Effect of Atmospheric Temperature and Relative Humidity on the Incidence of Covid-19 in the United States and Inferences about Infection Mitigation in Climate-Controlled Buildings

Benoit Cushman-Roisin¹, Vikrant S. Vaze, and Alexis R. Abramson
Thayer School of Engineering, Dartmouth College
Hanover, New Hampshire

12 July 2020

ABSTRACT

We investigated whether Covid-19 incidence data across the United States in early May 2020 might possibly reveal correlations with climatic variables during the preceding months, particularly temperature and relative humidity. Aside from the already well established and dominant correlations with population density, age and comorbidities, we found weaker but nonetheless statistically significant correlations of Covid-19 cases and fatalities with outdoor temperature and relative humidity. Both correlations are negative, implying that an increase of either temperature or relative humidity corresponds to lower numbers of Covid-19 cases and deaths. The correlations are robust in the sense that they persist through univariate and multivariate correlation analyses.

From this finding, we infer for climate-controlled buildings that some non-negligible degree of reduction of Covid-19 transmission could be achieved by adjustments to their HVAC systems. In particular, we note that a simultaneous increase of 0.5°C in temperature and an increase from 50% to 70% in relative humidity has the potential to reduce SARS-CoV-2 virus transmission by at least 3% and likely by 39% while keeping the occupants comfortable and productive, and avoiding damage to the building’s envelope. Similar adjustments could also be considered in public transportation vehicles equipped with climate-control.

We caution, however, that an adjustment to settings in a climate-controlled building does not imply that climate-control is necessarily an advantage. There is evidence that natural ventilation may be preferable to avoid recirculation of the virus inside the building. We only recommend that, if the building (or vehicle) must be climate-controlled, then it is preferable to raise slightly its temperature and relative humidity settings. Our recommendation does not substitute for the much more important directives issued by the medical profession such as social distancing, wearing of a mask, and frequent hand washing.

1. Introduction

Covid-19, the disease caused by the novel corona virus SARS-CoV-2, progressed rapidly in the early months of 2020 across the world, including the United States where it has been responsible for numerous deaths. Although it was quickly realized that the disease is more severely affecting people older than 65, especially those with pre-existing chronic conditions (comorbidities), people living in densely populated areas, and those commuting by means of public transportation, it is noted that after removal of the aforementioned

¹ Benoit.Cushman-Roisin@dartmouth.edu
factors, the incidence of the disease remains far from uniform. For example, the number of deaths per 100,000 persons differs by more than a factor 2 (as of early May 2020) between Atlanta and Omaha despite similar demographics (respective population density of 3,549 and 3,517 people per square mile, 4.12% and 4.45% of people 65 or older with at least two chronic conditions, and household median income of $57,597 and $56,406). This prompts us to ask what other factors may be at play. In particular, could environmental factors such as temperature and humidity play a role?

Such a question had already been addressed before the emergence of SARS-CoV-2 for a variety of viruses in a series of studies dating back to the 1940s. Memarzadeh (2011) reviewed more than 120 studies and concluded, for all viruses combined, that “there is no conclusive evidence suggesting a defined minimum or maximum relative humidity (RH) that reduces viral survival to the point where a virus is less able to survive or is affected in its ability to cause an infection” “but there is pervasive evidence in the literature that the survival of viruses and other infectious agents depends partially on levels of RH.” There are reports showing “opposing conditions for transmission of viruses ranging from low RH and high RH with temperature a secondary factor.” Specifically, transmission of respiratory syncytial viruses (RSV) exhibit different behaviors based on latitude and seasonality. A study by Songer (1967) showed that sensitivity to relative humidity appears to be a characteristic of the virus, with some RSV incidence peaking at both low and high temperatures, and with RSV incidence positively correlated with high relative humidity in some studies and negatively in others. Thompson et al. (2003) found that RSV activity is uniform throughout the year in areas with persistently high temperatures and high humidity as well as in areas where temperatures remained consistently low, but in temperate climates where seasonality is pronounced, RSV activity is maximum in winter, correlating with lower temperatures. This gives hints as to possible behaviors of the novel SARS-CoV-2 under various temperatures and relative humidity levels.

A recent literature review by Moriyama et al. (2020), which considers the novel SARS-CoV-2, remarks that relative humidity affects all transmission pathways but has the most pronounced effect on airborne transmission and affects all infectious droplets with respiratory viruses independent of their source. The authors express the hope that “The precise relationship between temperature, humidity, and COVID-19 will become more evident as the Northern Hemisphere reaches the summer months.” In the interim, based on pre-Covid-19 studies, it recommends that indoor relative humidity be maintained between 40% and 60%. This conclusion, however, that humidity affects Covid-19 transmission is not supported by Luo et al. (2020) who explored a possible relation between absolute humidity and transmission of Covid-19 across provinces of China in January 2020 only to conclude from a set of 37 observations that “changes in weather alone (i.e., increase of temperature and humidity as spring and summer months arrive in the North Hemisphere) will not necessarily lead to declines in COVID-19 case counts.”

