. The studies included measure NDVI in various ways including contemporaneous and cumulative NDVI exposure. The (Ji et al, 2019) study is the only study to assess changes over time and did not find a significant association with mortality using this exposure definition. However, they defined changes in NDVI by grouping areas into significant increases, no significant change, and significant decreases based on the slope coefficient from a linear regression of annual average NDVI from 2000-2014. We observed substantial inter-annual changes in NDVI that jumped up and down across the years. We averaged two 5-year periods to minimize some of this meteorological noise and attempt to capture changes due to urbanization, climate change, and greenspace interventions that will impact residents’ current and future greenspace exposure. Additionally, health impacts of other environmental exposures, such as PM2.5, commonly assume that it is appropriate to apply a spatial exposure-response function to temporal changes in exposure.1

To address your comment, we’ve added text discussing the temporal aspect of the exposure measurement to the introduction section:

“The nine longitudinal studies included in this meta-analysis had follow-up periods ranging from four to 18 years and measured urban greenspace using NDVI. Three studies defined greenspace using the average NDVI value from the greenest season of each year within the study period, while four others uses the greenest day or greenest month from a representative year or years.23–31”

We’ve also added text to the discussion section to be explicit about the limitations of our approach with respect to the temporal nature of the exposure measurement:

“Furthermore, the timescale on which exposure to higher levels of NDVI improves health is unknown. The studies included in the meta-analysis range in follow-up time from four to 18 years. If the changes in NDVI across the two time periods do not reflect true trends but rather temporary increases or decreases, our results will not be applicable to future heath projections.”

The authors apply one single risk function to all the cities all over the world. This risk function is based on the nine studies, most of which are conducted in developed countries in temperate climate zones. The derived relationship might be applicable to similar cities. For example, however, in developing countries in tropical climate zones, higher NDVI could be associated with higher exposure to vector-borne diseases transmitted by mosquitoes (e.g., malaria). Therefore, applying the relationship to global cities is questionable and probably inadequate.

We agree that the use of one global exposure-response function is a limitation of this study. However, the current evidence base linking greenspace and all-cause mortality does not support a city-specific approach. To address this, we chose a large-scale meta-analysis to be as generalizable as possible. While ideally, we would have city-, age-, gender-, and socioeconomic-specific risk curves, the approach we use here reflects the current state of the science and is commonly used in multi-city and multi-country health impact assessments. Many studies quantifying the health impacts of global environmental exposures use a generalized exposure-response function. For example, the Global Burden of Disease study uses a global relative risk when quantifying the associated health burden of ambient air pollution including particulate matter2 and nitrogen dioxide.3 Additionally, many health impact assessments from the past five years of greenspace4-11, air pollution9-12, noise9-10, and heat10-11 have relied on a single exposure-response functions.

References:

1. Dockery DW, Pope CA, Xu X, et al. An Association between Air Pollution and Mortality in Six U.S. Cities. *New England Journal of Medicine*. 1993;329(24):1753-1759. doi:10.1056/NEJM199312093292401
2. Ambient particulate matter pollution—Level 4 risk. <https://www.thelancet.com/pb-assets/Lancet/gbd/summaries/risks/ambient-particulate-matter-pollution.pdf>
3. https://www.healthdata.org/sites/default/files/methods\_appendices/2021/wozniak\_air\_no2\_writeup\_GBD2021\_030222\_AC\_updated030624.pdf
4. Brochu P, Jimenez MP, James P, Kinney PL, Lane K. Benefits of Increasing Greenness on All-Cause Mortality in the Largest Metropolitan Areas of the United States Within the Past Two Decades. Front Public Health. 2022 May 10;10:841936. doi: 10.3389/fpubh.2022.841936. PMID: 35619828; PMCID: PMC9127575.
5. Barboza EP, Cirach M, Khomenko S, Iungman T, Mueller N, Barrera-Gómez J, et al. Green space and mortality in European cities: a health impact assessment study. The Lancet Planetary Health. 2021 Oct;5(10):e718–30.
6. Garber MD, Guidi M, Bousselot J, Benmarhnia T, Dean D, Rojas-Rueda D. Impact of native-plants policy scenarios on premature mortality in Denver: A quantitative health impact assessment. Environment International. 2023 Aug;178:108050.
7. Dean D, Garber MD, Anderson GB, Rojas-Rueda D. Health implications of urban tree canopy policy scenarios in Denver and Phoenix: A quantitative health impact assessment. Environmental Research. 2024 Jan 15;241:117610.
8. Kondo MC, Mueller N, Locke DH, Roman LA, Rojas-Rueda D, Schinasi LH, et al. Health impact assessment of Philadelphia’s 2025 tree canopy cover goals. The Lancet Planetary Health. 2020 Apr 1;4(4):e149–57.
9. Pereira Barboza E, Montana F, Cirach M, Iungman T, Khomenko S, Gallagher J, et al. Environmental health impacts and inequalities in green space and air pollution in six medium-sized European cities. Environmental Research. 2023 Nov;237:116891.
10. Mueller N, Rojas-Rueda D, Khreis H, Cirach M, Andrés D, Ballester J, et al. Changing the urban design of cities for health: The superblock model. Environment International. 2020 Jan;134:105132.
11. Pereira Barboza E, Nieuwenhuijsen M, Ambròs A, Sá THD, Mueller N. The impact of urban environmental exposures on health: An assessment of the attributable mortality burden in Sao Paulo city, Brazil. Science of The Total Environment. 2022 Jul;831:154836.
12. Southerland VA, Brauer M, Mohegh A, Hammer MS, Van Donkelaar A, Martin RV, et al. Global urban temporal trends in fine particulate matter (PM2·5) and attributable health burdens: estimates from global datasets. The Lancet Planetary Health. 2022 Feb;6(2):e139–46.
13. Malashock DA, Delang MN, Becker JS, Serre ML, West JJ, Chang KL, et al. Global trends in ozone concentration and attributable mortality for urban, peri-urban, and rural areas between 2000 and 2019: a modelling study. The Lancet Planetary Health. 2022 Dec;6(12):e958–67.

