Country comparison of the COVID-19 Pandemic: What happened in Latin America, and why?

GalileoViolini

PUCMM

To follow the development of COVID-19 Pandemic the usual indicators have been: the number of confirmed contagions, that of the patients requiring intensive therapy, that of fatalities, and that of recoveries. All of them, except perhaps the second, present elements of uncertainties in their correlation to the pandemic.

A way to attempt to reduce these uncertainties is the simultaneous consideration of an indicator in different countries, comparing the different results. In this paper we shall make use of the number of confirmed contagions, as one which may allow to draw interesting conclusions.

For example, the study of its behaviour during the triggering phase of the pandemic has suggested that in some cases one can infer whether the initial development of the outbreak is due to a single parent case, or to a cluster of bursty events[1]. As another use of this indicator, one can mention the estimate of the number of asymptomatic or low symptoms cases[2].

The reason that make advantageous the comparison of the development of the pandemic in several countries is that one can assume a rather reasonable working hypothesis, namely the expectation that it follows a similar pattern irrespective of the country. On turn, this implies that if one finds appreciable deviations from a common pattern, one can expect that they are due to reasons that can be understood, suggested, studied and proven.

Of course, for the explanation of such differences one can and should look at the most different sectors. These include demography, ethnical differences, virology, economics, health system efficiency, sociology and environmental features, just to make an indicative, non-exhaustive list.

Even if certain indications can be obtained from the analysis of few cases, such as those of the initial development of the epidemic[3,4], it is clear that with the increase of the number of the contagions, possible differences of behaviour of the outbreak between different countries acquire weight and significance.

In this note we shall present evidence that one such difference exists at continental level when one compares the initial outbreak of the pandemic in Latin America with that in the countries where it has been more important.

Of course, the disadvantage of looking at the initial development is that data are less abundant and more subject to fluctuations, but in our analysis we shall see that the consideration of several countries not only should overcome this difficulty, but also provided a broad evidence that the difference is genuine.

The development of the number of the contagions N is expected to be described by a logistic curve. A good approximation of this curve in the initial phase is an exponential. Within this approximation, log N exhibits a linearly increasing behaviour, which eventually will become that of the logarithm of the logistic function.

In this study, we shall perform a linear analysis of $\log_{10}$ N, using the daily data of the confirmed cases of several countries[5].

We shall carry on two different types of fit to the data. A first analysis will deal with the initial development of the outbreak, and a second with the variation of the results, when one analyzes the evolution of the number of the contagions between a fixed initial day and a variable last day.

We find that, as one could expect because of the logistic behaviour of N, and as noticed in a previous paper[6] the slope of the line bestfitting the data decreases with the increase of the day chosen to end the studied period.

From the first type of analysis, we can prove that there is a significant difference between the way the pandemic started to spread in Latin America and in the group of countries where its actual development is larger. A preliminary result in this sense had been announced in a previous paper of limited diffusion.[7] Possible reasons of this difference will be discussed.

The second type of analysis is made possible by the fact that, having abundant data over a longer time period it is possible to analyze the variation of the slope of the straight bestfitting line in function of the variation of the considered period of time. As we shall show, this variation, already pointed out in reference 6, in most cases is linear. Such feature is present in the data referring to the evolution of Hubei pandemic, where one can find that the significant reduction of the new contagions took place when the slope reached a value of the order of 0.002 -0.003.

It may be observed that in each fit the value of the slope of log N allows to estimate the duplication time of N in that moment[8]. In practice, since we use $\log_{10}$ and consider daily N variations, the duplication time (in days) may be easily calculated dividing 0.3 by the value of the slope.

In our analysis we shall compare two groups of countries. The first group includes the sixteen countries having, at the date of 11 April 2020, a significant spread of the pandemic. For this we considered countries for which N>13000.[5] Only one of them, Brazil, belongs to Latin America. In Latin America the outbreak burst later, and, as we shall see, with a different development. Therefore for our comparison we can only find countries with a lower value of N, and this led us to select the group of the seven Latin American countries where, at the same date, N was larger than 2200.

We assume for $\log_{10} N$ a linear parametrization,

$$\log_{10} N = \beta x + a$$

where x is the number of the days elapsed since the day before the first day whose N data is fitted.

Our first comparison refers to the development of the outbreak between 50 and 1500 contagions. In the preliminary analysis we were referring to[7], the two groups of countries present minor differences, because of the slightly different criterion of choice.