In a larger study, Jüni et al. (2020) asked similar questions by comparing 144 geopolitical areas (such as entire countries, provinces and states) across the world with a reported total of 375,609 Covid-19 cases as of 20 March 2020. They correlated the number of newly reported cases during a week in late March with weather characteristics 14 days earlier, accounting for a time lag between exposure and symptoms. A univariate analysis detected a weak negative correlation between epidemic growth and relative humidity (a 10% increase in RH is estimated to reduce the epidemic growth by 9%, with 95% confidence interval of 4% to 15%) and no significant correlation with temperature. In a multivariate analysis, they found that the correlation with relative humidity “attenuated and became mostly nonsignificant.” This
study, however, has shortcomings: The 144 areas were selected in a non-random and non-exhaustive process introducing the possibility of bias, the disease data reflected an incremental approach with only a single week of exposure to the virus, and the weather data were taken for the capital city of each geopolitical area which for the larger areas is not necessarily representative of the entire territory.

In view of such disparate conclusions, the American Society of Heating, Refrigerating and Air-Conditioning Engineers (ASHRAE) – an organization whose recommendations and standards for indoor ventilation, temperature and relative humidity are highly regarded across the world – issued a position document on April 14th, 2020 (ASHRAE, 2020) not recommending “indoor temperature and humidity set points for the purpose of controlling infectious aerosol transmission,” limiting itself to advise that “designs that achieve higher ventilation rates will reduce risks.”

Clearly, what is needed is a study of Covid-19 cases and deaths over a longer and cumulative duration and with non-biased geographical coverage divided in much smaller areas.

2. Data Set

In order (1) to avoid any bias in selecting cities or geographical areas but with specific interest in the United States, (2) to have much finer spatial resolution than previous investigations, and (3) to encompass a longer-term and cumulative set of Covid-19 cases, we obtained numbers of Covid-19 reported cases and deaths for the 306 Hospital Referral Regions (HRRs) compiled by The Dartmouth Institute for Health Policy & Clinical Practice² as of 4 May 2020. The cumulative population in these 306 HRRs is 320,963,092 or 97% of the current U.S. population. In addition to these Covid-19 case and death numbers, we gathered for every HRR a series of demographic and climatic data. No data on confinement and prophylactic measures were included.

The demographic data consist of population in the HRR, population density, median age, household median income, and percentage of people at elevated risk (persons 65-years old or older and with at least two existing chronic conditions). The numbers were obtained as follows: Population from the 3,432 so-called Hospital Service Areas (as compiled by the Dartmouth Atlas Project³) aggregated to the 306 Hospital Referral Regions; population density from The New York Times as made available to and posted by the Dartmouth Atlas⁴; median age and 2017 household median income from city-data.com; and percentage of people at elevated risk from the Hospital Service Areas aggregated to the HRRs (weighted by population) as provided by the Dartmouth Atlas Project⁵.

The climatic data cover the 3-month period of February-March-April 2020 considered as the months during which most, if not all, Covid-19 infections took place in the United States leading to the cumulative data of May 4th, 2020, recalling that the first Covid-19 fatalities in the United States occurred in mid-February at the Life Care Center of Kirkland in the State of Washington, presumably resulting from

³ https://www.dartmouthatlas.org/covid-19/², because these population numbers are more recent than those provided at https://github.com/Dartmouth-DAC/covid-19-hrr-mapping/tree/master/HRR-mapping/sample-output
⁵ https://www.dartmouthatlas.org/covid-19/
infections in early February. The data consist of monthly minimum and maximum temperature (in degree Celsius), monthly average relative humidity (in %), and monthly precipitation (in mm). The numbers were obtained primarily from the Weather Atlas\(^6\) with missing data obtained from weather-and-climate.com, city-data.com, or The Southeast Regional Climate Center\(^7\).

For the sake of comparison, the May 4\(^{th}\) 2020 numbers of Covid-19 cases and deaths were pro-rated to a reference population of 100,000, the temperatures and relative humidity levels were averaged over the aforementioned period of three months, and the precipitation amounts were summed over the same period of three months. The compilation resulted in a database with 19 entries and 5 calculated values for each of the 306 Hospital Referral Regions of the United States. Table 1 lists the extreme, median and average values to give an idea of the spread in values in the database.

