Role of air temperature and humidity in the transmission of SARS-CoV-2 in the United States

Yiqun Ma[1,2], Sen Pei[3*], Jeffrey Shaman[3], Robert Dubrow[1,2], Kai Chen[1,2*]

1 Department of Environmental Health Sciences, Yale School of Public Health, 60 College Street, New Haven, CT, 06520-8034, USA

2 Yale Center on Climate Change and Health, Yale School of Public Health, 60 College Street, New Haven, CT, 06520-8034, USA

3 Department of Environmental Health Sciences, Mailman School of Public Health, Columbia University, New York, NY 10032, USA.

*Corresponding authors: K. Chen, Yale School of Public Health, 60 College Street, New Haven, CT, 06511, USA, kai.chen@yale.edu; S. Pei, Mailman School of Public Health, Columbia University, New York, NY 10032, USA, sp3449@cumc.columbia.edu

Abstract

Improved understanding of the effects of meteorological conditions on the transmission of SARS-CoV-2, the causative agent for COVID-19 disease, is urgently needed to inform mitigation efforts. Here, we estimated the relationship between air temperature or specific humidity (SH) and SARS-CoV-2 transmission in 913 U.S. counties with abundant reported infections from March 15 to August 31, 2020. Specifically, we quantified the associations of daily mean temperature and SH with daily estimates of the SARS-CoV-2 reproduction number (*Rt*) and calculated the fraction of *Rt* attributable to these meteorological conditions. Both lower temperature and lower SH were significantly associated with increased *Rt*. The fraction of *Rt* attributable to temperature was 5.10% (95% eCI: 5.00 - 5.18%), and the fraction of *Rt* attributable to SH was 14.47% (95% eCI: 14.37 - 14.54%). These fractions generally were higher in northern counties than in southern counties. Our findings indicate that cold and dry weather are moderately associated with increased SARS-CoV-2 transmissibility, with humidity playing a larger role than temperature.

Introduction

Since emerging in Wuhan, China, the novel severe acute respiratory syndrome coronavirus 2 (SARS-CoV-2), the causative agent of coronavirus disease 2019 (COVID-19), has produced a major global pandemic. As of November 12, 2020, approximately 10.6 million COVID-19 cases and 243 thousand deaths had been reported in the U.S.[1], more than any other country. The decreased stability of SARS-CoV-2 in warmer temperatures and higher humidity in laboratory experiments[2,3], and the documented seasonality of influenza[4] and infections caused by other coronaviruses[5-7], lead to the hypothesis that lower air temperature and lower humidity are associated with increased SARS-CoV-2 transmission. Quantifying this effect on a population level is urgently needed to help inform public health control efforts, including transmission prevention and communication with the public[8].

Numerous preliminary studies have found either positive or negative associations of air temperature and humidity with COVID-19 cases[9-13]. However, given the large number of undocumented SARS-CoV-2 infections[14], the variations in the lag between infection and symptom onset, and the inconsistent lag between testing and reporting, using daily new confirmed cases may not be optimal for examining meteorological effects[15]. As a result, a few studies have used the reproduction number to estimate SARS-CoV-2 transmissibility[16-18]. One study reported high daily air temperature and high daily relative humidity (RH) to be associated with a reduced daily effective reproduction number (the mean number of new infections caused by a single infected person in a population in which some individuals may no longer be susceptible due to acquired immunity[19]) for SARS-CoV-2 in both China and the U.S.[16]. However, two early studies focused on the first few of months of the pandemic found no association between temperature or humidity and the basic reproduction number (the mean number of new infections caused by a single infected person in a population in which everyone is assumed to be susceptible and no public health measures have been implemented)[17,18].

Early analyses, in particular, should be interpreted with caution[8], as the range of temperature and humidity measurements during the short observation period at the beginning of the pandemic was

relatively narrow in most studies[9-12,16-18], thus limiting the ability to detect associations between these meteorological variables and SARS-CoV-2 transmission. In addition, many previous studies (whether using COVID-19 cases or reproduction number as the outcome) controlled for no or only a few potential confounders[9-13,17,18], which include other environmental factors, socioeconomic factors, temporal changes in population immunity, and implementation of public health interventions.

Furthermore, although most early studies found an association between air temperature or humidity and COVID-19 incidence, the fraction of cases or deaths attributable to meteorological conditions remains unclear. One modeling study predicted that as long as most of the population is susceptible to infection, any role of humidity in SARS-CoV-2 transmission would be overwhelmed by the lack of population immunity[20]. This prediction is supported by the rapid transmission of SARS-CoV-2 regardless of climate zone, including warmer locations such as tropical Brazil, India and southern states in the U.S. during the northern hemisphere summer[1].

Here we investigate the association between air temperature or specific humidity (SH; the mass of water vapor in a unit mass of moist air [g/kg]) and SARS-CoV-2 transmission, as measured by the reproduction number *Rt* (the mean number of new infections caused by a single infected person, given the public health measures in place, in a population in which everyone is assumed to be susceptible). We estimate *Rt* in the 913 counties with at least 400 cumulative cases as of August 31, 2020 and calculate the fraction of *Rt* attributable to temperature or SH, adjusting for a wide range of potential confounders.

Results

Distribution of meteorological factors and Rt

From March 15 to August 31, 2020, a total of 4,903,520 cases of COVID-19 were reported in the 913 study counties (Extended Data Table 1). We estimated the county-specific *Rt* using a dynamic metapopulation model informed by human mobility data that represents the transmission of SARS-CoV-2 in the U.S. (see Methods). Mean daily *Rt* averaged over all counties and days during the study period was 1.40 and ranged from 0.46 to 5.43. Daily air temperature and SH also ranged widely (air temperature: -14.61 - 39.98 °C; SH: 0.99 - 22.15 g/kg). Union County, New Jersey had the highest *Rt* averaged over the

study period (Fig. 1a). The largest number of cumulative cases per 100,000 people was observed in Chattahoochee County, Georgia, while Taylor County, Florida had the lowest number (Fig. 1b). Southern counties generally were hotter and more humid than northern counties; whereas the western U.S., coastal counties generally were cooler and more humid than inland counties (Fig. 1c-d).

