Supplemental Digital Content of:

**Associations between Greenness and Predicted COVID-19 like Illness Incidence in the U.S. and U.K.**

Kelly Chen1\*, Jochem O Klompmaker1,2\*, Charlotte J. Roscoe1, Long H Nguyen3, David A. Drew3,4, Peter James1,5, Francine Laden1,2, Daniela Fecht6, Weiyi Wang6, John Gulliver7, Jonathan Wolf8, Claire J. Steves9, Tim D. Spector9, Andy T. Chan3,4,10, Jaime E. Hart 1,2

1. *Department of Environmental Health, Harvard T.H. Chan School of Public Health, Boston, MA, USA*
2. *Channing Division of Network Medicine, Department of Medicine, Brigham and Women’s Hospital and Harvard Medical School, Boston, MA, USA*
3. *Clinical & Translational Epidemiology Unit, Massachusetts General Hospital and Harvard Medical School, Boston, MA, USA*
4. *Division of Gastroenterology, Massachusetts General Hospital and Harvard Medical School, Boston, MA, USA*
5. *Department of Population Medicine, Harvard Medical School and Harvard Pilgrim Health Care Institute, Boston, MA, USA*
6. *MRC Centre for Environment and Health, School of Public Health, Imperial College London, London, UK*
7. *Centre for Environmental Health and Sustainability, George Davies Centre, University of Leicester, Leicester, UK*
8. *ZOE Global, London, UK*
9. *Kings College Department of Twin Research and Genetic Epidemiology, King’s College London, London, UK*
10. *Department of Immunology and Infectious Disease, Harvard T.H. Chan School of Public Health, Boston, MA, USA*

*\* contributed equally as first author*

**Corresponding author**

Jochem O Klompmaker

jklompmaker@hsph.harvard.edu

Department of Environmental Health, Harvard T.H. Chan School of Public Health, Boston, MA, USA

Channing Division of Network Medicine, Department of Medicine, Brigham and Women’s Hospital and Harvard Medical School, Boston, MA, USA

**Conflicts of interest**

The authors declare they have no actual or potential competing financial interests.

**Sources of financial support**

This work was supported by grant U01 CA176726 from the National Cancer Institute, grant U01 HL145386 from the Heart, Lung and Blood Institute, grant R01 HD057368 from the Eunice Kennedy Shriver National Institute of Child Health & Human Development, and grants R01 ES028033, R01 ES028712, R01 ES029840 and P30 ES000002 from the National Institute of Environmental Health Sciences.

**Data sharing statement**

The COVID Symptom Study (CSS) smartphone application data used in this study cannot be made available publicly due to privacy issues. We used data from Landsat 8 to assess greenness. This data is freely available in Google Earth Engine. A statistical analysis plan is available from the corresponding author (Jochem O Klompmaker) on reasonable request.

**eTable 1.** Baseline characteristics of all cohort participants and by U.S. and U.K. participants from March 24, 2020 to November 30, 2020 after excluding those who did not report baseline information, reported any symptoms or a positive COVID-19 test at or within 24 hours of enrollment, or had <24 hours of follow-up time.