<table>
<thead>
<tr>
<th>Variable</th>
<th>Minimum</th>
<th>Maximum</th>
<th>Median</th>
<th>Average</th>
<th>Unit</th>
</tr>
</thead>
<tbody>
<tr>
<td>Population density</td>
<td>175 (Anchorage AK)</td>
<td>28,211 (Bronx NY)</td>
<td>2,738</td>
<td>3,649</td>
<td>People per square mile</td>
</tr>
<tr>
<td>Median age</td>
<td>23.9 (New Brunswick NJ)</td>
<td>74.1 (Sun City AZ)</td>
<td>35.2</td>
<td>35.3</td>
<td>years</td>
</tr>
<tr>
<td>Percentage of people 65 or older and with at least 2 chronic conditions</td>
<td>2.34% (Provo UT)</td>
<td>15.12% (Ocala FL)</td>
<td>5.39%</td>
<td>5.61%</td>
<td>%</td>
</tr>
<tr>
<td>Household median income</td>
<td>24,540 (Camden NJ)</td>
<td>181,766 (Hinsdale IL)</td>
<td>47,556</td>
<td>51,041</td>
<td>$/household</td>
</tr>
<tr>
<td>Average temperature</td>
<td>–3.12 (Grand Forks ND)</td>
<td>23.63 (Honolulu HI)</td>
<td>8.04</td>
<td>8.52</td>
<td>°C</td>
</tr>
<tr>
<td>Total precipitation</td>
<td>24.1 (El Paso TX)</td>
<td>419.0 (Tupelo MS)</td>
<td>221.0</td>
<td>221.6</td>
<td>mm</td>
</tr>
<tr>
<td>Average relative humidity</td>
<td>32.6 (Las Vegas NV)</td>
<td>80.0 (Olympia WA)</td>
<td>68.6</td>
<td>66.9</td>
<td>%</td>
</tr>
<tr>
<td>Covid-19 cases per 100,000 people</td>
<td>7.73 (Chico CA)</td>
<td>2,839.6 (White Plains NY)</td>
<td>143.2</td>
<td>283.5</td>
<td>Number of persons</td>
</tr>
<tr>
<td>Covid-19 deaths per 100,000 people</td>
<td>0.00 (Idaho Falls ID)</td>
<td>154.86 (New York NY)</td>
<td>5.20</td>
<td>14.8</td>
<td>Number of persons</td>
</tr>
</tbody>
</table>

Table 1. Minimum, maximum, median and average values of variables across the data set. Large differences between median and average values is symptomatic of distributions with long tails.

Needless to say, Covid-19 numbers are constantly evolving, and there is therefore some arbitrariness in selecting a particular time for carrying out an analysis. It stands to reason that, if trends exist because of underlying biophysics, these trends should be manifest at any stage of the progression of the pandemic once numbers are sufficiently large to establish statistical significance. The selected time for the analysis then becomes irrelevant. One particular significance of the chosen time period for this study is that the Federal social distancing guidelines in the United States expired on April 30\(^{th}\) and led to different states, cities and regions reopening at substantially different rates throughout the months of May and June. Therefore, we believe that the data after early May might be additionally noisy. This is confirmed by our

---

\(^6\) [https://www.weather-atlas.com/](https://www.weather-atlas.com/)

\(^7\) [https://sercc.com/climateinfo/historical/avgrh.html](https://sercc.com/climateinfo/historical/avgrh.html)
analysis demonstrating a dilution of the statistical significance of our correlations and regressions when we tried to include the month of May in our analysis (Section 3c).

3. Data Analysis and Results

Before starting the correlation and regression calculations, we first considered the characteristics of the distributions at hand and noted that the population density and the Covid-19 case and death numbers range from near zero values to very large positive values, exhibiting long tails. Therefore, natural logarithms of these three variables, respectively denoted LPD, LC100 and LD100\(^8\), were used for all subsequent correlation and regression analyses. The percentages of population at high risk, household median incomes, total precipitation, and relative humidity, too, were bounded below by zero. But, the bulk of their values was sufficiently away from zero and without long tails. Hence, these variables were used directly (i.e. without any transformation) for the subsequent correlation and regression analyses.

3a. Univariate analyses

We then proceeded with a series of univariate inquiries into the correlations between Covid-19 cases and deaths with separate demographic and climatic variables.

Not surprisingly, Covid-19 cases and deaths were found to be strongly correlated with population density. Correlation coefficient between the log of cases (LC100 variable) and log of population density (LPD) is 0.4207 with a 95% confidence interval of [0.3239, 0.5088], while that between the log of deaths (LD100) and log of population density (LPD) is 0.4368 with a 95% confidence interval of [0.3414, 0.5233]. We also noted a small positive correlation of median household income with cases (8%) and with deaths (3%), but neither of these was statistically significant (\(p\)-values of 15% and 59%, respectively).