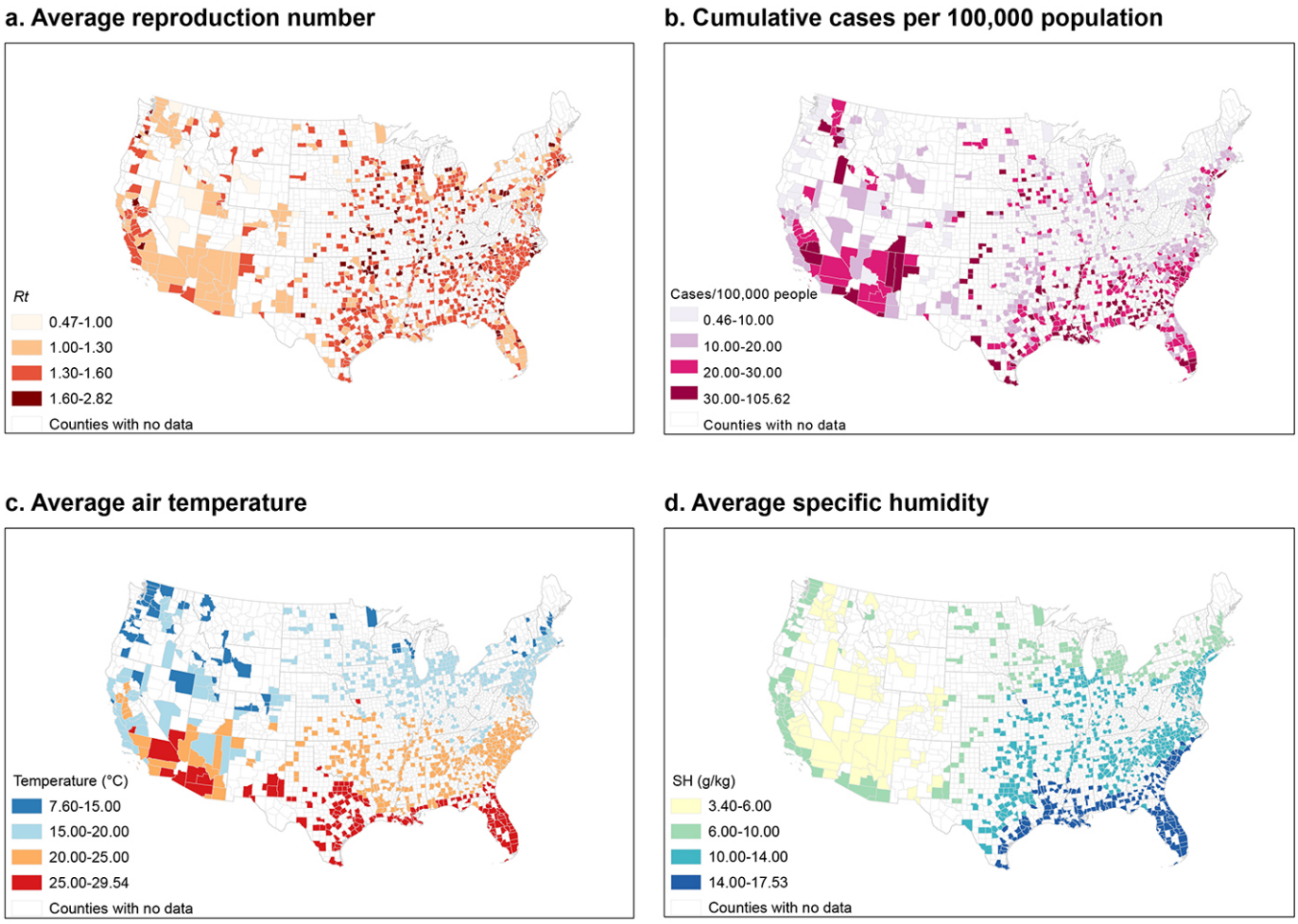
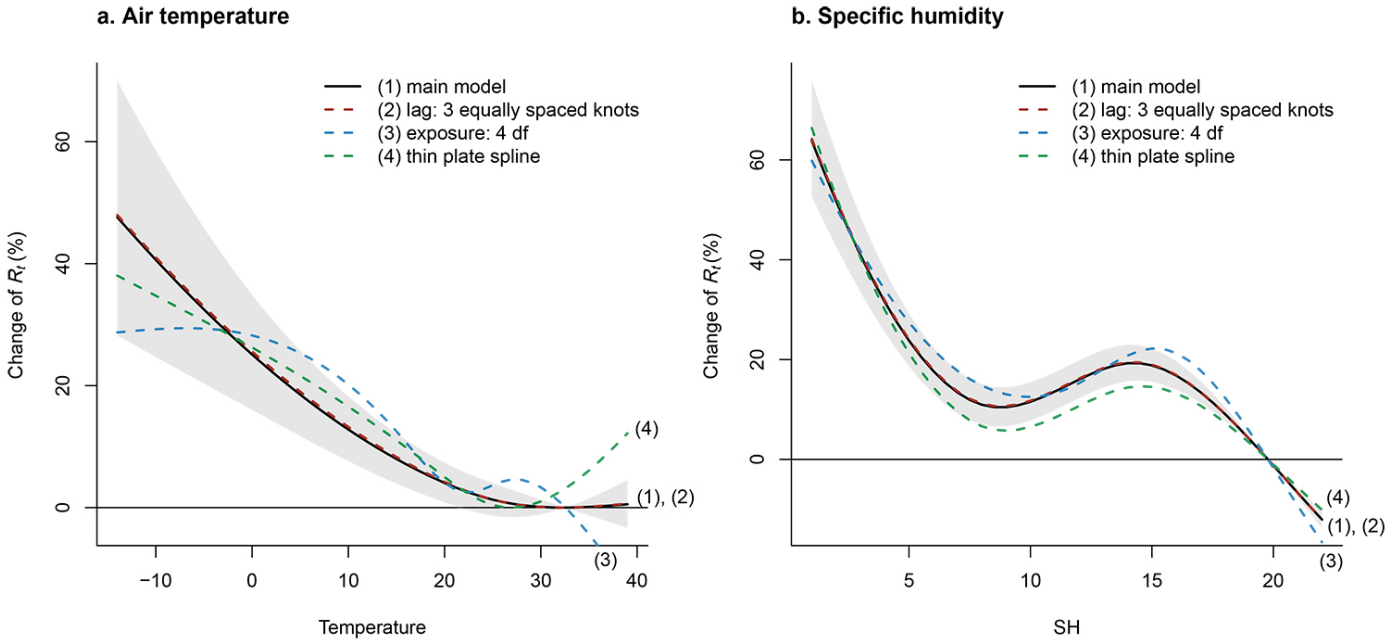


Fig. 1. Map of the distribution of reproduction number, cumulative cases, air temperature and specific humidity in study counties.