|  |  |  |  |
| --- | --- | --- | --- |
|  | **Full cohort** | **U.K. cohort** | **U.S. cohort** |
|  | **Median [25, 75 pct] / %** | **Median [25, 75 pct] / %** | **Median [25, 75 pct] / %** |
| **N (persons)** | 2,794,029 | 2,557,464 | 236,565 |
| **Age** | 46.0 [32.0, 59.0] | 46.0 [32.0, 58.0] | 55.0 [37.0, 67.0] |
| **Gender (%)** |  |  |  |
| Women | 60.6 | 60.3 | 64.5 |
| Men | 39.3 | 39.7 | 35.3 |
| Prefer not to say | 0.1 | 0 | 0.1 |
| (Missing) | 0 | 0 | 0 |
| **Race/ethnicity (%)** |  |  |  |
| White, non-Hispanic | 81.7 | 82.3 | 75.3 |
| Hispanic/Latino | 0.4 | 0 | 5.2 |
| Black | 0.6 | 0.4 | 2.4 |
| Asian | 2.1 | 1.9 | 3.7 |
| Mixed/Other Race | 2.6 | 2.6 | 2.8 |
| Prefer not to say | 0.1 | 0 | 0.9 |
| (Missing) | 12.5 | 12.7 | 9.7 |
| **BMI** | 25.3 [22.3, 29.1] | 25.2 [22.3, 29.0] | 25.8 [22.7, 30.0] |
| **Smoker status (%)** |  |  |  |
| Never smoker | 46.1 | 43.6 | 73 |
| Former smoker | 13.1 | 12.3 | 21.7 |
| Current smoker | 5 | 4.9 | 5.2 |
| (Missing) | 35.9 | 39.2 | 0.1 |
| **Has Diabetes (%)** | 3.1 | 2.9 | 5 |
| **Has Heart disease (%)** | 2.9 | 2.6 | 5.9 |
| **Has Lung disease (%)** | 10.1 | 10.1 | 10 |
| **Has Kidney disease (%)** | 0.7 | 0.7 | 1.4 |
| **Frontline healthcare worker, Yes (%)** | 4.9 | 4.5 | 9.2 |
| **Health problems requiring stay-at-home (%)** | 5.1 | 5.1 | 4.8 |
| **Regular use of mobility aid (%)** | 2.1 | 2.1 | 2 |
| **Health problems limiting activities (%)** | 8.5 | 8.4 | 9 |
| **Population Density (persons/mi2)** | 4347.5 [807.6, 10312.3] | 5007.2 [907.2, 10868.8] | 1662.9 [457.3, 4025.6] |
| **Urban/Rural** |  |  |  |
| Rural | 22.5 | 20.9 | 39.2 |
| Urban | 60.1 | 60.1 | 60.2 |
| Missing | 17.4 | 19.0 | 0.6 |
| **English IMD Income Score** | 7.0 [5.0, 9.0] | 7.0 [5.0, 9.0] | NA [NA, NA] |
| **English IMD Education Score** | 7.0 [5.0, 9.0] | 7.0 [5.0, 9.0] | NA [NA, NA] |
| **U.S. Census – Median Household Income** | 77333.0 [59088.0, 101061.0] | NA [NA, NA] | 77333.0 [59088.0, 101061.0] |
| **U.S. Census – Percent over 25 years with Bachelors** | 43.7 [29.7, 58.6] | NA [NA, NA] | 43.7 [29.7, 58.6] |
| **PM2.5 (µg/m3)** | 9.6 [8.6, 10.2] | 9.6 [8.8, 10.2] | 8.4 [6.9, 9.3] |

Abbreviations: IQR, interquartile range; BMI, body mass index; IMD, Index of Multiple Deprivation

Note: the median (25th – 75th percentile) area of the neighborhoods in the U.K. (lower layer super output areas in England and Wales and Datazones in Scotland) is 0.59 km2 (0.29 – 2.94) and the median area of the neighborhood in the U.S. (Zip Code Tabulation Areas) is 44.65 km2 (18.15 – 120.85).

**eTable 2.** Baseline characteristics of all cohort participants and by predicted COVID-19 like illness cases and COVID-19 identified by a PCR test cases from March 24, 2020 to November 30, 2020 after excluding those who did not report baseline information, reported any symptoms or a positive COVID-19 test at or within 24 hours of enrollment, or had <24 hours of follow-up time.