As we could have anticipated, we found no statistically significant correlation between Covid-19 cases and the percentage of people at elevated risk (persons 65-years old or older and with at least two existing chronic conditions) since, unlike progression of the disease within the body, transmission between persons does not discriminate with respect to a person’s age or health. Regression with median age was found to have low statistical significance. The conclusion is that median age is a far less useful indicator than the percentage of people at elevated risk. Put another way, median age is a poor proxy for risk.

Our primary aim is to evaluate possible correlations with climatic data. To give an idea of climatic trends amidst larger variations due to other factors, Figures 1 and 2 display scatter plots of LD100 versus temperature (T) and of LC100 versus relative humidity (RH). Both trend lines are negative.

---

\(^8\) To avoid a singularity, the zero-death values in Idaho Falls ID, Grand Forks ND and Rapid City SD were replaced by the finite but low value of 1 before taking the natural logarithm. Similarly, to avoid occasionally large negative values for the logarithms corresponding to death values between 0 and 1 per 100,000 people (there were 24 such locations in our database), we rounded up these values to 1 per 100,000 people, yielding zero as the lowest logarithmic values in the set.
Figure 1. Scatter plot of the natural logarithm of the number of deaths per 100,000 people (LD100) versus temperature (T). The trend is negative, meaning that, everything else being equal, deaths due to Covid-19 decrease with increased temperature.

Figure 2. Scatter plot of the natural logarithm of the number of cases per 100,000 people (LC100) versus relative humidity (RH). Here again, the trend is negative, meaning that, everything else being equal, cases due to Covid-19 decrease with increased relative humidity. It was also found that the trend line was only very weakly sensitive to the cluster of low RH values (nine values between 32 and 41%).

The logarithms of both Covid-19 cases (LC100) and deaths (LD100) were correlated separately with temperature (T, in °C), precipitation (P, in mm) and relative humidity (RH, expressed as a %). Table 2 recapitulates the findings.
Although small in magnitude, all correlations are statistically significant with all $p$-values lower than 0.05, implying that the hypothesis that the population correlation is zero may be rejected at the 5% significance level, or, put the other way, the correlation is statistically significant at a greater than 95% confidence level.

3b. Multivariate analyses

To better untangle the role of climatic factors amidst dominant demographic factors, it is essential to perform multivariate analyses where Covid-19 data are regressed with one or several climatic factors and one or several demographic factors. Since virus transmission and infection do not discriminate among people, there is no expectation of a correlation between Covid-19 cases and the percentage of people at elevated risk, and indeed we found none. Thus, the $O65C2P$ variable (percentage of people over 65 years of age and with at least two chronic conditions) was only used in regressions of death data.

Table 3 below recapitulates the various regression analyses that we performed, with up to four independent variables. As a rule, statistical significance of the coefficients for demographic and climatic data is weaker with Covid-19 cases than with Covid-19 deaths, leading us to consider the case numbers as less reliable than the death numbers. This is presumably because diagnosed case data are marred by a lack of systematic testing or of control sampling of the general population making the number of cases an unreliable proxy for the incidence of the disease across the population. So, from this point forward, our attention is restricted to the question of how Covid-19 deaths depend on demographic and climatic factors.
What stands clearest (in terms of the lowest \( p \)-values and thus highest statistical significance levels) is the positive dependence of deaths on population density (\( LPD \) variable). Next in order of significance is the positive dependence on the percentage of people with elevated risk (\( O65C2P \) variable). Neither of these findings is surprising. We then note that correlations of deaths with temperature and/or relative humidity have relatively weak but consistently negative coefficients and are statistically significant (1 – \( p \) value at around 95% or higher levels).

We retain as most valuable the quadruple regression of Covid-19 deaths with population density, percentage of people at elevated risk, temperature and relative humidity. This is because it removes the leading dependencies on the two primary demographic factors and includes the two environmental variables that one can control in an indoor environment. Based on the slope factors and standard errors listed in Table 3 and the intercept number (not listed in Table 3), we can express this quadruple regression mathematically as:

