Environmental temperature does not correlate with COVID-19 spreading and death toll

Rafael Franco^{1,2} and Miguel Galo-Fernández³

1. Department of Biochemistry and Molecular Biomedicine. University of Barcelona. Barcelona. Spain.
2. CiberNed. Centro de Investigación en Red, Enfermedades Neurodegenerativas. National Health Institute Carlos iii, Madrid. Spain.
3. Statistical advisor (www.galo.es). miguel@galo.es. Murcia. Spain.

Correspondence

Rafael Franco

Department of Biochemistry and Molecular Biomedicine.

Diagonal 643.

Facultat de Biologia. Prevosti Building

rfranco123@gmail.com; rfranco@ub.edu

Tel +34 934012108

Abstract

Background. Appearance of outbreaks due to some viruses is seasonal; in temperate (non-tropical) areas of the World, viruses are usually silent in summer while they reappear in winter. Recent attempts to putatively prove this hypothesis has led to promising data whose analysis lacks scientific rigor.

Methods. Using data (cases and deaths) and the average temperature in April of a country in which COVID-19 has badly hit, Spain, we have applied statistical tests to look for the correlation , or lack thereof, between these parameters.

Results: The kinetics of current acute COVID-19 pandemics does not correlate with temperature. Using absolute number of cases and death rate and average temperature in April in either Spanish provinces/regions or in Catalanian *comarcas* (*comarca* being similar to Counties in the United States), we find that the data are independent. What our data tells is that temperature is not a key factor in current evolution of this particular pandemic.

Conclusions. The pandemic curve in a territory is not differentially affected by environmental temperature. While it is a fair wish, there cannot be any scientific evidence supporting the hypothesis of SARS-CoV-2 disappearance in summer. Only time will tell.

Keywords

Temperature, SARS-CoV-2, COVID-19, incidence rate, time-series regression, Barcelona, Spain.

Introduction

COVID-19 pandemic is affecting world population in a way that was not known in recent history. The serious condition having a death toll of at least 1%¹ and affecting the economy of population worldwide² is managed mainly by confinement to avoid person-to-person infection, i.e. to reduce infectivity rate.

It is known that the first year of appearance of a new virus the affectation is higher due to the lack of previous exposure to the pathogen. When the pathogen reappears, the immune system of previously-exposed individuals impedes infection. Then, already exposed individuals rarely become infected nor transmit the infection in the next season (see³ and references therein). This natural mechanism of defense against new pathogens is reinforced by the development of vaccines or of specific drugs, both of which are only available upon testing for safety and efficacy, i.e. some years after the appearance of the new virus. In this scenario the main objective is to wait for infection decay and be ready for reappearance. For many pathogens outbreaks are seasonal. The main example is provided by the influenza virus; in temperate (not tropical) areas of the World flu appears in winter and disappears in summer⁴. We recommend the review of⁵ to understand flu epidemiology.

To the best of our knowledge there is not any real explanation of why the influenza virus becomes undetectable in the months with higher temperature. Probably there are multiple factors that impact, among them a lower latency time. In fact RNA viruses such as the influenza and SARS-type, have a limited life when outside an infected body; accordingly higher temperatures inactivate quickly the virus outside the infected individual thus decreasing the infective power. Also, the ultraviolet rays of the Sun, which appears more hours in summer than in winter, kill pathogens. It is likely that human bodies in summer have more mechanism of defense against infection. The mechanism are unclear but one that has been put forward is a higher activity of epithelial cell of the lung to trap/remove particles. Taken together it is tempting to speculate that COVID-19 will disappear in summer. This will be known in the future; in the meantime we can try to make correlations between temperature and COVID-19 infection. This paper aims at testing this possibility by correlating cases and temperature in a Country with very high COVID-19 incidence, Spain. As of today May 12, the death rate per million population is 572, right after San Marino, Belgium and Andorra. If we take only Madrid and Catalonia regions the death toll per million population is the highest in the world and comparable to that of San Marino (>1,200) (<https://www.worldometers.info/coronavirus/>; date: May 12, 2020). Our results indicate that cases/deaths in this first COVID-19 infection in Spain do not correlate with temperature.