Table 1a

Country	β	R^2
United States	0.1233	0.9958
Spain	0.1434	0.9888
Italy	0.1649	0.9724
Germany	0.1347	0.9808
France	0.1285	0.9798
Iran	0.2011	0.9990
United Kingdom	0.1086	0.9958
Turkey	0.2600	0.9872
Switzerland	0.1209	0.9804
Belgium	0.1117	0.9676
Netherlands	0.1065	0.982
Canada	0.1017	0.9888
Austria	0.1254	0.9953
Portugal	0.1317	0.9904
Russia	0.0955	0.9952

Table 1b

Country	β	R^2
Brasil	0.0848	0.9644
Chile	0.0747	0.9697
Ecuador	0.0666	0.9109
Peru	0.0605	0.9788
Mexico	0.0661	0.9872
Panama	0.0627	0.9764
Dominican Republic	0.0667	0.9633
Colombia	0.0722	0.9636

The differences between the values of β of the two groups of countries are evident. The most striking one is that the mean values are 0.1235 and 0.06929, respectively, and this essentially means a factor 2 of difference in the duplication time. A difference is observed for any pair of countries one from Table 1a and one from Table 1b. Indeed, no value of β in Table 1b is larger than any value of β in Table 1a. Moreover, a significant difference is also observed in the time elapsed to go from $N \approx 50$ to $N \approx 1500$, which is 15.5 and 11 days, respectively.

From these results, one can draw the conclusion that the rapidity of the initial diffusion of the pandemic in Latin America has been slower than in the other group of countries we considered.This conclusion is confirmed if we compare the global growth of the contagions between 20 March 2020 and 11 April 2020 with that between 5 March 2020 and 20 March 2020 (Table 2).[8]

Table 2

Date/Region population	Europe 750 M	Asia exc. China 3100 M	North America 500 M	South America 430 M	Africa 1350M	CenterAmerica Caribbean 90 M
March 5	5930	11978	276	34	43	6
March 20	1298950	35803	20208	2252	1075	494
April 11	1360660	206219	553516	51887	13357	8159

The reason of this difference can only be determined by a specific dedicated research, but it seems clear that one can safely exclude that it may be a fluctuation. Therefore, let us shortly discuss about possible explanations. Some of them may be quantified, others can only be discussed in a qualitative way.

Two differences are well known, one is the average age of the population and the other the average temperature of the countries, and both of them have been object of consideration to explain different behaviours of this pandemic in different countries, such as the high Italian lethality[2,9], or the correlation between confirmed cases of contagions and environmental factors, such as temperature, humidity and UV radiation.[10]

The average age of the populations within each group of countries is very different. It is approximately 29 years in the Latin American group and 41 years in the other[7], and this could explain a possible different resilience to COVID-19. Instead, it seems unlikely an ethnical explanation, considering that the ethnic composition of the Latin American countries is very different, with countries with large Afroamerican population, countries with large presence of descendants of the original American populations and that a similar result holds also for a country like Argentina[7] where the non Caucasian population is a minority.

Regarding temperature we refer to the data of reference 7, with the modifications required by those in the composition of the two groups of countries. Therefore, we considered for each country the minimum and maximum March average temperature (in C°) of one or two cities where the pandemic spreaded more:

First group: New York (2,11), Barcelona (10,17), Madrid (6,16), Milan (7,14), Berlín (1,9), Munich (1,9), Paris (5,13), Colmar (2,12), Teheran (7,16), Birmingham (3,10), Istanbul (5,12), Lausanne (4,11), Antwerp (3,13), Mons (3,11), Tilburg (2,10), Amsterdam (2,10), Montreal (-6,2), Toronto (-2,5), Ischgl (-1,6), Lisbon (10,18), Moscow (-5,1), Saint Petersburg (-4,2). Averages: 2.5 and 10.4.

Second group: Sao Paulo (20,28), Rio de Janeiro (24,30), Santiago (9,28), Guayaquil (24,31), Lima (20,27), Mexico City (8,26), Panama City (25,32), Santo Domingo (20,30), San Francisco de Macorís (19.3,24.5), Bogota (8,20). Averages 17.7 and 27.6.

Indeed, the average March 2020 temperatures are very different.

Another possible explanation could refer to the testing number and protocols, because these are indicators with tremendous differences. However, it is very difficult to estimate the relevance of this data.

Another qualitative difference that possibly might have played a major role for this difference regards the social distancing and lockdown measures. These measures, although similar, namely border closure, university and school closure, quarantine and lockdown, were implemented more or less at the same moment (mid March) in the countries of either group, but when the actual level of development of the outbreak was very different. Concerning the importance of one of these measures, the border closure to flight, it has been estimated that it averted 71 % more contagions.[11]

However, there is also another intriguing possibility. It is something which was mentioned as a possibility in the context of the discussion of the pandemic in Lombardy, namely the existence of different types of viruses[12].

This possibility received support from the recent analysis of Forster et al. [13] that identified three different types of virus and associated them with different geographical diffusion of the pandemic. This could perhaps offer a clue to explain the Latin American difference. Actually, the Latin American pandemic has been imported. If the three types of virus had different effects on the health condition of the infected persons, this might have produced a selection of one of them as the most frequent among flying passengers, possibly the mildest one and this might have been reflected in the outbreak.