<table>
<thead>
<tr>
<th>Covid-19 incidence per 100,000 people</th>
<th>simultaneously regressed with</th>
<th>independent variable</th>
<th>Slope factor</th>
<th>Standard error</th>
<th>( t )-value</th>
<th>( p )-value</th>
<th>Goodness of fit (adjusted ( R^2 ))</th>
</tr>
</thead>
<tbody>
<tr>
<td>Log of deaths</td>
<td>temperature, and relative humidity</td>
<td>( T )</td>
<td>-0.02420</td>
<td>0.01189</td>
<td>-2.035</td>
<td>0.0427</td>
<td>0.0252</td>
</tr>
<tr>
<td></td>
<td></td>
<td>( RH )</td>
<td>-0.02284</td>
<td>0.009809</td>
<td>-2.329</td>
<td>0.0205</td>
<td></td>
</tr>
<tr>
<td></td>
<td>population density, % people at elevated risk, and temperature</td>
<td>( LPD )</td>
<td>0.8129</td>
<td>0.09410</td>
<td>8.639</td>
<td>3.36 \times 10^{-16}</td>
<td>0.2098</td>
</tr>
<tr>
<td></td>
<td></td>
<td>( O65C2P )</td>
<td>12.871</td>
<td>4.3113</td>
<td>2.985</td>
<td>0.00306</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>( T )</td>
<td>-0.02111</td>
<td>0.01115</td>
<td>-1.893</td>
<td>0.05929</td>
<td></td>
</tr>
<tr>
<td></td>
<td>population density, % people at elevated risk, and relative humidity</td>
<td>( LPD )</td>
<td>0.8105</td>
<td>0.09407</td>
<td>8.616</td>
<td>3.94 \times 10^{-16}</td>
<td>0.2113</td>
</tr>
<tr>
<td></td>
<td></td>
<td>( O65C2P )</td>
<td>12.393</td>
<td>4.2413</td>
<td>2.922</td>
<td>0.00374</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>( RH )</td>
<td>-0.01842</td>
<td>0.009044</td>
<td>-2.037</td>
<td>0.04257</td>
<td></td>
</tr>
<tr>
<td></td>
<td>population density, % people at elevated risk, temperature, and relative humidity</td>
<td>( LPD )</td>
<td>0.7913</td>
<td>0.09413</td>
<td>8.406</td>
<td>1.72 \times 10^{-15}</td>
<td>0.2187</td>
</tr>
<tr>
<td></td>
<td></td>
<td>( O65C2P )</td>
<td>14.532</td>
<td>4.3590</td>
<td>3.334</td>
<td>0.0000964</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>( T )</td>
<td>-0.02182</td>
<td>0.01109</td>
<td>-1.967</td>
<td>0.05005</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>( RH )</td>
<td>-0.01896</td>
<td>0.009005</td>
<td>-2.105</td>
<td>0.03609</td>
<td></td>
</tr>
</tbody>
</table>

**Table 3.** Regression analyses of logarithms of Covid-19 cases and deaths with various combinations of demographic and climatic variables. Multivariate regression analyses involving precipitation data yielded very low statistical significance for the precipitation variable and hence were omitted from this table.
\[
\ln(\text{Covid-19 deaths per 100,000}) \\
= -(3.839 \pm 2.040) + (0.7913 \pm 0.1845) \ LPD + (14.532 \pm 8.544) \ O65C2P \\
- (0.02182 \pm 0.02174) \ T - (0.01896 \pm 0.01765) \ RH ,
\]

with errors corresponding to 95% confidence intervals (1.96 times the standard error). For modest variations in only temperature and relative humidity, the differential form of this relation is:

\[
\frac{\Delta(\text{Covid-19 deaths per 100,000})}{\text{Covid-19 deaths per 100,000}} = -(0.02182 \pm 0.02174) \Delta T - (0.01896 \pm 0.01765) \Delta RH ,
\]

implying that every added degree Celsius to the temperature and every added percentage point to the relative humidity is associated with a decrease in the incidence of Covid-19 deaths by (2.2±1.1)% and (1.9±0.9)%, respectively. Thus, for example, a 2°C temperature rise with simultaneous 5% increase in relative humidity is associated with a decrease in Covid-19 deaths somewhere between 7.1% and 20.5%, all other factors being equal.

As remarked earlier, the number of deaths is a more reliable indicator of the incidence of Covid-19 than case count, and we may assume that the relative decrease given by Equation (2) can be applied to Covid-19 incidence. Thus, we write:

\[
\frac{\Delta(\text{Covid-19 incidence per 100,000})}{\text{Covid-19 incidence per 100,000}} = -(0.02182 \pm 0.02174) \Delta T - (0.01896 \pm 0.01765) \Delta RH .
\]

3c. Extending data to June 2020

In the time taken to complete the preceding data collection and analysis, early June Covid-19 data became available, and we decided to use them to test the sensitivity of our preceding results based on early May Covid-19 data. Thus, we augmented our database with the June 4th Covid-19 numbers and with May climatic variables, and we compared the correlations of June Covid-19 data with the four preceding months of climatic data to the earlier correlations of May Covid-19 data with the three preceding months of climatic data.