Methods

Data collection

All data are taken from official National and/or regional sources, mainly from the INE (Spanish National Institute for Statistics) and METEOCAT (Catalonian service for weather data). No data of temperature is available for 2020, therefore we took data from the most recent year for which temperatures are available. Data are available in the accompanying Excel. We assumed that differences on average temperatures in April in 2020 versus those in previous years are not biased in any specific Spanish location. We have confirmed by comparing average temperature in some few years backwards, that despite global warming, average temperatures have not significantly changed in Spanish regions in short periods of time. Figures were prepared using Excel and its linear regression utility.

Statistical analysis

The IBM SPSS® Statistics package was used to calculate the Pearson correlation coefficient and the p value in contrast of hypothesis on correlation between i) cases/deaths versus temperature and ii) deaths/population. Significance was considered when $p < 0.05$.

Results

Apparent positive correlation between cases and April average temperature in Catalonia

As indicated in Methods we took the number of cases and of COVID-19-related death from official sources and the temperatures publicly available; in Catalonia there is not any information for 2019 or 2020 but for 2018. All data used in statistical analysis are found in Supplementary material (Excel file; available upon request to the corresponding author). Within Catalonia data were taken by *comarcas* that are roughly equivalent to what counties are in the US. Figure 1 shows the point cloud chart of deaths per 10,000 population and temperature variation (respect to average) and the line resulting from linear regression fit. The temperature taken for a given *comarca* is the average temperature in April. Despite the apparent positive correlation that is suggested by the positive slope of the linear plot, statistical analysis shows that the null hypothesis holds, i.e. the two compared sets of data are independent.

The Pearson correlation coefficient was 0.107 when the pair of death rate/temperature variation data were taken into account. This value and the number of data pairs (39) allows calculating the p value. Indeed the two sets of data are independent because the p value is 0.581. As a control the same type of analysis was performed taking into account the death rate and the population. The Pearson correlation coefficient was 0.980 that leads to a extremely small p value when $n=39$. Then, as expected, the death toll and population are correlated.

Similar results are obtained with cases as deaths are roughly proportional to cases. It is however important that cases/deaths are counted using the same criteria. This requirement was not met in April but homogeneous criteria across Catalonia and Spain was ruling when the data was collected by May 1, 2020.

Figure 1. Plot of deaths versus April temperature variation in Catalanian “comarcas” (see Supplementary material for raw data). Death are presented per 10,000 population and X axis indicate temperature variation for each “comarca” in April respect to April average temperature in Catalonia (in year 2018). Death toll as of May 1, 2020.

The apparent positive correlation contrasts with visual inspection of the heat map of deaths in Catalanian “comarcas”. Temperatures in April in Catalonia are colder going North (up in Figure 2) and West (left in Figure 2). Areas hardest hit by the epidemic are not necessarily those in North West areas. By comparing relatively important cities with similar latitude, Lleida with is the capital of a province and Barcelona, another capital of province, it is true that cases are most abundant in the warmer city (Barcelona). However, this trend is completely opposite when we one goes South from Barcelona, where warmer cities do present much less cases than in Barcelona or than -the even colder- Girona. Again it is unlikely that COVID-19 incidence in Catalonia in April correlated with temperature.

Figure 2. Heat map of death rates in Catalonia. Relatively low death rates (normalized by million population) are in grey color scale and relatively high death rates are in pink/magenta color scale; the darker the color the higher the death rate. The diameter of the pink circle is roughly proportional to the normalized death rate. Taken on May 12, 2020 from National.cat on-line newspaper (https://www.elnacional.cat/es/salud/coronavirus-mapa-contagios-catalunya-municipios_485916_102.html). A similar image but taken on May 1st if available in ⁶.

Apparent negative correlation between cases and April average temperature in Spain

We next wanted to address whether the apparently positive trend in Catalonia was also found in the whole Country, Spain. We took the number of cases and of COVID-19-related death from official sources and the temperatures publicly available; in Spain there is not any information for 2020 but for 2019. All data used in statistical analysis are found in Supplementary material (Excel file; available upon request to the corresponding author). Within Spain data is taken by provinces except when death toll is not available per provinces but for various closely-located provinces constituting a region. Then we used either data from provinces except for Catalunya, *Galicia*, *Comunidad Valenciana* and *Pais Vasco*. It should be reminded that Catalonia has data for every *comarca* but not for provinces; therefore we used average data for the whole region (and similarly for *Galicia*, *Comunidad Valenciana* and *Pais Vasco*). Data from these regions were also included in the analysis. Figure 3 shows the point cloud chart of deaths per 10,000 population and temperature variation (respect to average) and the line resulting from linear regression fit. The temperature taken for a given province/region is the average temperature in April.

