

National Consumption of Antimalarial Drugs and COVID-19 Deaths Dynamics: An Ecological Study

5

Maxime Izoulet, CEMI-EHESS, Paris.

10 1. Introduction

COVID-19 (Coronavirus Disease-2019) is an international public health problem with a high rate of severe clinical cases. Several treatments are currently being tested worldwide. This paper focuses on anti-malarial drugs such as chloroquine or hydroxychloroquine, which have been currently reviewed by a systematic study as a good potential candidate¹ and that has been reported as the most used treatment by a recent survey of physicians². We compare the dynamics of COVID-19 daily deaths in countries using anti-malaria drugs as a treatment from the start of the epidemic versus countries that do not, the day of the 3rd death and the following 10 days. We show that the first group have a much slower dynamic in daily deaths than the second group. This ecological study is of course only one additional piece of evidence in the debate regarding the efficiency of anti-malaria drugs, and it is also limited as the two groups certainly have other systemic differences in the way they responded to the pandemic, in the way they report death or in their population that better explain differences in dynamics (systematic differences that may also explain their choice to rely on anti-

¹ Cortegiani A, Ingoglia G, Ippolito M, Giarratano A, Einav S, *A systematic review on the efficacy and safety of chloroquine for the treatment of COVID-19*. Journal of Critical Care, 10 march 2020, available online:

https://www.ncbi.nlm.nih.gov/pubmed/32173110?fbclid=IwAR3inu2KMU9p8a1z__S3ucf2WWeFrXKzazXFcIkthX6TcOakbqseDzgeB0c

² Richardson Valerie, *Hydroxychloroquine rated 'most effective therapy' by doctors for coronavirus: Global survey*. The Washington Times, 2 April 2020. Available online:

<https://www.washingtontimes.com/news/2020/apr/2/hydroxychloroquine-rated-most-effective-therapy-do/?fbclid=IwAR2e0CFgalMskxIYyAS2VFBmFGQeFyEyT3AtiASUT6O2FB9Xo5J3zOHKbcQ>

malaria drugs in the first place). Nevertheless, the difference in dynamics of daily deaths is so striking that we believe that the urgency context commands presenting the results before delving into further analysis. In the end, this data might ultimately be either a piece of evidence in favor of anti-malaria drugs or a stepping stone in understanding further what other ecological aspects place a role in the dynamics of COVID-19 deaths.

2. Method

In this study, we set up two groups of 16 countries and study the dynamics of the number of deaths between the day of the 3rd death and the following 10 days. The first group is made up of countries that we know use or produce chloroquine or hydroxychloroquine on a massive scale during this period. The second group consists of countries that did not use or produce chloroquine or hydroxychloroquine in large quantities during the period under consideration. When we calculate the averages of each of the two groups, we find very marked differences in their temporal dynamics (see results).

We then use Box and Jenkins' methodology to apply ARIMA (Auto Regressive Integrated Moving Average) models to these time series, compare the model parameters obtained for each group of countries, and make forecasts of the means of the two groups from these results. Unsurprisingly, the ARIMA models predict a stabilization of the number of deaths for the group of countries using chloroquine and a large increase for the group of countries not using it. The 60 countries most affected by the epidemic (in terms of number of cases) were studied one by one in descending order to determine whether or not they were conducting a national strategy for the large-scale use or production of chloroquine at the beginning of the epidemic in the country (around the 3rd death)³. If there was no evidence of such a strategy, or even if sources indicated a strategy to the contrary, the country was classified in the "control group" group, until a panel of 16 countries was obtained in order to have a large sample, provided that daily death data were available for the 10 days following the third death. The second group was constituted with the 16 countries among the 60 most affected in terms of number

³ List available at the 13th of April (see detailed list in annex): <https://www.worldometers.info/coronavirus/#countries>

of cases for which sources indicate the massive use or production of chloroquine at the beginning of the epidemic in the country (around the 3rd death), provided that they have daily death data for the 10 days following the 3rd death. The different groups of countries were constituted according to the information available in the international press on their use or mass production of such drugs over the period under consideration. 16 countries thus constitute each of the two groups (see figure 1 and figure 2).

For each of the two groups, the number of daily deaths is noted each day from the 3rd death in the country and the following 10 days. Then the average of the daily deaths is established for each day for each group of countries. For the group without chloroquine, an average is also calculated by removing China and another by removing China and Spain, as these two countries have the two most explosive time series and may be seen as outliers. The trends do not change substantially.