Let us now move to the second part of this study and compare the current values of β, in order to have also an idea of the possible future developments of the pandemic.

For this we studied the evolution of the contagions from the day they reached a level N≈200 to 11 April 2020.

Table 3 presents this result for each country, accompanied by two additional informations, namely the day which corresponds to x=1 and the R^2 of the fit. We use red color for Latin American countries.

Table 3

Country	Day x=1	β	R^2
United States	March 4	0.0946	0.956
Spain	March 3	0.0785	0.916
Italy	February 23	0.0565	0.895
France	March 2	0.0716	0.961
Germany	March 2	0.0739	0.933
Iran	February 26	0.0485	0.856
United Kingdom	March 6	0.0780	0.980
Turkey	March 19	0.1025	0.937
Belgium	March 7	0.0680	0.966
Switzerland	March 6	0.0594	0.905
Netherlands	March 7	0.0610	0.953
Canada	March 14	0.0810	0.955
Brazil	March 16	0.0715	0.955
Portugal	March 14	0.0724	0.937
Austria	March 11	0.0576	0.894
Russia	March 19	0.0811	0.995
Chile	March 16	0.0630	0.951
Peru	March 18	0.0644	0.985
Ecuador	March 19	0.0579	0.924
Mexico	March 21	0.0592	0.989
Panama	March 21	0.0524	0.976
Dominican Republic	March 22	0.0540	0.950
Colombia	March 21	0.0509	0.977

The results of Table 3 show:

1- That the determination coefficients have very satisfactory values
2- that the β values of Latin American countries are not much dispersed, with an average value of 0.059
3- that the β values of the other countries are more dispersed, with an average value of 0.074

Nevertheless, it must be recalled that the analysis of Latin American data is based on a much smaller number of data than the other. This fact must be kept in mind for the following discussion.

We mentioned that β allows to calculate the duplication time at a certain moment. The study of the Chinese and South Korean development suggests that when β reaches certain typical values, this indicates that the outbreak is entering its final phase. Typical values are β = 0.03-0.04 to identify the beginning of the stabilization of the epidemic, and β ≈ 0.025 for few new contagions.

We fitted the data using as first day that indicated in Table 3, with different final day. In practically all cases, the decrease of β, as function of the variation of the final day is very well described as linear, β = bx + c. The results were always very good with the only exception of Peru, where the data of the last days seem to contradict the previous development.

Table 4 presents the results of the linear fits of the β values obtained keeping the first day fixed at the date indicated in Table 3 and with variable final day. In Table 4 we display the values of b and c with the R^2 of the fit. We also indicate the day corresponding to x=1, i.e. the earliest final day considered in the individual fit of logN, and finally the day when, according to the fit the value β= 0.025 would be reached.

Table 4

Country	**Day x=1**	**β**	**c**	**R^2**	**β =0.025**
United States	March 11	-0.0010	0.1346	0.7862	June 28
Spain	March 11	-0.0025	0.1568	0.984	May 3
Italy	March 11	-0.0017	0.1093	0.997	April 30
France	March 11	-0.0015	0.1178	0.994	May 12
Germany	March 9	-0.0017	0.1325	0.983	May 11
Iran	March 11	-0.0024	0.1129	0.912	April 17
United Kingdom	March 13	-0.007	0.1029	0.842	July 1
Turkey	March 27	-0.0051	0.1751	0.986	April 26
Belgium	March 15	-0.0013	0.1044	0.993	May 23
Switzerland	March 13	-0.0018	0.1139	0.958	May 1
Netherlands	March 15	-0.0014	0.1017	0.996	May 8
Canada	March 26	-0.0027	0.1208	0.999	April 23
Brazil	March 28	-0.0024	0.1047	0.954	April 30
Portugal	March 24	-0.0027	0.1168	0.999	April 27
Austria	March 24	-0.0027	0.1025	0.999	April 22
Russia	April 4	-0.0006	0.0903	0.877	July 22
Chile	March 28	-0.0022	0.0949	0.997	April 29
Peru	March 29	-0.00001	0.0623	Very poor	
Ecuador	March 28	-0.0030	0.0928	0.914	April 19
Mexico	March 30	-0.0013	0.0751	0.968	May 8
Panama	March 29	-0.0016	0.0715	0.979	April 27
Dominican Republic	March 30	-0.0027	0.0876	0.986	April 22
Colombia	March 30	-0.0014	00681	0.92	April 23

How seriously can we consider the predictions of the last column?

An idea may come from a similar prediction we made in ref. 4 and 14. These predictions were about when the data considered in those paper (written about one month ago the former and about two weeks ago the latter) would have predicted that β values of 0.03, 0.035 and 0.04 would be reached.