These maps display the distribution of the daily reproduction number (*Rt*), daily air temperature, and daily specific humidity (SH) averaged over the study period, and the cumulative cases per 100,000 population, in 913 U.S. counties.

Associations between meteorological factors and Rt

We estimated the complex non-linear and temporally delayed associations of meteorological factors with the SARS-CoV-2 *Rt* using a generalized additive mixed model adjusting for spatiotemporal variations in *Rt* and potential measured confounders, described in detail in Methods. We then calculated the optimum values of temperature and SH, which correspond to the lowest *Rt*. We found an approximately linear inverse temperature-*Rt* relationship (Fig. 2a), with lower air temperatures significantly associated with increased transmission of SARS-CoV-2 when below the optimum temperature (32.57 °C). No significant associations were observed for temperatures above the optimum value. The relationship between SH and *Rt* was non-linear (Fig. 2b). Higher SH was significantly associated with decreased transmission, except for an increasing trend from approximately 9 to 15 g/kg. The optimum SH was estimated to be 19.78 g/kg. Compared with the optimum value, the 10^{th} percentile of the distribution of air temperature (8.8 °C) or SH (4.5 g/kg) was associated with a 14.10% (95% CI: 8.59 - 19.89%) and 27.49% (95% CI: 21.93 - 33.30%) increase of *Rt*, respectively. Effect estimates showed a decreasing trend in the lag dimension, diminishing to a small non-significant effect on lag day 13 (Extended Data Fig. 1). Sensitivity analyses showed the estimated relationships between air temperature or SH and *Rt* were generally consistent under different modeling choices (Fig. 2a-b).



107

108 **Fig. 2. The associations of air temperature (°C) and specific humidity (g/kg) with *Rt*, under**
109 **different choices of model.**

110 This figure shows the estimated exposure-response curves for the associations of air temperature (°C) and
111 specific humidity (g/kg) with reproduction number (*Rt*) for SARS-Cov-2, with different modelling
112 choices: (1) main model with 95% confidence interval (grey area): tensor product smooths to control for
113 the temporal and spatial variations, and a cross-basis term for air temperature and SH, which is defined by
114 natural cubic splines with 3 df for both the exposure-response and lag-response association, with a
115 maximum lag of 13 days; (2) redefine the lag dimension using a natural cubic spline and 3 equally placed
116 internal knots in the log scale; (3) change the df to 4 in the cross-basis term for air temperature or SH in
117 the exposure-response function; (4): use a thin plate spline to control for geographical coordinates and
118 time instead of using the tensor product smooths.

119

120 ***Fractions of Rt attributable to meteorological factors***

121 Based on the estimated exposure-response curves and daily county-specific *Rt*, we further calculated the
122 fraction of *Rt* attributable to temperature or SH (i.e., the attributable fraction (AF), which can be
123 interpreted as the fraction of *Rt* attributable to the deviation of temperature or SH from the optimum
124 value). Across all 913 counties over the entire study period, the AF for temperature was 5.10% (95%

empirical confidence intervals [95% eCI]: 5.00 - 5.18%), and the AF for SH was 14.47% (95% eCI: 14.37 - 14.54%) (Extended Data Table 2). The AF for temperature showed an increasing trend from south to north (Fig. 3a). The county with the highest AF for temperature was Whatcom County, Washington (12.88%) and the county with the lowest AF for temperature was Hidalgo County, Texas (0.52%). The AF for SH showed an increasing trend from south to north in the eastern U.S., whereas in the western U.S., the AF for SH was lower in counties in coastal states than in counties in interior states (Fig. 3b). The county with the highest AF for SH was Nye County, Nevada (27.47%), and the county with the lowest AF for SH was Plaquemines Parish, Louisiana (7.24%). The AF for temperature was the largest in March and April, and the lowest in July and August (Fig. 3c). The AF for SH showed a modest decline between March and August (Fig. 3d).

Sensitivity analyses indicate that the AF for air temperature remains robust when excluding socioeconomic factors and when additionally adjusting for smoking and obesity prevalence, long-term air pollution, or short-term air pollution (Extended Data Table 2). However, the estimated AF for temperature decreased from 5.11% to 3.55% after additionally adjusting for daily ultraviolet (UV). The AF for SH was robust across all sensitivity analyses.