3. Results

The graphical projection of the mean curves indicates a divergence in the dynamics of the daily death curves of the two groups of countries which is very clear for the period studied (i.e. from the beginning of the epidemic) (see figure 3).

The average curve for countries using antimalarial drugs is rather stable or slightly increasing, the curve for countries not using those treatments is on the contrary strongly increasing. Moreover, the simple regression curves clearly indicate this difference in trend. The average of countries with widespread chloroquine use is fairly well modelled ($R^2 = 0,73$) by a slightly ascending polynomial regression, whereas the average of countries without chloroquine is very well modelled ($R^2 = 0,98$) by an exponential regression. Modelling and forecasting using ARIMA (Auto Regressive Integrated Moving Average) models are widely used in time series econometrics. Introduced by Box and Jenkins, they allow an excellent modelling of time series based on the data themselves and without including any theoretical *a priori* on these data. They therefore allow excellent modelling of the internal dynamics of these data and are highly predictive, which tends to validate their relevance. They are widely used in macroeconomics and finance, but also in many other fields, in biology, geophysics, astronomy, etc...

Let's say an ARIMA (p,d,q) process:

$$(1.1) \quad (1 - L)^d X_t = \frac{\theta(L)}{\phi(L)} \varepsilon_t$$

With $\phi_i (i = 1, \dots, p)$ the reals corresponding to the autoregressive coefficients, $\theta_j (j = 1, \dots, q)$ the reals corresponding to the moving average coefficients, of the order of integration d and $(\varepsilon_t \sim WN(0, \sigma^2))$ the residuals behaving as white noise, with zero mean and variance σ^2 , constant and less than infinity.

Following Box and Jenkins' methodology for specifying, estimating and validating the ARIMA modelling, the application to the mean time series of the two groups of countries using the R⁴ software gives the results of **the figure 4**.

The Akaike Information Criterion (AIC), for each model selected, is the best relative to other alternative models that were also calculated in this study, i.e. it is closest to zero, indicating the quality of the model specification. This criterion is calculated as follows:

$$(1.2) \quad AIC = \log \widehat{\sigma_\varepsilon^2} + \frac{2(p+q)}{T}$$

This modelling then allows a 10-day forecast of the evolution of the death dynamics for each of the two groups of countries. We obtain the results in *R* (the first column shows the number of days after the first day with 3 deaths, the second column shows the estimated forecast values, the third and fourth columns show the low and high values of the 80% confidence intervals of the forecast, and the fifth and sixth columns show the 95% confidence intervals)

-For the "antimalarial drugs group", see **figure 5 and figure 6**.

-For the group "without chloroquine", see **figure 7 and figure 8**.

⁴ Hyndman R. *et al.*, *Forecasting Functions for Time Series and Linear Models*, 2020. <https://cran.r-project.org/web/packages/forecast/forecast.pdf>

Forecasts reinforce early visual observations. For the group of countries "antimalarial drugs group", the forecast of the ARIMA model (1,0,1) indicates a stabilization of the death curve. For the "control group" countries, the ARIMA model's forecast (0,2,0) indicates a very significant acceleration in the number of deaths. It should be noted that beyond D+10, such an acceleration is already visible in the actual data of many countries for which this statistic is available.

To validate model's specification, residuals distribution is then tested, in order to control they behave as a white noise, i.e. they are not autocorrelated. This verification is done through the autocorrelations of residuals plotting in R .

Autocorrelation function is a X_t process of k order that can be writing as follow :

$$(1.3) \quad \rho_k = \frac{\frac{1}{T} \sum_{t=1}^{T-k} (X_t - \bar{X})(X_{t+k} - \bar{X})}{\frac{1}{T} \sum_{t=1}^T (X_t - \bar{X})^2}$$

For ARIMA (1,0,1) applied to "antimalarial drugs group", we obtain autocorrelations of **the figure 9**.

No autocorrelation is significant, residuals are behaving as a white noise, it indicates the validity of the model.

For ARIMA (0,2,0) applied to "control group", we obtain autocorrelations of **the figure 10**.

In the same way, no residuals autocorrelation is significant. Residuals are behaving as a white noise, model specification and estimation is then validated.