Despite the apparent negative correlation suggested by the negative slope of the linear plot, statistical analysis shows that the null hypothesis holds, i.e. the two compared sets of data are independent.

The Pearson correlation coefficient was 0.049 when the pair of death rate/temperature variation data were taken into account. This value and the number of data pairs (41) allows calculating the p value. Indeed the two sets of data are independent because the p value is 0.761. As a control the same type of analysis was performed taking into account the death rate and the population. The Pearson correlation coefficient was 0.848 that leads to a extremely small p value when $n=41$. Then, as expected, also in Spain the death toll and population are correlated.

Similar results are obtained with cases as deaths are roughly proportional to cases. A note of caution is needed as before along April 2020 the criteria to count COVID-19 cases and number of deaths changed. At the date of this study (May 1st) the criteria to count cases and deaths was the same all around Spain (by Government rules).

Figure 3. Plot of deaths versus April temperature variation in Spanish provinces/regions (see Supplementary material for raw data). Death are presented per 10,000 population and X axis indicate temperature variation for each province/region in April respect to April average temperature in Spain (in year 2019). Data from regions were considered when official death rates were not available for individual provinces in given region. Death toll as of May 1, 2020.

Discussion

Data on this paper shows that there is not any correlation between the spread of SARS-CoV-2 averaged on the month when COVID-19 peaked in Spain and the average temperature in different regions of the Country. Just having a look to Figure 2 and taken into account the four main cities in Catalonia, it is evident that cases are fewer in warmer Tarragona than in Barcelona but also in colder Girona and much colder Lleida⁶. Furthermore, the number of cases and normalized number of deaths are fewer in Lleida, which is colder than Girona. Going to data in overall Spain, there is a seemingly negative correlation between death and average temperature that might lead to accept that the

higher the temperature the lesser the death toll. But in Catalonia whose capital is Barcelona in which the report by ⁷ seems to prove such negative correlation, our careful analysis shows the opposite, i.e. the higher the temperature the worse the COVID-19-related numbers. In fact all those data can only be reconciled by the fact that there is not any correlation. Linear regression assumes that data can be properly fitted to a straight line but this is not the case as clearly deduced from the position of the “experimental” points in Figures 1 and 3. Distribution of data points is not linear and in fact it can neither be adjusted to any other another analytical curve (quadratic, exponential, logarithmic, etc).

Google Scholar search for “temperature and COVID-19” provides 9690 hits as of May 12, 2020. The first hit is a recent report stating that high temperature and high humidity reduces COVID-19 transmission. The paper seemingly demonstrates that 1°C of temperature increase decreases by 0.0225 what authors call “*daily effective reproductive number*” (R), being the decrease of 0.0158 when the humidity increases by 1% ⁸. Although there are flaws in the statistics and the approach misses that viral transmission is not immediate, these numbers are modest when compared with those of ⁷ who seemingly demonstrate that for every °C of temperature increase there is (in Catalonia) a 7.5% reduction in the COVID-19 “incidence” rate.

Wishes are not legitimate if fueling predictions disguised of rigorous research. Not surprisingly virtually all the reports go in the same direction. The fact that other SARS viruses slow down progression depending on the temperature and humidity ⁹ is irrelevant in the case of SARS-CoV-2. Common behavior of previous viruses cannot serve to predict the behavior of a new virus. But even today when we know that influenza viruses disappear in summer, there is controversy on the real correlation between flu spread and temperature, humidity and wind speed ¹⁰. There are two types of misunderstanding underlying data handling in the ⁸, in the ⁷ and in virtually all the recent papers “demonstrating” correlation between COVID-19 cases and temperature. Actually they have in common the analysis of day-to-day variations in temperature (or in humidity). The now famous epidemic curve has a similar inverted U shape in all countries. Neither the rise nor the decay has anything to do with temperature; in fact, the first highest slope in the rising of COVID-19 cases in April was in warmer Italy, and few days later Italy was surpassed by warmer Spain. In the most acute phase on pandemics Spain had the highest slope despite being the warmer country compared with Countries whose peak of COVID-cases were preceding Spain. In summary, virus spreading in a Country with high average temperature in April was higher than in colder countries. A second reason is also related with immediate data collection as any viral infection takes days to be detected. Virus transmission depends on the viral exposure that, for COVID-19, in turn depends on the confinement measures and how they are enforced. Moreover, virus-derived symptoms do not immediately appear but take an undetermined number of days. Last but not least, it is required that the cases are real and, much unfortunately, there has been a shortage of tests in Spain, thus impeding to know the real number of infected people and missing asymptomatic cases.