|  |  |  |  |
| --- | --- | --- | --- |
|  | **Full cohort** | **Predicted COVID-19 like illness** | **COVID-19 identified by a PCR test** |
|  | **Median [25, 75 pct] / %** | **Median [25, 75 pct] / %** | **Median [25, 75 pct] / %** |
| **N (persons)** | 2,794,029 | 143,340 | 22,784 |
| **Age** | 46.0 [32.0, 59.0] | 37.0 [26.0, 51.0] | 44.0 [30.0, 55.0] |
| **Gender (%)** |  |  |  |
| Women | 60.6 | 59.4 | 63.7 |
| Men | 39.3 | 40.5 | 36.2 |
| Prefer not to say | 0.1 | 0.0 | 0.0 |
| (Missing) | 0.0 | 0.0 | 0.0 |
| **Race/ethnicity (%)** |  |  |  |
| White, non-Hispanic | 81.7 | 79.2 | 86.1 |
| Hispanic/Latino | 0.4 | 0.4 | 0.5 |
| Black | 0.6 | 0.5 | 0.6 |
| Asian | 2.1 | 2.0 | 2.9 |
| Mixed/Other Race | 2.6 | 3.4 | 2.7 |
| Prefer not to say | 0.1 | 0.1 | 0.1 |
| (Missing) | 12.5 | 14.4 | 7.2 |
| **BMI** | 25.3 [22.3, 29.1] | 25.4 [22.0, 29.8] | 25.6 [22.6, 29.8] |
| **Smoker status (%)** |  |  |  |
| Never smoker | 46.1 | 36.7 | 43.3 |
| Former smoker | 13.1 | 9.2 | 11.6 |
| Current smoker | 5.0 | 5.2 | 3.5 |
| (Missing) | 35.9 | 48.8 | 41.5 |
| **Has Diabetes (%)** | 3.1 | 2.9 | 3.1 |
| **Has Heart disease (%)** | 2.9 | 2.1 | 2.1 |
| **Has Lung disease (%)** | 10.1 | 13.8 | 11.4 |
| **Has Kidney disease (%)** | 0.7 | 0.7 | 0.7 |
| **Frontline healthcare worker, Yes (%)** | 4.9 | 6.8 | 13.4 |
| **Health problems requiring stay-at-home (%)** | 5.1 | 5.9 | 3.9 |
| **Regular use of mobility aid (%)** | 2.1 | 2.3 | 1.4 |
| **Health problems limiting activities (%)** | 8.5 | 11.4 | 7.0 |
| **Population Density (persons/mi2)** | 4347.5 [807.6, 10312.3] | 5475.5 [1254.4, 11521.1] | 5007.3 [1251. 6, 10706.5] |
| **Urban/Rural** |  |  |  |
| Rural | 22.5 | 18.9 | 18.9 |
| Urban | 60.1 | 64.3 | 65.0 |
| Missing | 17.4 | 16.9 | 16.2 |
| **English IMD Income Score** | 7.0 [5.0, 9.0] | 7.0 [4.0, 9.0] | 7.0 [4.0, 9.0] |
| **English IMD Education Score** | 7.0 [5.0, 9.0] | 7.0 [4.0, 9.0] | 7.0 [5.0, 9.0] |
| **U.S. Census – Median Household Income** | 77333.0 [59088.0, 101061.0] | 71242.0 [55485.0, 92550.0] | 70457.0 [54543.0, 91723.0] |
| **U.S. Census – Percent over 25 years with Bachelors** | 43.7 [29.7, 58.6] | 38.8 [26.6, 54.3] | 37.7 [25.4, 52.6] |
| **Summer NDVI** | 0.52 [0.37, 0.63] | 0.46 [0.35, 0.55] | 0.46 [0.35, 0.55] |

Abbreviations: IQR, interquartile range; BMI, body mass index; IMD, Index of Multiple Deprivation

**eTable 3.** Baseline characteristics of all included cohort participants and of individuals that were excluded because they reported any COVID-19 symptoms or a positive COVID-19 test within 24 hours of enrollment, or had <24 hours of follow-up time.

|  |  |  |
| --- | --- | --- |
|  | **Full cohort** | **Excluded individuals with <24hours of follow-up time** |
|  | **Median [25, 75 pct] / %** | **Median [25, 75 pct] / %** |
| **N (persons)** | 2,794,029 | 1,477,166 |
| **Country (%)** |  |  |
| U.K. | 91.5 | 92.3 |
| U.S. | 8.5 | 7.7 |
| **Age** | 46.0 [32.0, 59.0] | 37.00 [28.00, 49.00] |
| **Gender (%)** |  |  |
| Women | 60.6 | 47.8 |
| Men | 39.3 | 27.6 |
| Prefer not to say | 0.1 | 24.6 |
| (Missing) | 0 | 0 |
| **Race/ethnicity (%)** |  |  |
| White, non-Hispanic | 81.7 | 60.8 |
| Hispanic/Latino | 0.4 | 1.6 |
| Black | 0.6 | 0.6 |
| Asian | 2.1 | 0.5 |
| Mixed/Other Race | 2.6 | 1.8 |
| Prefer not to say | 0.1 | 0.1 |
| (Missing) | 12.5 | 34.7 |
| **BMI** | 25.3 [22.3, 29.1] | 25.4 [22.2, 29.7] |
| **Smoker status (%)** |  |  |
| Never smoker | 46.1 | 50.0 |
| Former smoker | 13.1 | 7.7 |
| Current smoker | 5 | 5.3 |
| (Missing) | 35.9 | 37.1 |
| **Has Diabetes (%)** | 3.1 | 1.8 |
| **Has Heart disease (%)** | 2.9 | 1.5 |
| **Has Lung disease (%)** | 10.1 | 8.1 |
| **Has Kidney disease (%)** | 0.7 | 0.5 |
| **Frontline healthcare worker, Yes (%)** | 4.9 | 5.9 |
| **Health problems requiring stay-at-home (%)** | 5.1 | 4.5 |
| **Regular use of mobility aid (%)** | 2.1 | 1.5 |
| **Health problems limiting activities (%)** | 8.5 | 7.6 |
| **Population Density (persons/mi2)** | 4347.5 [807.6, 10312.3] | 5592.28 [1254.39, 11754.43] |
| **Urban/Rural** |  |  |
| Rural | 22.5 | 12.0 |
| Urban | 60.1 | 40.6 |
| Missing | 17.4 | 47.4 |
| **English IMD Income Score** | 7.0 [5.0, 9.0] | 6.0 [4.0, 9.0] |
| **English IMD Education Score** | 7.0 [5.0, 9.0] | 7.0 [4.0, 9.0] |
| **U.S. Census – Median Household Income** | 77333.0 [59088.0, 101061.0] | 72206.0 [54785.0, 95348.0] |
| **U.S. Census – Percent over 25 years with Bachelors** | 43.7 [29.7, 58.6] | 38.7 [25.2, 54.8] |
| **Summer NDVI** | 0.52 [0.37, 0.63] | 0.45 [0.35, 0.55] |