The regression analyses on the augmented data set did not yield robust results similar to those with the earlier data set. In particular, the quadruple regression of Covid-19 deaths \((LD100)\) with population density \((LPD)\), percentage of people at risk \((O65C2P)\), temperature \((T)\), and relative humidity \((RH)\) yielded \(p\)-values rising from 0.050 to 0.113 for temperature and from 0.036 to 0.296 for relative humidity. Put another way, the confidence values associated with having a non-zero correlation with temperature dwindled from 95% to 89% and with relative humidity from 96% to 70%. The slope coefficients also dropped slightly and standard errors rose slightly. For June Covid-19 deaths, the quadruple regression was found to be:
\[
\ln(\text{Covid-19 deaths by June per 100,000}) \\
= -(4.461 \pm 2.082) + (0.7997 \pm 0.1868) \ LPD + (13.632 \pm 8.696) \ O65C2P \\
- (0.01966 \pm 0.02424) \ T - (0.00953 \pm 0.01785) \ RH .
\] (4)

As mentioned earlier, one likely interpretation is that, since our analyses do not include variables that model gradual relaxation of confinement directives, different states, cities and regions reopening at substantially different rates throughout the months of May and June made the later data much noisier from a statistical standpoint. Another likely interpretation is that, due to the combined effects of confinement (particularly strict for persons at risk), and of adhesion to sanitary protocols such as washing hands, social distancing and wearing of face masks, the number of Covid-19 cases and deaths had become by June 2020 much less sensitive to temperature and relative humidity. This does not mean, however, that the transmission of the virus lost its dependence on these atmospheric factors; it only says that the dependence became obscured by other factors not retained in our regression analyses.

4. Implications and Recommendations

According to a study conducted by the U.S. Environmental Protection Agency (EPA, 1989\(^9\)), people in the United States spend the vast majority (87%) of their time indoors, nearly 21 hours per day. If we subtract the 14 hours between 6pm to 8am when people tend to be at home, the estimation is that people spend, on a daily basis, 7 out of 10 hours being inside a building that may not be their home and 3 hours outside, possibly in their cars or using public transportation. Lately, because of confinement measures, people have spent additional time in their homes, inside which the humidity level is fairly similar to that outdoors.

With few exceptions nowadays, office and commercial buildings have a controlled indoor climate with little variation in temperature and relative humidity to ensure comfort. If people were spending their entire time inside such buildings, atmospheric data would have little to no impact on virus transmission, but our preceding analysis does show statistically significant correlations, and we must conclude that the three or so hours spent outdoors do affect the transmission of SARS-CoV-2. If merely 3 hours of exposure to ambient environmental conditions do have an impact, it stands to reason that the 7 or so hours spent inside a climate-controlled building should have a similar, if not stronger, impact on the transmission of the virus.

This brings us to ask like others before us (Moriyama et al., 2020; ASHRAE, 2020) whether some adjustments to indoor climate control could be recommended in order to mitigate transmission of SARS-CoV-2 inside buildings. To be acceptable, such adjustments would not only have to be effective in reducing virus transmission but also meet two important criteria, namely maintenance of occupants’ comfort (with effect on employees’ productivity) and avoidance of damage to the building.

Environmental comfort inside buildings is largely controlled by five factors, namely lighting, rate of ventilation, purity of the air, temperature, and relative humidity. Lighting does not concern us here. The

\(^9\) This source is admittedly dated, but there does not seem to be more recent statistics of the time spent indoor. In any case, we only use the number as indicative.
rate of ventilation has already been addressed by Memarzadeh (2011) and the American Society of Heating, Refrigerating and Air-Conditioning Engineers (ASHRAE, 2020) who have both determined that a higher rate of ventilation reduces risk. Sufficient purity of the air is assured by avoidance of mold and proper filtration; some (Schoen, 2020) are recommending stronger filtration to combat transmission of SARS-CoV-2, but it is not clear whether a tighter filter many meters away in the mechanical room would be effective against a virus particle that may not make it that far before being biologically disabled or removed from the air flow by settling on a surface. What we are asking here is whether there might be permissible changes in room temperature and/or relative humidity and what benefit could be derived from such changes.

To explore this possibility, we begin with the definition of the comfort level inside a building based on temperature and relative humidity. There does not seem to be a single standard that fits all situations. ASHRAE recommends that operative temperatures range between 68.5°F (20°C) and 75°F (24°C) in the winter, and from 75°F (24°C) to 80.5°F (27°C) in the summer, and indoor relative humidity be maintained at or below 65%, with 50% being ideal (ANSI/ASHRAE Standard 55-2017; document not public and available for sale; Harriman et al., 2008, pages 69-76). There is no one-size-fits-all comfort zone because physiological comfort depends not only on temperature and humidity but also on the level of air ventilation, the level of physical activity of occupants, and seasonal clothing. For the sake of the discussion below, we selected the following definition of the comfort zone in the \((T, RH)\) space of the psychrometric chart: Temperature \(T\) between 20°C (68°F) and 26.7°C (80°F), and relative humidity \(RH\) between 30% and 70%, with truncation of the upper-right corner as depicted in Figure 3 below.