According to the US centers for disease control (CDC), nonpharmacological interventions in a viral pandemics such as any of those produced by influenza viruses aim at: “1) *delay the exponential growth in incident cases and shift the epidemic curve to the right in order to “buy time” for production and distribution of a well-matched pandemic strain vaccine*, 2) *decrease the epidemic peak*, and 3) *reduce the total number of incident cases, thus reducing community morbidity and mortality*” ¹¹. These are sensible rules to implement while waiting that SARS-CoV-2 will, in a temperate Country like Spain, disappear in summer. Disappearance or not will be empirically ascertained. Even if it disappears for some months preparation for next-season outbreaks

also requires interventions such an epidemiological study to detect and make a profile for asymptomatic carriers and to early detect the beginning of next COVID-19 outbreak ⁶.

References

1. Onder G, Rezza G, Brusaferro S. Case-Fatality Rate and Characteristics of Patients Dying in Relation to COVID-19 in Italy. *JAMA - J. Am. Med. Assoc.* American Medical Association; 2020.
2. McKibbin WJ, Fernando R. The Global Macroeconomic Impacts of COVID-19: Seven Scenarios. *SSRN Electron J.* Elsevier BV; 2020 Mar 5;
3. Hayward AC, Wang L, Goonetilleke N, et al. Natural T cell-mediated protection against seasonal and pandemic influenza: Results of the flu watch cohort study. *Am J Respir Crit Care Med* [Internet]. American Thoracic Society; 2015 Jun 15 [cited 2020 May 12];**191**(12):1422–1431. Available from: <http://www.atsjournals.org/doi/10.1164/rccm.201411-1988OC>
4. Roussel M, Pontier D, Cohen J-M, Lina B, Fouchet D. Quantifying the role of weather on seasonal influenza. *BMC Public Health* [Internet]. BioMed Central Ltd.; 2016 May 26 [cited 2020 May 12];**16**(1):441. Available from: <http://www.ncbi.nlm.nih.gov/pubmed/27230111>
5. Cannell JJ, Zaslloff M, Garland CF, Scragg R, Giovannucci E. On the epidemiology of influenza [Internet]. *Virology*. 2008 [cited 2020 May 12]. p. 29. Available from: <http://www.ncbi.nlm.nih.gov/pubmed/18298852>
6. Franco R. COVID-19 Management in Spain. Errors to Avoid and the Need of A Post-Confinement Longitudinal Epidemiological Study. Preprints; 2020 May 8;
7. Tobías A, Molina T. Is temperature reducing the transmission of COVID-19? *Environ Res.* Academic Press Inc.; 2020 Jul 1;**186**(April 18):In the Press.
8. Wang J, Tang K, Feng K, Lv W. High Temperature and High Humidity Reduce the Transmission of COVID-19. *SSRN Electron J.* Elsevier BV; 2020 Mar 11;
9. Chan K, Malik Peiris J, Lam S, Poon L, Yuen K, Seto W. The Effects of Temperature and Relative Humidity on the Viability of the SARS Coronavirus. *Adv Virology*. 2011;**Article ID**:1–7.
10. Peci A, Winter AL, Li Y, et al. Effects of absolute humidity, relative humidity, temperature, and wind speed on influenza activity in Toronto, Ontario, Canada. *Appl Environ Microbiol.* American Society for Microbiology; 2019 Mar 1;**85**(6).
11. Centers for Disease Control and Prevention. Interim Pre-pandemic Planning Guidance : Community Strategy for Pandemic Influenza Mitigation in the United States — Interim Pre-Pandemic Planning Guidance : Community Strategy for Pandemic Influenza Mitigation in the United States — Early , Targeted , La [Internet]. *CDC stacks* 2007. Available from: <http://pandemicflu.gov/professional/community/%0Ahttps://stacks.cdc.gov/view/cdc/11425>