Abbreviations: IQR, interquartile range; BMI, body mass index; IMD, Index of Multiple Deprivation

**eTable 4.** Hazard ratios (and 95% CIs) for predicted COVID-19 like illness incidence presented by a 0.1 increase in 3-month moving average neighborhood NDVI, among participants in the full COVID Symptom Study cohort (*n* = 2,794,029) among U.K. participants (*n* = 2,557,464) and among U.S. participants (*n* = 236,565) during follow-up from March 24, 2020 to November 30, 2020.

|  |  |  |  |
| --- | --- | --- | --- |
|  | **Full cohort** | **U.K. cohort** | **U.S. cohort** |
| **HR (95% CI) a** | **HR (95% CI) a** | **HR (95% CI) a** |
| **Main model** | 0.965 (0.960, 0.970) | 0.964 (0.959, 0.969) | 0.994 (0.975, 1.013) |
| **Summer NDVI b** | 0.971 (0.966, 0.975) | 0.968 (0.964, 0.973) | 0.992 (0.974, 1.010) |
| **Not adjusted for month c** | 0.823 (0.819, 0.827) | 0.815 (0.811, 0.818) | 0.935 (0.918, 0.951) |
| **Adjusted for season instead of month d** | 0.874 (0.870, 0.878) | 0.867 (0.863, 0.871) | 0.970 (0.951, 0.988) |
| **Not adjusted for population density e** | 0.955 (0.951, 0.959) | 0.953 (0.948, 0.957) | 0.999 (0.982, 1.015) |
| **Adjusted for county/authority district f** | 0.974 (0.969, 0.980) | 0.975 (0.969, 0.980) | 0.995 (0.973, 1.016) |
| **Adjusted for PM2.5 g** | 0.966 (0.961, 0.971) | 0.962 (0.957, 0.967) | 1.004 (0.983, 1.025) |
| **COVID-19 identified by PCR test h** | 0.936 (0.925, 0.947) | 0.928 (0.917, 0.940) | 0.995 (0.956, 1.036) |

Abbreviations: HR, hazard ratio; CI, Confidence Interval

a Stratified by age group, country, and calendar month at study entry, andadjusted for: calendar month, gender, race, BMI, comorbidities (including diabetes, heart disease, lung disease, and kidney disease), current smoking status, frontline healthcare worker status, population density (quartiles), neighborhood level-income (quintiles), neighborhood-level education (quintiles)

b Model using summer neighborhood NDVI instead of a 3-month moving average neighborhood NDVI