An ideal setting, close to the “productivity maximum” of employees in an office building (Kosonen & Tan, 2004), is \(T = 23.5°C\) and \(RH = 50\%\), in the center of the comfort zone as indicated by an asterisk on Figure 3. Taking this central point as the most likely setting in a climate-controlled building, we use it as the reference point and label it with 0% to indicate the point of no change. We then use Equation (3) with various positive changes \(\Delta T\) in temperature and \(\Delta RH\) in relative humidity to estimate the relative decrease in Covid-19 incidence that could result from those changes. The top point of the comfort zone (labeled A on Figure 3) corresponds to \(T = 24°C\) and \(RH = 70\%\). If the climate control system of the building is set at these values, the changes from the center point are \(\Delta T = 0.5°C\) and \(\Delta RH = 20\%\), and the projected benefit reduction according to Equation (3) is

\[
\text{Relative decrease in Covid-19 incidence per 100,000} = (0.02182 \pm 0.02174)(0.5) + (0.01896 \pm 0.01765)(20) = 0.390 \pm 0.364 = 39\% \pm 36\%.
\] (5)

For point B, with \(T = 26.7°C\) and \(RH = 55\%\) and thus \(\Delta T = 3.2°C\) and \(\Delta RH = 5\%\), the expected benefit is projected to be 16.5% ± 15.8%. For point C with same temperature increase of 3.2°C as for point B but no increase in humidity, the expected benefit is only 7.0% ± 7.0%. Note that the points considered here do not jeopardize the integrity of the building as they all correspond to levels of relative humidity that do not exceed 70%.
Figure 3. The psychrometric chart of air with water vapor at standard atmospheric pressure on which is delimited the comfort zone for indoor environments (yellow region) with ideal set point in the center (marked by an asterisk) and selected points, labeled A, B, C and D, of slightly higher temperature and/or relative humidity that should reduce the risk of Covid-19 incidence. Source of unannotated chart: Carrier Corporation, 1975.

Venturing a bit out of the comfort zone to point D in Figure 3 (at 26°C and 70%), the expected benefit rises to 43% ± 41%, which is higher than the 39% ± 36% of point A. Thus, given the values of the coefficients and the permitted increases of temperature and relative humidity, it appears that an increase in relative humidity is potentially more effective than an increase in temperature, and the recommendation is to aim for corner A of the comfort zone, corresponding to 24°C (75°F) and 70% relative humidity.

Aside from biological comfort, a word may be said about employee productivity since estimates have been made to assess productivity loss when temperature and relative humidity are less than ideal. Kosonen and Tan (2004) calculated that productivity is optimal at $T = 24°C$ and $RH = 50\%$, and, from there, that 1 additional degree in temperature (from 24°C to 25°C) causes a 1.9% in employee productivity loss and a 3°C increase an escalating 15.0% loss. An increase of relative humidity from 50% to 65% causes a 0.3% loss at 24°C, a 3.4% loss at 25°C, and an 18% loss at 27°C. Thus, from the perspective of employee productivity, temperature is the more sensitive factor of the two, and, if one parameter is to be changed, it is preferable to increase relative humidity before temperature. Fortunately, this coincides with the preferred direction for greater mitigation of SARS-CoV-2 transmission. In sum, the more effective way to reduce SARS-CoV-2 transmission is to increase relative humidity before temperature, and this may be accomplished while staying in the comfort zone and with low productivity loss.
We recommend adopting the settings of point A of Figure 3, which are $T = 24^\circ C$ ($75^\circ F$) and $RH = 70\%$ corresponding to a dew point of $18.2^\circ C$ ($64.7^\circ F$), with estimated Covid-19 incidence reduction of 39% ± 36% (minimum of 3%) and concomitant productivity loss of only 0.6%.

5. Conclusions

We asked the question whether ambient temperature and relative humidity may influence the transmission of the SARS-CoV-2 virus, and for this purpose we compiled a database that combines Covid-19 incidence, climatic variables, and demographic data. The spatial coverage is the set of 306 Hospital Referral Regions of the United States, accounting for 97% of the US population, thus eliminating any bias in the selection of the localities aside from the choice of the United States. The Covid-19 incidence data were the numbers of cases and deaths per 100,000 people as of May 4th 2020. The demographic data were population density, median age, household median income, and percentage of people at elevated risk (persons 65-years old or older and with at least two existing chronic conditions).