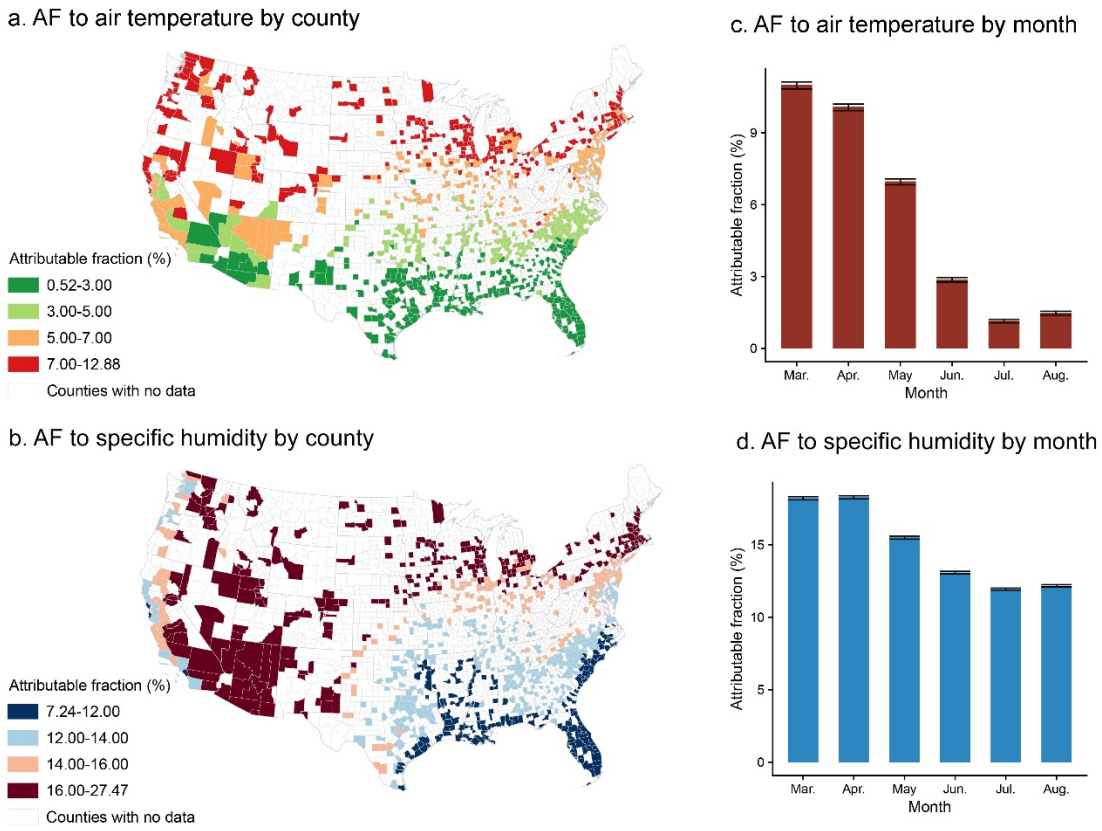


Fig. 3. Fractions of *Rt* attributable to meteorological factors by county and month

a, b: the distribution of the fraction of reproduction number (*Rt*) attributable to temperature or specific humidity (i.e., attributable fraction [AF]) in each county; c, d: the distribution of AF across months in the study period. The black lines represent the 95% confidence interval, which were calculated by 1000 Monte Carlo simulations.

Discussion

Using estimated reproduction numbers for 913 U.S. counties and controlling for temporal and spatial trends and other potential confounders, we assessed the associations of air temperature and SH with the transmission of SARS-CoV-2 and estimated the fractions of *Rt* attributable to temperature and SH. We found both lower air temperature and lower SH to be significantly associated with increased *Rt*. During

the study period, 5.10% of *Rt* was attributable to the deviation of temperature from its optimum value and 14.47% of *Rt* was attributable to the deviation of SH from its optimum value. Temperature and SH contributed more to transmission of SARS-CoV-2 in colder and drier counties and months than in warmer and more humid counties and months. In March (the coldest month of our study period of March-August), the AF for temperature was 11.00% and the AF for SH was 18.22% (Fig. 3). We can anticipate higher AFs during the colder and drier months of January and February.

Associations of lower temperature and lower humidity with increased COVID-19 outcomes have been reported by many previous studies. An early study in Wuhan, China reported that higher temperature and RH were associated with decreased COVID-19 deaths[21]. Many multicity analyses in China also supported such negative associations[9,11,12,22]. For example, using data of daily confirmed case counts from 30 provincial capital cities of China, Liu et al. found that lower temperature and lower absolute humidity were associated with higher COVID-19 case counts[11]. Later, with the rapid spread of COVID-19 around the world, studies in other countries emerged[23-25]. In the early stages of this pandemic in the U.S., a state-level study of daily COVID-19 case counts observed a declining trend of reported cases with increasing temperature up to 52 ºF[23]. Based on data from 166 countries worldwide, another study reported that a 1 °C increase in temperature and a 1% increase in RH were associated with a 3.08% and 0.85% reduction in daily new cases, respectively[25]. However, many of these earlier studies were limited by short study periods (e.g. 1-2 months), use of daily confirmed cases or deaths across countries for which there were varying reporting biases, failure to account for the time lag between observed weather conditions and when cases or deaths were recorded, or failure to account for time delays between infection acquisition and case confirmation[15].

By representing the transmissibility of SARS-CoV-2, the estimated daily reproduction number serves as a better outcome than daily case counts. While case counts are subject to the influence of reporting delay and underreporting, which vary across locations and are thus difficult to control, the reproduction number is a direct estimate of the transmission rate of SARS-CoV-2, quantifying the average number of infections

caused by one infection in the population. A small number of studies previously analyzed the association between temperature or humidity and reproduction number[16-18]. Wang et al. found that a 1 °C increase in temperature was associated with a reduction in the effective reproduction number by 0.023 in China and 0.020 in the U.S., and a 1% increase in RH was associated with a reduction in the effective reproduction number by 0.0078 in China and 0.0080 in the U.S.[16]. These associations are consistent with our findings but were not supported by two studies in China that examined the basic reproduction number: the first found no association between temperature and SARS-CoV-2 transmission[18]; the second found no association between absolute humidity and SARS-CoV-2 transmission[17]. However, these early studies were limited by short observation periods at the beginning of the pandemic, and they did not account for variations of testing capacity, reporting, human mobility, and population susceptibility in estimating SARS-CoV-2 transmissibility.

In our study, *Rt* was estimated using a dynamic metapopulation model informed by human mobility data. This mechanistic model accounted for unreported infections, reporting delays, and county-to-county movement. We explicitly estimated the population susceptibility in each county, and removed its influence in the calculation of *Rt*[26] (see Methods). Further, model estimated population susceptibility has been validated against independent seroprevalence studies[26]. Thus, our estimations account for spatial heterogeneity in population immunity.