4. Limitations

Introduced in the 1970s by Box and Jenkins⁵, ARIMA models are so-called a-theoretical models⁶, which seek predictive efficiency by focusing on the past data of a time series, without worrying about the causes of these past data. They are therefore not able to explain all the explanatory variables of a temporal evolution, but they are very effective in describing the internal dynamics of the evolution. Nor are they an instrument of proof, but rather a statistical index updating a dynamic. Here they make it possible to highlight two very distinct dynamics from the very first days of the outbreak, which is very useful since this highly contagious epidemic has a strong internal dynamic. They have been already used for modelling the spread of the epidemic, notably in India⁷.

Of course, they do not model, and do not claim to model, all the parameters that explain a temporal evolution. On the other hand, they are often highly predictive⁸ and outweigh many models with more explanatory variables, which is a very important criterion of overall model validity. It should also be noted that while many sources exist to determine the health action of governments, including their use or mass production of chloroquine from the onset of the crisis, quantitative data are lacking and do not allow for more in-depth temporal analyses and causality tests. There also might be systematic differences between the two groups - in particular political differences, urban differences or differences in other strategy aspects such as testing. There is strong evidence for places like South Korea and Japan that mass testing is an effective strategy to control the epidemic, and our study might be a proxy for testing strategies. All these aspects should be examined in a late study.

⁵ Box, George; Jenkins, Gwilym (1970). *Time Series Analysis: Forecasting and Control*. San Francisco: Holden-Day, 1970.

⁶ Gujarati Damodar N., *Econométrie*, Traduction de la quatrième édition américaine par Bernard Bernier, Collection Ouvertures économiques, éditions de Boeck, 2012.

⁷ Choudhary Ishan, *Forecasting COVID-19 cases in India: How many cases are going to get detected by 7th April 2020 in India?* Medium, 29 march 2020, link: <https://towardsdatascience.com/forecasting-covid-19-cases-in-india-clc410cfc730>

⁸ Lardic S. et Mignon V. : *Econométrie des séries temporelles macroéconomiques et financières*, Economica, 2002.

180

5. Conclusion

185

190

We find major differences in death rates, with countries using antimalarial drugs faring better than those which do not. This analysis is of course only one additional piece of evidence in the debate regarding the efficiency of anti-malaria drugs, and it is also limited as the two groups certainly have other systemic differences in the way they responded to the pandemic. Nevertheless, the differences in dynamics is so striking that we believe that the urgency context commands presenting this ecological study before delving into further analysis. In the end, this data might ultimately be either a piece of evidence in favor of anti-malaria drugs or a stepping stone in understanding further what other ecological aspects place a role in the dynamics of COVID-19 deaths.

List of figures

195

	Number of Daily Deaths										
Antimalarial drugs group	D=Day of the third death	D+1	D+2	D+3	D+4	D+5	D+6	D+7	D+8	D+9	D+10
Algeria	1	1	0	1	2	2	2	4	2	0	2
Dominican Republic	1	0	0	3	4	0	10	8	11	3	9
Egypt	2	2	0	1	1	2	4	5	1	1	3
Greece	2	1	0	1	0	1	4	3	2	2	3
India	1	0	1	1	0	2	3	0	2	8	0
Indonesia	3	1	0	0	2	12	6	7	6	10	1
Malaysia	1	5	2	4	2	4	3	3	1	8	2
Morocco	1	0	1	0	1	1	5	12	2	1	7
Pakistan	1	0	2	1	1	1	1	2	1	2	7
Panama	2	0	3	0	2	1	5	3	7	3	3
Russia	2	0	1	0	4	1	8	7	6	4	9
Serbia	1	3	1	2	3	3	7	5	3	8	5
South Korea	4	2	3	1	1	3	1	4	7	4	3
Tunisia	2	0	1	1	0	2	1	0	1	1	2
UAE	1	2	1	2	0	1	1	0	1	1	0
Ukraine	1	0	0	0	0	0	2	0	0	4	1
Overall	26	17	16	18	23	36	63	63	53	60	57
Mean by country	1,6	1,1	1,0	1,1	1,4	2,3	3,9	3,9	3,3	3,8	3,6

Figure 1: Number of daily deaths after day with 3 deaths, “antimalarial drugs group”.