The climatic variables were the monthly means of temperature and relative humidity and the total precipitation for the months of February, March and April 2020. Arguably, it would have been preferable to use the actual weather data for early 2020 rather than climatic averages, but we did not have the human and time resources to allow us to do so. Nonetheless, we did find significant relationships, and it is to be expected that the use of actual weather data for early 2020 would only have made those sharper and more precise. In a prolongation study, we added the Covid-19 incidence data of a month later (June 4th) and added the climatic data for the month of May.

We performed a series of regression analyses of the Covid-19 incidence data versus a series of combination of demographic and climatic data. Not surprisingly, the strongest correlations were with population density and the percentage of people at elevated risk, but, although weaker, correlations with temperature, precipitation and relative humidity were found to be statistically significant, too. The multivariate regression of Covid-19 deaths versus population density, percentage of people at elevated risk, temperature and relative humidity yielded the following fit (Equation (1) above):

\[
\text{Natural logarithm of Covid-19 deaths per 100,000} \\
= -(3.839 \pm 2.040) + (0.7913 \pm 0.1845) \ LPD + (14.532 \pm 8.544) \ O65C2P \\
- (0.02182 \pm 0.02174) \ T - (0.01896 \pm 0.01765) \ RH, \\
\]

in which $LPD$ is the natural logarithm of the population density (people per square miles), $O65C2P$ is the percentage of people at least 65 years old and with at least two chronic conditions, $T$ is temperature (in degree Celsius), and $RH$ is relative humidity (expressed as a percentage). All correlations are statistically significant with a likelihood of 95% or higher. We note that the correlations with temperature and relative humidity are each negative, implying that an increase in either temperature or relative humidity or both is associated with reduced transmission of the SARS-CoV-2 virus. The prolongation study yielded similar, but weaker, conclusions, likely due to certain non-climatic factors including differential reopening and better sanitary protocols, factors not retained in the present study.
This finding prompted us to explore what possible adjustments could be made to temperature and relative humidity in climate-controlled buildings that would reduce Covid-19 transmission between occupants while not jeopardizing their comfort and, for employees, also their productivity. We estimate that, if relative humidity is raised from the ideal value of 50% to the still comfortable value of 70% and if temperature is raised by 0.5°C to 24°C (75°F), these changes could be associated with Covid-19 transmission reduction by at least 3% and likely by 39% while affecting employee productivity by less than 1%

The preceding considerations were made in the context of buildings, but it is evident that they may also be applied to climate-controlled vehicles used in public transportation such as municipal buses, subway cars, and trains. An increase of relative humidity in those vehicles may be beneficial to the same extend that it may be beneficial in a climate-control building. There is no claim made here, however, that such a measure should be in lieu of more important measures such as social distancing, frequent hand washing, and the wearing of a mask.

We caution, however, that an adjustment to temperature and humidity settings in a climate-controlled building or vehicle does not imply that climate-control is necessarily an advantage. There is emerging evidence (Qian et al. 2020; Frémont, 2020; World Health Organization, 2020) that climate-controlled indoor spaces account for the vast majority of places where new Covid-19 cases are occurring post-confinement. The implication is that natural ventilation is preferable to avoid the recirculation of the virus inside an enclosed space. This may explain why Covid-19 post-confinement has made a much greater resurgence in the southern USA where most offices and restaurants are climate-controlled than in Europe where they are not. We merely recommend that, if the building (or vehicle) must be climate-controlled, then it is preferable to raise slightly its temperature and relative humidity settings.

Our finding based on climatic data that an increase in temperature and relative humidity may lower transmission of the SARS-CoV-2 virus needs to be framed within the context of far more important and well established factors such as age, comorbidities, population density, and adherence to prophylactic measures. Our recommendation does not substitute for the much more important directives issued by the medical profession including social distancing, wearing of a mask, and frequent hand washing.

Acknowledgements

Lew Harriman of Mason-Grant Consulting provided pertinent literature and generous advice. His assistance and criticisms are much appreciated. Covid-19 data were kindly provided by Dr. Elliott S. Fisher, Professor of Health Policy and Medicine at The Dartmouth Institute for Health Policy & Clinical Practice.

REFERENCES


doi.org/10.1503/cmaj.200920

doi.org/10.1016/j.enbuild.2004.06.021

doi/10.1101/2020.02.12.20022467


www.medrxiv.org/content/10.1101/2020.04.04.20053058v1


doi.org/10.1001/jama.289.2.179

irp-cdn.multiscreensite.com/c4e267ab/files/uploaded/kt34RqduTfGjxf3soeQ_EPA_Report%20to%20Congress%2
www.who.int/news-room/commentaries/detail/transmission-of-sars-cov-2-implications-for-infection-prevention-precautions