Another strength of our study was adjustment for a wide range of demographic and socioeconomic factors in the main analysis, as well as for smoking and obesity, air pollution, and UV radiation in sensitivity analyses. We also thoroughly controlled for spatially and temporally heterogenous unmeasured confounders, such as implementation of and compliance with public health measures[26], by simultaneously controlling for temporal and spatial variations and including a random intercept to further account for unmeasured county-level confounding (see Methods). This approach accounted for substantial differences in the epidemic curves among counties (see Extended Data Fig. 2).

Our findings are supported by laboratory evidence on the stability of SARS-CoV-2 as a function of temperature and humidity. It has been reported that the virus' half-life in human nasal mucus and sputum is shorter under conditions of higher temperature and RH than under conditions of lower temperature and RH[2]. Similar findings were reported by other studies testing virus stability in virus transport medium[3], in aerosols, and on various surfaces[27]. Further, the SARS-CoV-2 half-life was found to be longer at lower temperatures, and at both 22 °C and 27 °C, the half-life decreased as RH increased from 40% to 65% but increased as RH increased from 65% to 85%[28]. This result is consistent with the non-linear relationship between SH and *Rt* observed in our study (Fig. 2b), in which there was an increasing trend of *Rt* from 9 to 15 g/kg of SH superimposed on the overall decreasing trend.

The associations between temperature and humidity and SARS-CoV-2 transmissibility may be mediated by airway antiviral defenses. Inhalation of cold and dry air can impair mucociliary clearance, a crucial mechanism for elimination of inhaled pathogens[29]. Further, during the colder winter months people spend more time indoors, which may facilitate virus transmission[30]. During these months, whether indoors or outdoors, people are exposed to less UV radiation from the sun and therefore may produce less vitamin D and other UV-induced mediators of immune function[31].

We found that SH contributes more to SARS-CoV-2 transmission than temperature, which is consistent with studies of influenza[32]. SH is more strongly associated with the observed seasonality of influenza in temperate regions than either temperature or RH[32-34]. In developed countries, such as the U.S., people spend approximately 90% of their time indoors[35], especially during winter[30]. Although indoor temperature is usually controlled, indoor humidity generally is not, and closely mirrors outdoor levels[36-38], perhaps explaining why ambient outdoor SH is more strongly associated with SARS-CoV-2 transmission than ambient outdoor temperature. However, it remains unclear whether SH is the causative modulator of SARS-CoV-2 transmission or is simply a useful indicator of the indoor environment and the combined effects of temperature and RH.

In the sensitivity analyses, after adjusting for daily UV radiation, the estimated AF for temperature decreased by about 30% (Extended Data Table 2), indicating that UV radiation acted as a confounder. This result is consistent with a recent study that found higher levels of solar UV radiation have a stronger association than temperature or humidity with a decreased growth rate (exponential increase in cases) of COVID-19[39]. In contrast, the fraction of *Rt* attributable to SH remained stable after adjusting for UV radiation. Although it is unclear why UV radiation would serve as a confounder for temperature, but not for SH, this result does suggest that SH is a more robust predictor than temperature.

Several limitations should be noted. First, this is an ecological rather than an individual-level study, thus making the study susceptible to the ecological fallacy. Second, due to the high correlation between temperature and SH, we were unable to explore whether the effects of temperature and humidity are independent. Third, our study period was restricted to March-August; if we had been able to include an entire year, including the colder months of November-February, our AF estimates for the entire study period would likely have been larger.

Conclusion

Our findings indicate that cold and dry weather are moderately associated with increased SARS-CoV-2 transmissibility in the U.S., with absolute humidity (i.e., SH) playing a greater role than temperature. More extensive public health interventions are needed to mitigate the increased transmissibility of SARS-CoV-2 in winter months.

Methods

Data collection

We extracted hourly air temperature and SH from the North America Land Data Assimilation System project[40], a near real-time dataset with a 0.125° × 0.125° grid resolution. We spatially and temporally averaged these data into daily county-level records. SH is the mass of water vapor in a unit mass of moist air (g/kg).

Other characteristics of each county, including geographic location, population density, demographic structure of the population, socioeconomic factors, intensive care unit (ICU) bed capacity, health risk factors, and long-term and short-term air pollution were collected from multiple sources. Geographic coordinates, population density, median household income, percent of people older than 60 years, percent black residents, percent Hispanic residents, percent owner-occupied housing, and percent residents aged 25 years and over without a high school diploma were collected from the U.S. Census Bureau[41]. The prevalence of smoking and obesity among adults in each county was obtained from the Robert Wood Johnson Foundation's 2020 County Health Rankings[42]. Total ICU beds in each county were derived from Kaiser Health News[43]. We extracted annual $PM_{2.5}$ concentrations in the U.S. from 2014 to 2018 from the $0.01° \times 0.01°$ grid resolution $PM_{2.5}$ estimation provided by the Atmospheric Composition Analysis Group[44], and calculated average $PM_{2.5}$ levels during this 5-year period for each county to represent long-term $PM_{2.5}$ exposure (Extended Data Fig. 3). Short-term air quality data during the study period, including daily mean $PM_{2.5}$ and daily maximum 8-hour O_3, were obtained from the United States Environmental Protection Agency[45]. Daily downward UV radiation at the surface was extracted from the European Centre for Medium-Range Weather Forecasts ERA5 climate reanalysis[46], with data available before August 2020.

Estimation of reproduction number

We estimated the daily reproduction number (*Rt*) in all 3,142 U.S. counties using a dynamic metapopulation model informed by human mobility data[26]. *Rt* is the mean number of new infections caused by a single infected person, given the public health measures in place, in a population in which everyone is assumed to be susceptible. In the metapopulation model, two types of movement were considered: daily work commuting and random movement. The transmission dynamics are depicted by a set of ordinary differential equations[26].

We explicitly simulated reported and unreported infections, for which separate transmission rates are defined, and allowed transmission rates and ascertainment rates to vary across different counties. To infer

key epidemiological parameters, we fit the transmission model to county-level daily cases and deaths reported from March 15, 2020 to August 31, 2020. The estimated reproduction number is computed as $R_t = \beta D[\alpha + (1-\alpha)\mu]$, where β is county-specific transmission rate, μ is the relative transmissibility of unreported infections, α is the county-specific ascertainment rate, and D is the average duration of infectiousness. To avoid possible inaccurate estimation for counties with few cases, we inferred *Rt* in the 913 U.S. counties with at least 400 cumulative confirmed cases as of August 31, 2020 (Fig. 1). Details of the model fitting and *Rt* estimation are reported elsewhere[26].