	Number of Daily Deaths										
Control group	D=Day of the third death	D+1	D+2	D+3	D+4	D+5	D+6	D+7	D+8	D+9	D+10
Austria	2	1	0	2	0	2	8	5	7	3	18
Belgium	3	0	0	1	0	6	0	4	7	16	30
Brazil	3	3	4	7	7	9	12	13	18	15	22
Canada	3	4	1	3	0	7	1	4	2	10	3
China	8	16	15	24	26	26	38	43	46	45	57
France	1	1	0	3	2	7	3	11	3	15	13
Germany	1	3	2	1	4	4	9	2	16	24	16
Iran	2	2	2	4	4	3	7	8	9	11	12
Italy	1	4	4	1	5	4	8	12	11	27	28
Netherlands	2	1	0	1	0	5	2	8	4	19	15
Spain	1	5	2	7	13	6	19	31	47	63	98
Sweden	1	4	1	2	1	5	4	1	6	13	22
Switzerland	1	1	3	4	2	1	5	8	6	10	13
Turkey	2	5	12	9	7	7	15	16	17	16	23
United Kingdom	1	2	1	2	2	1	10	14	20	16	33
United States	5	3	2	1	3	4	3	4	4	8	3
Overall	37	55	49	72	76	97	144	184	223	311	406
Mean by country	2,3	3,4	3,1	4,5	4,8	6,1	9,0	11,5	13,9	19,4	25,4
Mean without China	1,9	2,6	2,3	3,2	3,3	4,7	7,1	9,4	11,8	17,7	23,3
Mean without China and Spain	2,0	2,4	2,3	2,9	2,6	4,6	6,2	7,9	9,3	14,5	17,9

Figure 2: Number of daily deaths after day with 3 deaths, “control group”.

200

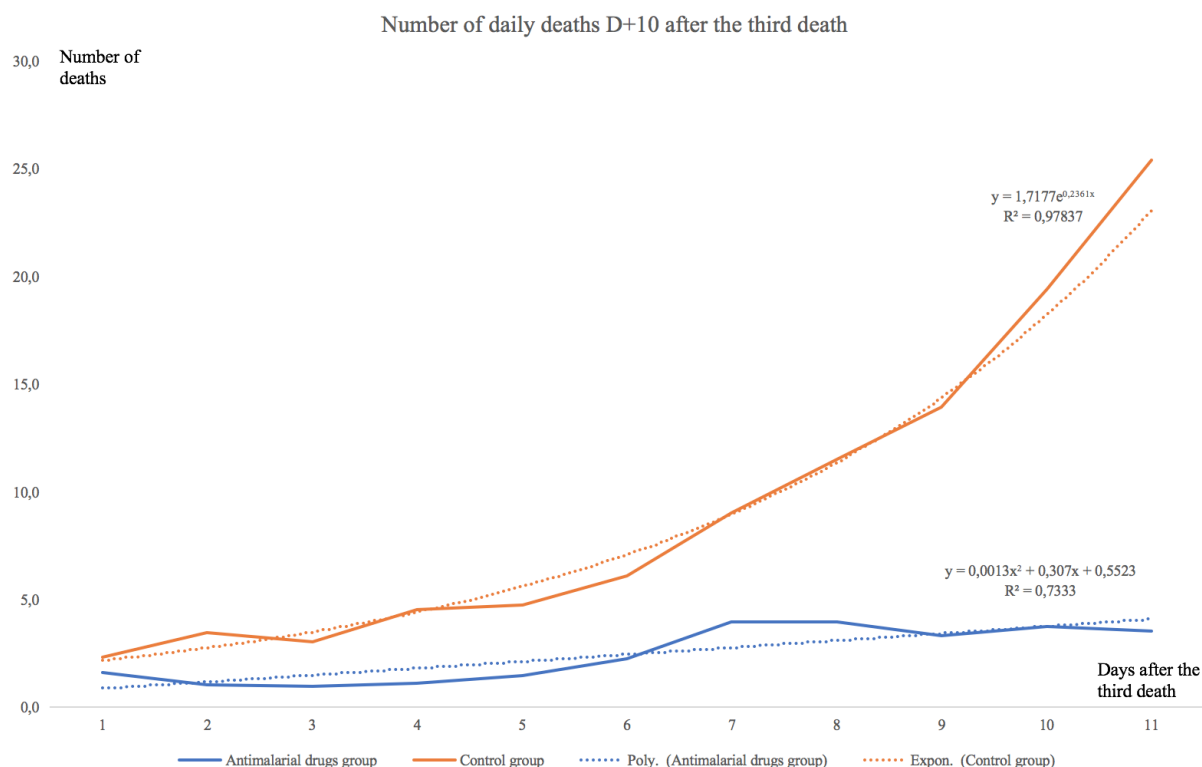


Figure 3: Means of the number of daily deaths for each group

	Selected models	ϕ_1	d	θ_1	σ^2	Log likelihood	AIC
Antimalarial drugs group	ARIMA (1,0,1)	0,6501	0	1	0,2077	-8,95	25,9
Control group	ARIMA (0,2,0)	-	2	-	2,183	-16,28	34,57