Table 5 presentes the predictions made in ref 14 and compares them with the results of Table 4.

Table 5

Country	**β=0.002 day (ref.8)**	**β=0.025day (this paper)**	**β=0.003 day (ref.8)**
Spain	April 29	May 3	April 23
Germany	May 21	May 11	May 14
Austria	April 21	April 22	April 17
UK	July 28	July 1	July 12
France	May 14	May 12	May 8
Switzerland	May 1	May 1	April 19
Belgium	May 10	May 23	May 3
Netherlands	May 15	May 8	May 6
Italy	April 27	April 30	April 22
Iran	April 5	April 17	April 2
South Korea	April 1	April 5 real	March 28

The agreement, given the approximations inherent this type of prediction seems to be rather satisfactory.

I am grateful to Drs. B.Pirouz and C. Rodriguez for many fruitful discussions and comments.

References

1. B. Pirouz, G. Violini, Why Italy and Iran? Why in some country COVID19 outbreak burst? DOI: 10.13140/RG.2.2.36010.36809/1

2. B. Pirouz, G. Violini, A novel approach to evaluate the asymptomatic COVID19 contagions DOI: 10.13140/RG.2.2.23633.07525/1

3. G. Violini, Desarrollo del contagio de Covid-19 e implicaciones para la República Dominicana, DOI: 10.13140/RG.2.2.31681.71522

4. G. Violini, Características Globales del Patrón de Evolución y Contención de la Pandemia COVID-19 DOI: 10.13140/RG.2.2.21615.38569

5. https://coronavirus.jhu.edu/map.html; https://coronavirus.politologue.com

6. G. Violini, ¿Entró en la etapa decreciente el contagio de Covid-19? DOI:10.13140/RG.2.2.14904.49924

7. G. Violini, Evolución comparativa de la pandemia en América Latina, DOI:10.13140/RG.2.2.21136.15368

8. G. Violini, ¿Es todavía posible controlar el desarrollo del COVID-19 en África y América? ¿Nos enseñan algo las experiencias italiana y coreana? DOI 10.13140/RG.2.2.32654.92483
9. S. Ferro, https://www.lavoce.info/archives/64469/quanti-sono-davvero-icontagiati-in-italia March 24, 2020

10. Mao Wang, Aili Jiang, Lijuan Gong, Lina Luo, Wenbin Guo, Chuyi Li, Jing Zheng, Chaoyong Li, Bixing Yang, Jietong Zeng, Youping Chen, Ke Zheng, Hongyan Li, Temperature significant change COVID-19 Transmission in 429 cities, **doi:** https://doi.org/10.1101/2020.02.22.20025791

11. C. R. Wells, P. Sah, S. M. Moghadas, A. Pandey, A. Shoukat, Yaning Wang, Zheng Wang, L. A. Meyers, B. H. Singer, and A. P. Galvani, Impact of international travel and border control measures on the global spread of the novel 2019 coronavirus outbreak, PNAS March 31, 2020 117 (13) 7504-7509

12. L. Cuppini, “È vero che in Lombardia si muore di più perché il virus è mutato?”, *Corriere della Sera*, 21 marzo 2020.

13. P. Forster, L. Forster, C. Renfrew, and M. Forster, Phylogenetic network analysis of SARS-CoV-2 genomes, https://doi.org/10.1073/pnas.2004999117

14. G, Violini, Duplicación del contagio explicada a mi nieto e intento de prever su evolución, DOI: 10.13140/RG.2.2.23099.77603

15. Pirouz, B.; Haghshenas, S,S.; Haghshenas, S,S.; Piro, P. Investigating a Serious Challenge in the Sustainable Development Process: Analysis of Confirmed cases of COVID-19 (New type of Coronavirus) Through a Binary Classification using Artificial Intelligence and Regression Analysis. Sustainability, 2020b, 12(6), https://doi.org/10.3390/su12062427

16. Jingyuan Wang , Ke Tang, Kai Feng and Weifeng Lv, High Temperature and High Humidity Reduce the Transmission of COVID-19, https://arxiv.org/ftp/arxiv/papers/2003/2003.05003.pdf

17. Behzad Pirouz, A. Golmohammadi, H. Saeidpour Masouleh, G. Violini, and Behrouz Pirouz, Relationship between Average Daily Temperature and Average Cumulative Daily Rate of Confirmed Cases of COVID-19, Medrxiv, doi.org/10.1101/2020.04.10.20059337

18. N. Yilmaz, E. Eren, Z. Kalayci, F. Saribek, C. Oz, Relationship between ultraviolet, hot and humidity with Covid -19outbreak https://www.researchgate.net/publication/340352717_Relationship_between_ultraviolet_hot_and_humidity_with_Covid_-19_outbreak .