Statistical analysis

All statistical analyses were conducted with R software (version 3.6.1) using the *mgcv* and *dlnm* package.

Exposure-response curves

Given the potential non-linear and temporally delayed effect of air temperature or SH, a distributed lag non-linear model (DLNM)[47] combined with generalized additive mixed models (GAMM) was applied to estimate the associations of daily mean temperature or daily mean SH with SARS-CoV-2 *Rt*. Because of the high correlation between air temperature and SH (r = 0.80, Extended Data Table 3), we analyzed these two variables separately. The full model can be expressed as:

$$\log(E(R_{i,t})) = \alpha + te(s(\text{latitude}_i, \text{longitude}_i, \text{k=30}), s(\text{time}_t, \text{k=30})) + \text{cb.temperature (or cb.SH)} + \beta_1(\text{population density}_i) + \beta_2(\text{percent black residents}_i) + \beta_3(\text{percent Hispanic residents}_i) + \beta_4(\text{precent people older than 60 years}_i) + \beta_5(\text{median household income}_i) + \beta_6(\text{percent owner-occupied housing}_i) + \beta_7(\text{percent residents older than 25 years without a high school diploma}_i) + \beta_8(\text{number of ICU beds per 10,000 people}_i) + u_i$$

where $E(R_{i,t})$ refers to the expected *Rt* in county i on day t, and α is the intercept. The time trend was controlled by a flexible natural cubic spline over the range of study dates with a maximum of 30 knots; a thin plate spline with a maximum of 30 knots was used to control the coordinates of the centroid of each county. Due to the unique pattern of the non-linear time trend of *Rt* in each county (Extended Data Fig.

2), we constructed tensor product smooths (*te*) of the splines of geographical coordinates and time, to better control for the temporal and spatial variations. Cb.temperature or cb.SH is a cross-basis term for the mean air temperature or mean SH. We modeled exposure-response associations using a natural cubic spline with 3 degrees of freedom (df), and modeled the lag-response association using a natural cubic spline with an intercept and 3 df with a maximum lag of 13 days. We adjusted for county-level characteristics, including population density, percent black residents, percent Hispanic residents, percent people older than 60 years, median household income, percent owner-occupied housing, percent residents older than 25 years without a high school diploma, and number of ICU beds per 10,000 people. The random effect of county (u_i) was considered in the model to further control for unmeasured county-level confounding. To obtain more precise estimates, we excluded from the analysis days during which *Rt* was less than 0.2.

Based on the estimated exposure-response curves, between the 1st and the 99th percentiles of the distribution of air temperature and SH, we determined the value of exposure associated with the lowest *Rt* to be the optimum temperature or the optimum SH, respectively. The natural cubic spline functions of the exposure-response relationship were then re-centered with the optimum temperature and SH as reference values. We report the cumulative relative risk of *Rt* associated with daily temperature or SH exposure in the previous two weeks (0 to 13 lag days) as the percent changes of *Rt* when comparing the daily exposure with the optimum reference values (i.e., the cumulative relative risk of *Rt* equals one and the percent change of *Rt* equals zero when the temperature or SH exposure is at its optimum value).

Attribution of *Rt* to temperature or SH

We used the optimum value of temperature or SH as the reference value for calculating the fraction of *Rt* attributable to temperature or SH; i.e., the attributable fraction (AF). For these calculations, we assumed that the associations of temperature and SH with *Rt* were consistent across the counties. For each day in each county, based on the cumulative lagged effect (cumulative relative risk) corresponding to the temperature or SH of that day, we calculated the attributable *Rt* in the current and next 13 days, using a

previously established method[48]. Specifically, in a given county, the *Rt* attributable to a temperature or a SH (x_t) for a given day *t* was defined as the attributable absolute excess of *Rt* ($AE_{x,t}$, the excess reproduction number on day *t* attributable to the deviation of temperature or SH from the optimum value) and the attributable fraction of *Rt* ($AF_{x,}$, the fraction of *Rt* attributable to the deviation of temperature or SH from the optimum value), each accumulated over the current and next 13 days. The formula can be expressed as

$AF_{x,t} = 1 - \exp(-\sum_{l=0}^{13} \beta_{x_t,l})$ and $AE_{x,t} = AF_{x,t} \times \sum_{l=0}^{13} \frac{n_{t+1}}{13+1}$, where n_t is the *Rt* on day *t*, and $\sum_{l=0}^{13} \beta_{x_t,l}$ is the overall cumulative log-relative risk for exposure x_t on day *t* obtained by the exposure-response curves re-centered on the optimum values. Then, the total absolute excess of *Rt* attributable to temperature or SH in each county was calculated by summing the absolute excesses of all days during the study period, and the attributable fraction was calculated by dividing the total absolute excess of *Rt* for the county by the sum of the *Rt* of all days during the study period for the county. The attributable fraction for the 913 counties combined was calculated in a similar manner at the national level. We derived the 95% empirical confidence intervals (95% eCI) for the attributable absolute excess and attributable fraction by 1000 Monte Carlo simulations[48]. We also calculated the attributable fractions by month in the study period.

Sensitivity analyses

We conducted several sensitivity analyses to test the robustness of our results: a) the lag dimension was redefined using a natural cubic spline and three equally placed internal knots in the log scale; b) an alternative four df was used in the cross-basis term for air temperature or SH in the exposure-response function; c) time trend and geographical coordinates were controlled by a thin plate spline, instead of the tensor product smooth; d) all demographic and socioeconomic variables were excluded from the model; e) adjustment for the prevalence of smoking and obesity among adults was included in the model; f) additional adjustment was made for the average $PM_{2.5}$ concentration in each county during 2014-2018[49];

g) additional adjustment was made for daily mean $PM_{2.5}$, daily maximum 8-hour O_3, and daily downward UV radiation at the surface. For daily covariates with available data in only some of the counties or study period, the results of sensitivity analyses were compared to the main model re-run on the same partial dataset.