Figure 4: ARIMA parameters specifications and estimations for each group


```
> forecast(arima(antimalarialdrugsgroup,order=c(1,0,1)))
      Point Forecast      Lo 80      Hi 80      Lo 95      Hi 95
12      3.023685  2.414098  3.633271  2.0914030  3.955966
13      2.839780  1.707117  3.972444  1.1075207  4.572040
14      2.720230  1.428703  4.011758  0.7450094  4.695451
15      2.642515  1.289449  3.995581  0.5731786  4.711851
16      2.591994  1.213749  3.970240  0.4841493  4.699839
17      2.559153  1.170404  3.947902  0.4352443  4.683062
18      2.537804  1.144640  3.930968  0.4071433  4.668464
19      2.523925  1.128900  3.918951  0.3904182  4.657433
20      2.514904  1.119092  3.910715  0.3801944  4.649613
21      2.509039  1.112896  3.905182  0.3738219  4.644256
```

Figure 5: Forecasting values and confidence intervals for an ARIMA (1,0,1) process applied to “antimalarial drugs group”

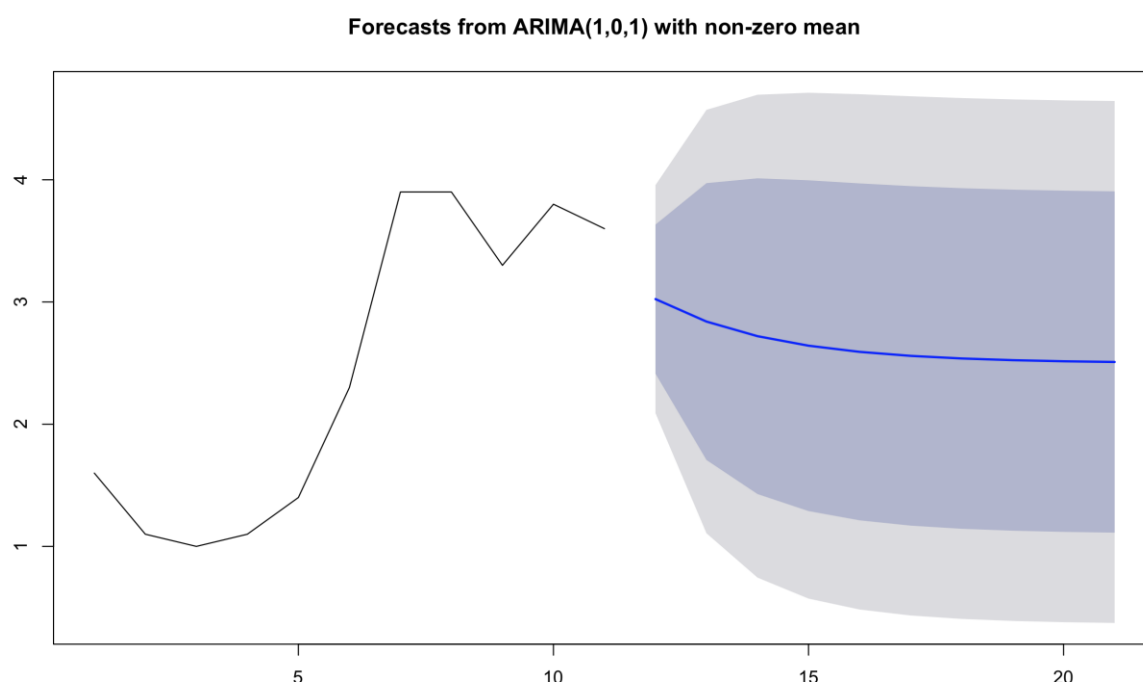


Figure 6: Forecasts plot for “antimalarial drugs group”

```
> forecast(arima(controlgroup, order=c(0,2,0)))
```