**Referee: 3**  
  
COMMENTS TO THE AUTHOR(S)  
This paper performs a health impact assessment of changes in urban NDVI in more than 1000 cities across the globe, and finds on average mild changes although with important variability between cities.  
  
It is clearly of interest to provide an idea of to which extent green space impact public health and the global scope is an additional strength of this study. The method is also appropriate overall although I think it comes with important uncertainty that is not really acknowledged (more on that in the below comments). It is not very clear to me however why these specific counterfactuals have been used (2014-18 vs 2019-23) instead of more usual counterfactual such as a fixed increase or some policy objective.

Thank you for taking the time to review our study and for your thoughtful feedback.

## Comments  
1. It seems that the exposure used here differs slightly from the studies estimating the HR, and I wonder to which extent this could impact the mortality estimates.  
- Has the exclusion of water pixels done in the epidemiological studies? It is mentioned in the introduction that blue spaces have alleged benefits on health, and that negative NDVIs tend to represent water. One could then imagine a kind of V-shaped relationship between NDVI and mortality, that could perhaps impact the estimation of HR. Maybe removing water pixels would change the potential range of NDVI here and therefore represent a slightly different exposure?

Thank you for your comment. Not all the studies included in the meta-analysis report how they have treated water in their analyses. However, four of the five studies (Wilker (2014), Crouse (2017), Nieuwenhuijsen (2018), and Ziljema (2019)) that do mention the handling of water removed water pixels from their analysis. We removed pixels representing water to be consistent with this majority.

To address your comment, we have added text to the methods section:

“Following the methods used by many of the studies included in the meta-analysis of greenspace and mortality that we use for our health impact assessment, we removed pixels representing water and clouds.”

- I cannot find any mention of "greenest season" NDVI in the meta-analysis of Rojas-Rueda et al. (2019) and was wondering to which extent it could change the picture here.

While the meta-analysis does not mention greenest season NDVI because the nine studies included in the meta-analysis use a range of different NDVI-based exposure metrics, the majority of the studies define greenspace in this way. Crouse (2017), Wilker (2014) and Nieuwenhuijsen (2018) used the greenest season average across all years in their study. Vienneau (2017) used the greenest season average from one year, Ziljema (2019) used the greenest season from four different years, Orioli (2019) used the greenest day, and Villeneuve (2012) averaged across the greenest cloud-free image from each year of the study period (8 years). James (2016) and Ji (2019) were the only two studies to look at annual average. They measured NDVI both contemporaneously (current season) and cumulatively (annual seasonal average over the study period, which consisted of 8 and 14 years, respectively).

To address your comment, we’ve expanded the language in the methods section to be clearer about the exposure metric used by the studies in the meta-analysis:

“Three studies defined greenspace using the average NDVI value from the greenest season of each year within the study period, while four others uses the greenest day or greenest month from a representative year or years.23–31 To align with the most commonly used exposure metric by the studies included in this meta-analysis, we therefore calculated the population-weighted greenest season NDVI.”