c Model not adjusted for calendar month

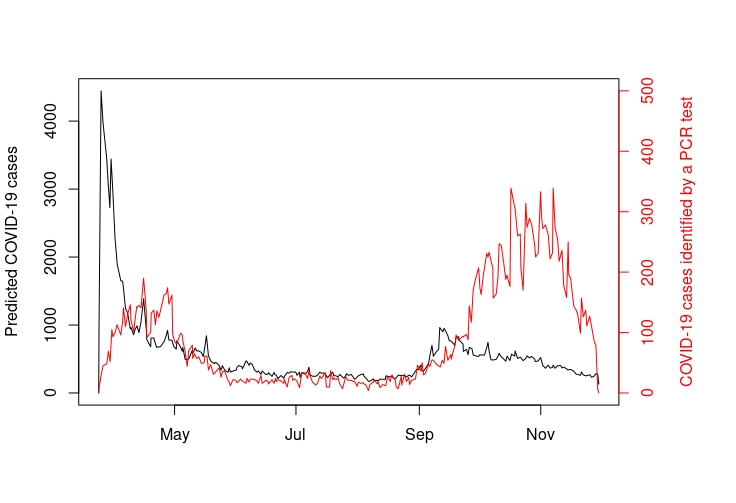
d Model adjusted for season instead of calendar month

e Model not adjusted for population density

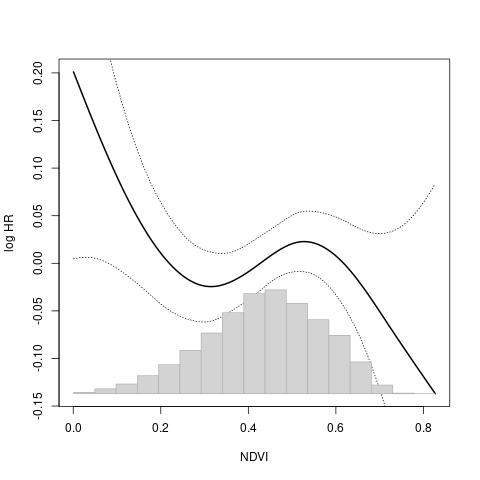
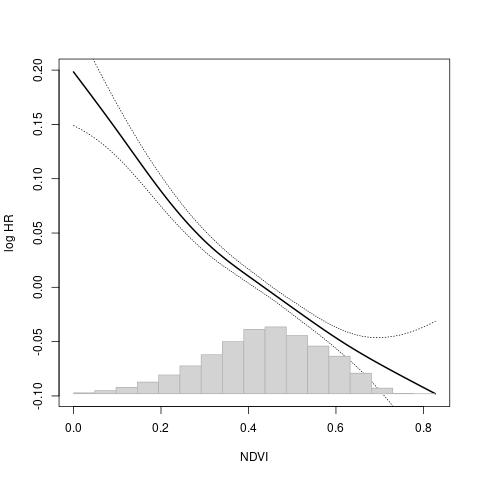
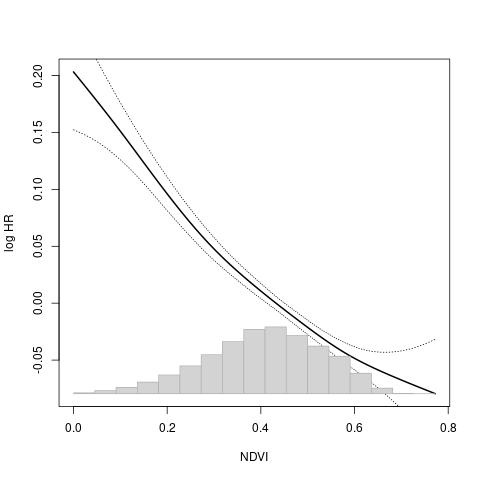
f Model included a frailty term for county/local authority (n=379)

g Model adjusted for PM2.5

h Model using incident instead of predicted COVID-19. In the U.K., there were 21,227 COVID-19 cases identified by a PCR test and in the U.S., there were 1557 COVID-19 cases identified by a PCR test.

****

**eFigure 1.** The number of predicted COVID-19 like illness cases (black) and incident COVID-19 cases identified by a PCR test (red) among participants in the COVID Symptom Study cohort during follow-up from March 24, 2020 to November 30, 2020. Note that the y-scale for predicted COVID-19 like illness cases differs from the y-scale for incident COVID-19 cases identified by a PCR test.

****

**eFigure 2.** The exposure-response curve of 3-month moving average neighborhood NDVI (3 degrees of freedom) with predicted COVID-19 like illness incidence in the full COVID Symptom Study cohort (left figure, *n* = 2,794,029), among U.K. participants (middle figure, *n* = 2,557,464) and among U.S. participants (right figure, *n* = 236,565) during follow-up from March 24, 2020 to November 30, 2020, in the fully adjusted model.