	Point Forecast	Lo 80	Hi 80	Lo 95	Hi 95
12	31.4	29.50637	33.29363	28.50394	34.29606
13	37.4	33.16571	41.63429	30.92420	43.87580
14	43.4	36.31467	50.48533	32.56392	54.23608
15	49.4	39.02814	59.77186	33.53761	65.26239
16	55.4	41.35643	69.44357	33.92222	76.87778
17	61.4	43.33588	79.46412	33.77331	89.02669
18	67.4	44.99422	89.80578	33.13331	101.66669
19	73.4	46.35349	100.44651	32.03594	114.76406
20	79.4	47.43177	111.36823	30.50882	128.29118
21	85.4	48.24421	122.55579	28.57513	142.22487

225 **Figure 7: Figure 5: Forecasting values and confidence intervals for an ARIMA**
(0,2,0) process applied to “control group”

230

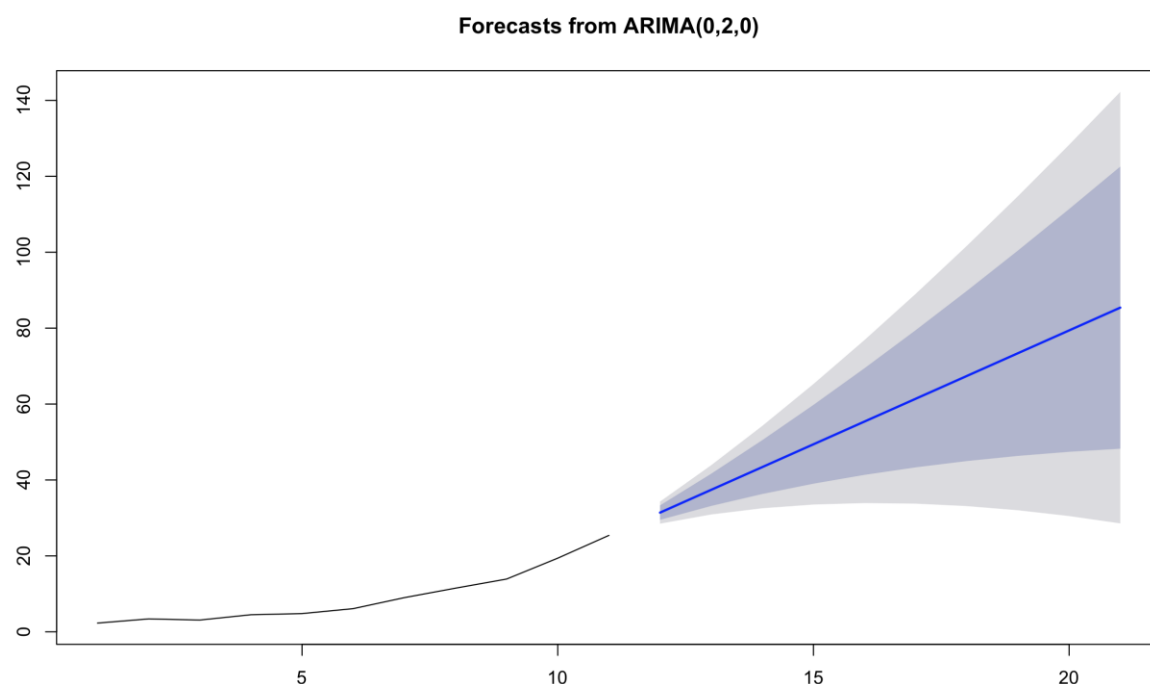


Figure 8: Forecasts plot for “control group”