Acknowledgements

Y.M. received funding from the China Scholarship Council (201906320022). S.P. and J.S. acknowledged funding from the National Institutes of Health (GM110748) and the National Science Foundation (DMS-2027369), as well as a gift from the Morris-Singer Foundation. R.D. received funding from the High Tide Foundation.

Author contributions

K.C. conceived of and supervised the conduct of this study and edited the manuscript. J.S. and R.D. contributed to the study design, interpretation of results, and manuscript revision. S.P. estimated the reproduction number and contributed to the writing. Y.M. conducted formal analyses and drafted the manuscript. All authors reviewed and approved the final version of this manuscript.

Competing interests

J.S. and Columbia University disclose partial ownership of SK Analytics. J.S. discloses consulting for BNI. All other authors declare no competing interests.

Data availability

Estimates of county-level reproduction number are available at https://github.com/shaman-lab/Counterfactual

The data sets used in the study are publicly available from the following locations:

Hourly air temperature and specific humidity data:
https://disc.gsfc.nasa.gov/datasets/NLDAS_FORA0125_H_002/summary?keywords=NLDAS

Population density, median household income, percent of people older than 60 years, percent black residents, percent Hispanic residents, percent owner-occupied housing, and percent residents aged 25 years and over without a high school diploma: https://www.census.gov/data/tables.html

U.S. county boundary: https://www.census.gov/geographies/mapping-files/time-series/geo/cartographic-boundary.html

Prevalence of smoking and obesity among adults in each county: https://www.countyhealthrankings.org/explore-health-rankings/rankings-data-documentation

Total ICU beds in each county: https://khn.org/news/as-coronavirus-spreads-widely-millions-of-older-americans-live-in-counties-with-no-icu-beds/

Annual $PM_{2.5}$ concentrations in the U.S. from 2014 to 2018: http://fizz.phys.dal.ca/~atmos/martin/?page_id=140

Short-term daily mean $PM_{2.5}$ and daily maximum 8-hour O_3: https://www.epa.gov/outdoor-air-quality-data/download-daily-data

Daily downward UV radiation at the surface: https://cds.climate.copernicus.eu/cdsapp#!/dataset/reanalysis-era5-single-levels?tab=overview

Code availability

R code for this analysis will be available at https://github.com/CHENlab-Yale/COVID-Climate

Reference

1. Dong, E., Du, H. & Gardner, L. An interactive web-based dashboard to track COVID-19 in real time. *The Lancet Infectious Diseases* **20**, 533-534 (2020).
2. Matson, M.J., *et al.* Effect of environmental conditions on SARS-CoV-2 stability in human nasal mucus and sputum. *Emerg Infect Dis* **26**, 2276–2278 (2020).
3. Chin, A.W.H., *et al.* Stability of SARS-CoV-2 in different environmental conditions. *Lancet Microbe* **1**, e10 (2020).
4. Sooryanarain, H. & Elankumaran, S. Environmental role in influenza virus outbreaks. *Annu Rev Anim Biosci* **3**, 347-373 (2015).
5. Tan, J., *et al.* An initial investigation of the association between the SARS outbreak and weather: with the view of the environmental temperature and its variation. *J Epidemiol Community Health* **59**, 186-192 (2005).
6. Abdul-Rasool, S. & Fielding, B.C. Understanding human coronavirus HCoV-NL63. *Open Virol J* **4**, 76-84 (2010).
7. Esper, F., Weibel, C., Ferguson, D., Landry, M.L. & Kahn, J.S. Coronavirus HKU1 infection in the United States. *Emerg Infect Dis* **12**, 775-779 (2006).
8. Carlson, C.J., Gomez, A.C.R., Bansal, S. & Ryan, S.J. Misconceptions about weather and seasonality must not misguide COVID-19 response. *Nat Commun* **11**, 4312 (2020).
9. Shi, P., *et al.* Impact of temperature on the dynamics of the COVID-19 outbreak in China. *Sci Total Environ* **728**, 138890 (2020).
10. Prata, D.N., Rodrigues, W. & Bermejo, P.H. Temperature significantly changes COVID-19 transmission in (sub)tropical cities of Brazil. *Sci Total Environ* **729**, 138862 (2020).
11. Liu, J., *et al.* Impact of meteorological factors on the COVID-19 transmission: A multi-city study in China. *Sci Total Environ* **726**, 138513 (2020).
12. Xie, J. & Zhu, Y. Association between ambient temperature and COVID-19 infection in 122 cities from China. *Sci Total Environ* **724**, 138201 (2020).
13. Pani, S.K., Lin, N.H. & RavindraBabu, S. Association of COVID-19 pandemic with meteorological parameters over Singapore. *Sci Total Environ* **740**, 140112 (2020).
14. Li, R., *et al.* Substantial undocumented infection facilitates the rapid dissemination of novel coronavirus (SARS-CoV-2). *Science* **368**, 489-493 (2020).
15. Smit, A.J., *et al.* Winter is coming: a southern hemisphere perspective of the environmental drivers of SARS-CoV-2 and the potential seasonality of COVID-19. *Int J Environ Res Public Health* **17**, 5634 (2020).
16. Wang, J., *et al.* High temperature and high humidity reduce the transmission of COVID-19. *SSRN* (2020).
17. Luo, W., *et al.* The role of absolute humidity on transmission rates of the COVID-19 outbreak. *medRxiv*, 2020.2002.2012.20022467 (2020).
18. Yao, Y., *et al.* No association of COVID-19 transmission with temperature or UV radiation in Chinese cities. *Eur Respir J* **55**, 2000517 (2020).
19. Gostic, K.M., *et al.* Practical considerations for measuring the effective reproductive number, Rt. *medRxiv* (2020).
20. Baker, R.E., Yang, W., Vecchi, G.A., Metcalf, C.J.E. & Grenfell, B.T. Susceptible supply limits the role of climate in the early SARS-CoV-2 pandemic. *Science* **369**, 315-319 (2020).
21. Peak, C.M., *et al.* Individual quarantine versus active monitoring of contacts for the mitigation of COVID-19: a modelling study. *The Lancet Infectious Diseases* (2020).
22. Qi, H., *et al.* COVID-19 transmission in Mainland China is associated with temperature and humidity: A time-series analysis. *Sci Total Environ* **728**, 138778 (2020).
23. Sehra, S.T., Salciccioli, J.D., Wiebe, D.J., Fundin, S. & Baker, J.F. Maximum daily temperature, precipitation, ultra-violet light and rates of transmission of SARS-Cov-2 in the United States. *Clin Infect Dis* (2020).