2. I would say the main shortcoming of this analysis is the lack of any uncertainty estimate (save for the grand mean). There seems to be a lot heterogeneity in the reported HRs in Rojas-Rueda et al. (2019) which means high variability of the overall one. Additionally, I can imagine there is some uncertainty related to exposure assessment, notably because of cloud cover. I think confidence intervals (or any measure of uncertainty) should be included in the analysis for the reader to get a sense of how uncertain the whole exercise here (and high uncertainty is not a flaw of any analysis). Note that it is quite easy to sample from the distribution of the HR and compute CIs from there.

Thank you for your comments. We agree that it is important to communicate the uncertainty of our results. To address your comment, we have added 95% confidence bounds to the reported health impact assessment results section (in addition to the graphs in the appendix that display error bounds).

While we have not sampled from the distribution of the HR, we have used the upper and lower bounds of both the baseline mortality rates and HR estimate to approximate the error that simulating these draws would create. We have found in previous work that this approach reduces computational burden in generating the uncertainty estimates while still closely approximating the uncertainty bounds were we to use another technique like Monte Carlo simulation.

There is also uncertainty related to the exposure assessment, because a range of approaches to using NDVI were used in the included studies. However, our approach reflects the most common method used by the meta-analysis studies. Cloudy pixels were removed in our analysis. Not all the studies in the meta-analysis mentioned how they handled clouds, but all those that did (N=5) used cloud-free images and an additional study restricted to positive values which removes cloudy images and water. Unfortunately, the satellite-derived NDVI estimates do not provide error estimates. The confidence bounds of the HR is the only estimate we have of error from the included studies using different NDVI definitions.

We have rounded all reported absolute results to the nearest 100 deaths so as not to overstate the precision.

3. Why restrict the analysis to cities larger than 500,000 inhabitants (or the biggest one of the country)? The GHS-UCDB is much more comprehensive than that and the NDVI is a global product so, unless I am missing something, nothing prevents from expanding the number of cities covered.

You are correct that the GHS-UCDB contains many more cities (~10,000 cities). We restricted to cities in this way to reflect the universe of cities included in The Lancet Countdown’s urban greenspace metric. We have added language to be clearer about how and why the study population was chosen in the methods section:

“We included the 1,041 cities for which urban greenspace was estimated by the Lancet Countdown on health and climate change. The Lancet Countdown included cities if they were the most populous in their country or had over 500,000 inhabitants.”

4. It would be a nice addition to provide absolute figures in the results, i.e. in this set of cities, how many deaths have been lost/averted due to NDVI changes? At the moment only death rates are reported. Note that Eq. 3 promises attributable numbers to the reader.

Thank you for your comment. In response to your comment, we have added some absolute results to the main text and included absolute results in Figure 4.

“In absolute terms, Eastern Asia had the largest health gains from changes in NDVI with an estimated 20,600 avoided deaths (95%CI: 13,300, 36,100) across all cities. Sub-Saharan Africa has the greatest health burden from urban greenspace changes, with a total of 9,100 more deaths (95% CI: 6,000, 15,800).”

“In absolute terms, there was an estimated 3,300 fewer (95% CI: 25,000 fewer- 16,600 more) greenspace-associated deaths globally. Continental cities had the greatest reductions, with an estimated 10,900 (95% CI: 7,300, 10,900) fewer deaths, while tropical cities had the greatest increases (17,300, 95% CI: 11,300, 29,800).”

## Minor  
- 112: "Joint Research Commission" is actually the "Joint Research Centre"

Thank you for catching this. We have updated the text to reflect this edit.

- I think the title of the paper is slightly misleading and should indicate that the paper looks at changes in NDVI. "A quantitative health impact assessment" promises a more expanded assessment than what is done in the paper.

Thank you for your comment. We have updated the study title to: “A health impact assessment of changes in NDVI on all-cause mortality across 1,041 global cities”

- The figures are a bit difficult to read and I suspect they would be difficult to analyse for someone with colour-blindness. I would suggest thinner and fainter lines for border in Figure 2 and boxplots in figure 3, with perhaps slightly enhanced points.

Thank you for your suggestions, we have updated the borders of the maps and boxplots in Figures 2, 3, and 4 and increased the point sizes in Figures 3 and 4.

- Similarly, I would suggest reordering the regions in Figure 3, so that regions from the same continent are next to each other. Another ordering could be by "average latitude" or something like that.

We have reordered Figure 3 and the regional graphs of Figure 5 to be arranged by average latitude per your suggestion.