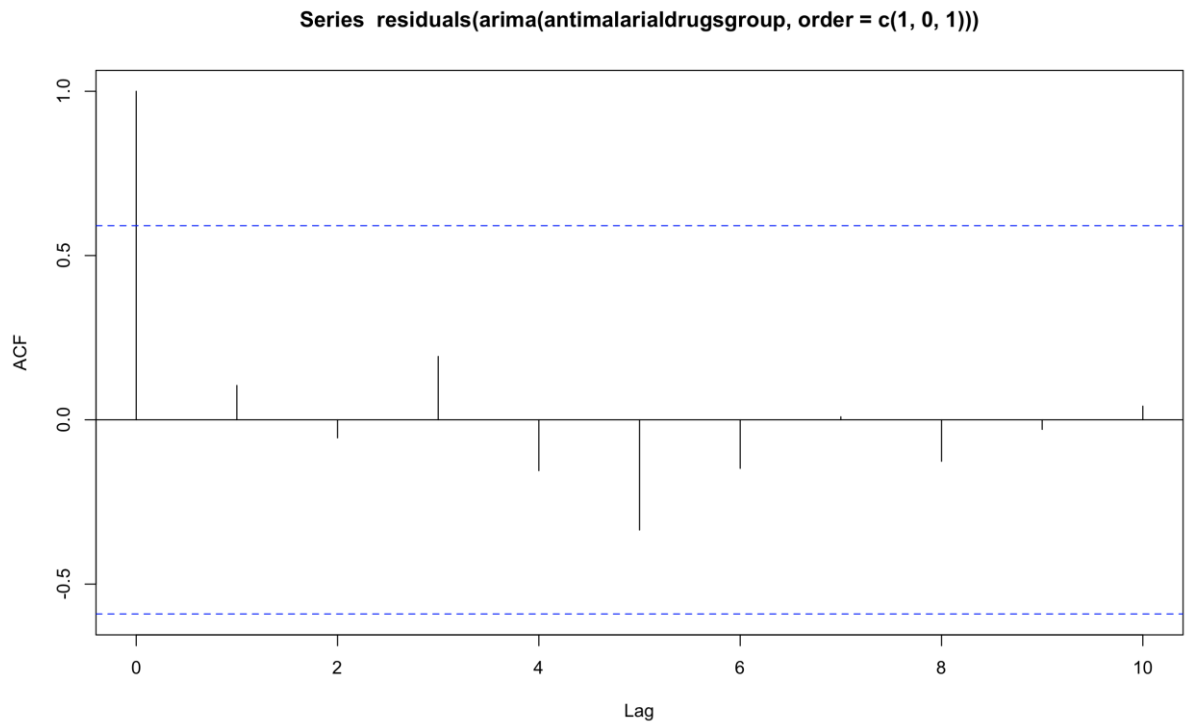


Figure 9: Autocorrelations of residuals for ARIMA (1,0,1) applied to “antimalarial drugs group”

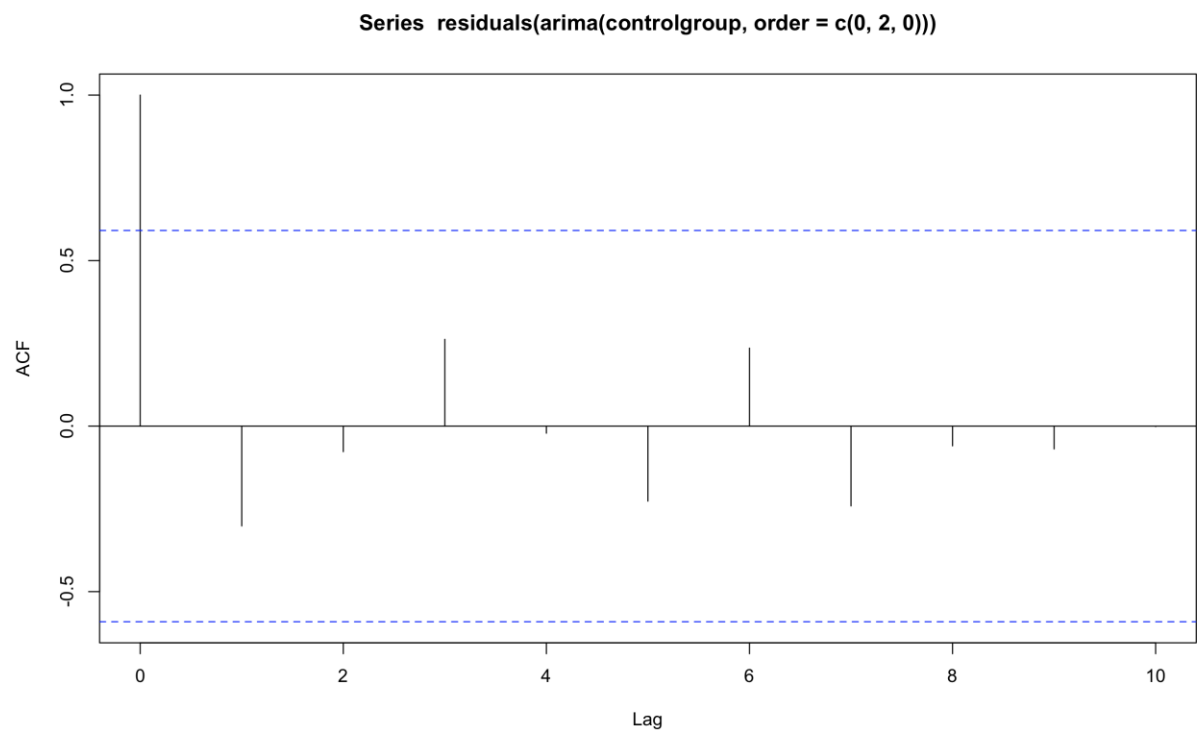


Figure 10: Autocorrelations of residuals for ARIMA (0,2,0) applied to “control group”

Bibliography:

245

Box, George; Jenkins, Gwilym (1970). *Time Series Analysis: Forecasting and Control*.
San Francisco: Holden-Day, 1970.