24. Runkle, J.D., *et al.* Short-term effects of specific humidity and temperature on COVID-19 morbidity in select US cities. *Sci Total Environ* **740**, 140093 (2020).
25. Wu, Y., *et al.* Effects of temperature and humidity on the daily new cases and new deaths of COVID-19 in 166 countries. *Sci Total Environ* **729**, 139051 (2020).
26. Pei, S., Kandula, S. & Shaman, J. Differential effects of intervention timing on COVID-19 spread in the United States. *Science Advances*, eabd6370 (2020).
27. van Doremalen, N., *et al.* Aerosol and surface stability of SARS-CoV-2 as compared with SARS-CoV-1. *N Engl J Med* **382**, 1564-1567 (2020).
28. Morris, D.H., *et al.* The effect of temperature and humidity on the stability of SARS-CoV-2 and other enveloped viruses. *bioRxiv*, 2020.2010.2016.341883 (2020).
29. Moriyama, M., Hugentobler, W.J. & Iwasaki, A. Seasonality of respiratory viral infections. *Annual Review of Virology* **7**, 83-101 (2020).
30. Shaman, J. & Galanti, M. Will SARS-CoV-2 become endemic? *Science*, eabe5960 (2020).
31. Hart, P.H., Gorman, S. & Finlay-Jones, J.J. Modulation of the immune system by UV radiation: more than just the effects of vitamin D? *Nature Reviews Immunology* **11**, 584-596 (2011).
32. Shaman, J. & Kohn, M. Absolute humidity modulates influenza survival, transmission, and seasonality. *Proc Natl Acad Sci U S A* **106**, 3243-3248 (2009).
33. Shaman, J., Goldstein, E. & Lipsitch, M. Absolute humidity and pandemic versus epidemic influenza. *Am J Epidemiol* **173**, 127-135 (2011).
34. Shaman, J., Pitzer, V.E., Viboud, C., Grenfell, B.T. & Lipsitch, M. Absolute humidity and the seasonal onset of influenza in the continental United States. *PLoS Biol* **8**, e1000316 (2010).
35. Klepeis, N.E., *et al.* The National Human Activity Pattern Survey (NHAPS): a resource for assessing exposure to environmental pollutants. *J Expo Anal Environ Epidemiol* **11**, 231-252 (2001).
36. Shaman, J., Kandula, S., Yang, W. & Karspeck, A. The use of ambient humidity conditions to improve influenza forecast. *PLoS Comput Biol* **13**, e1005844 (2017).
37. Quinn, A. & Shaman, J. Indoor temperature and humidity in New York City apartments during winter. *Science of The Total Environment* **583**, 29-35 (2017).
38. Nguyen, J.L., Schwartz, J. & Dockery, D.W. The relationship between indoor and outdoor temperature, apparent temperature, relative humidity, and absolute humidity. *Indoor Air* **24**, 103-112 (2014).
39. Merow, C. & Urban, M.C. Seasonality and uncertainty in global COVID-19 growth rates. *Proceedings of the National Academy of Sciences*, 202008590 (2020).
40. Cosgrove, B.A., *et al.* Real-time and retrospective forcing in the North American Land Data Assimilation System (NLDAS) project. *Journal of Geophysical Research: Atmospheres* **108**(2003).
41. United States Census Bureau. Tables. (2020).
42. Robert Wood Johnson Foundation. 2020 County Health Rankings. (2020).
43. Kaiser Health News. Millions Of Older Americans Live In Counties With No ICU Beds As Pandemic Intensifies. (2020).
44. van Donkelaar, A., Martin, R.V., Li, C. & Burnett, R.T. Regional estimates of chemical composition of fine particulate matter using a combined geoscience-statistical method with information from satellites, models, and monitors. *Environ Sci Technol* **53**, 2595-2611 (2019).
45. United States Environmental Protection Agency. Outdoor Air Quality Data. (2020).
46. European Centre for Medium-Range Weather Forecasts. ERA5 hourly data on single levels from 1979 to present. (2018).
47. Gasparrini, A., Armstrong, B. & Kenward, M.G. Distributed lag non-linear models. *Stat Med* **29**, 2224-2234 (2010).
48. Gasparrini, A. & Leone, M. Attributable risk from distributed lag models. *BMC Med Res Methodol* **14**, 55 (2014).

488 49. Wu, X., Nethery, R.C., Sabath, M.B., Braun, D. & Dominici, F. Air pollution and COVID-19
489 mortality in the United States: Strengths and limitations of an ecological regression analysis. *Sci*
490 *Adv* **6**(2020).

491

492

Extended Data Fig. 1. Lag-response relationships of air temperature (°C) or specific humidity (g/kg) with reproduction number (*Rt*).

These curves are computed for the 10th percentile of air temperature and specific humidity vs. the optimum values on different lag days; the grey areas display the 95% confidence interval. The effect estimates show a decreasing trend in the lag dimension, diminishing to a small non-significant effect on lag day 13.

Extended Data Fig. 2. Daily *Rt* from March 15 to August 31 in the largest county in each state.

Black dots represent the daily value of reproduction number (*Rt*) in the largest county in each U.S. state. Blue lines show the trend of *Rt* through time, fitted by local polynomial regression; the light blue areas display the 95% confidence interval.

Extended Data Fig. 3. Distribution of average $PM_{2.5}$ concentration during 2014-2018.

This map displays the county-level average $PM_{2.5}$ concentration during 2014-2018, extracted from the $PM_{2.5}$ estimation provided by Atmospheric Composition Analysis Group.