Charpentier A., *Différencier (indéfiniment) les séries temporelles ?*
<http://freakonometrics.hypotheses.org/1876>, 2009.

250

Choudhary Ishan, *Forecasting COVID-19 cases in India: How many cases are going to
get detected by 7th April 2020 in India?* Medium, 29 march 2020, link:
[https://towardsdatascience.com/forecasting-covid-19-cases-in-india-
c1c410cfc730](https://towardsdatascience.com/forecasting-covid-19-cases-in-india-c1c410cfc730)

255

Gujarati Damodar N., *Econométrie*, Traduction de la quatrième édition américaine par
Bernard Bernier, Collection Ouvertures économiques, éditions de Boeck, 2012.
Granger C.W.J., *Time series analysis, cointegration and applications*, Nobel Lecture,
December 8, 2003.

Hamilton J. D., *Time Series Analysis*, Princeton University Press, 1994.

260

Hayek F.A., *La théorie des phénomènes complexes*, in *Essais de philosophie, de science
politique et d'économie*, traduction française Christophe Piton, Bibliothèque
classique de la liberté, Paris, 2007.

Lardic S. et Mignon V. : *Econométrie des séries temporelles macroéconomiques et
financières*, Economica, 2002.

265

Mandelbrot, B., Hudson R. L., *The (Mis)Behaviour of Markets: A Fractal View of Risk,
Ruin and Reward*, Profile Book, 2008.

Mullard A., *Drug repurposing programmes get lift off*
Nat Rev Drug Discov, 11 (2012), pp. 505-506, 10.1038/nrd3776.

Pfaff B., *VAR Modelling*, 2020. <https://cran.r-project.org/web/packages/vars/vars.pdf>

270

Hyndman R. *et al.*, *Forecasting Functions for Time Series and Linear Models*, 2020.
<https://cran.r-project.org/web/packages/forecast/forecast.pdf>

Veenstra J.Q., *Fractional ARIMA (and Other Long Memory) Time Series Modeling*,
2018. <https://cran.r-project.org/web/packages/arfima/arfima.pdf>

Zhou P., Yang X.L., Wang X.G. , Hu B. , Zhang L. , Zhang W. , *et al.* *A pneumonia
outbreak associated with a new coronavirus of probable bat origin*

275

Nature (2020 Feb 3), 10.1038/s41586-020-2012-7

Annexes

280

All	Europe	North America	Asia	South America	
Country, Other	Total Cases	New Cases	Total Deaths	New Deaths	
World	1,866,654	+14,397	115,269	+1,075	
USA	560,433	+133	22,115	+10	
Spain	169,496	+2,665	17,489	+280	
Italy	156,363		19,899		
France	132,591		14,393		
Germany	127,854		3,022		
UK	84,279		10,612		
China	82,160	+108	3,341	+2	
Iran	73,303	+1,617	4,585	+111	

<u>Turkey</u>	56,956		1,198	
<u>Belgium</u>	30,589	+942	3,903	+303
<u>Netherlands</u>	26,551	+964	2,823	+86
<u>Switzerland</u>	25,503	+88	1,117	+11
<u>Canada</u>	24,383		717	
<u>Brazil</u>	22,318	+126	1,241	+18
<u>Russia</u>	18,328	+2,558	148	+18
<u>Portugal</u>	16,934	+349	535	+31
<u>Austria</u>	13,998	+53	368	+18
<u>Israel</u>	11,235	+90	110	+7
<u>Sweden</u>	10,948	+465	919	+20
<u>S. Korea</u>	10,537	+25	217	+3
<u>Ireland</u>	9,655		334	
<u>India</u>	9,240	+35	331	
<u>Peru</u>	7,519		193	
<u>Ecuador</u>	7,466		333	
<u>Japan</u>	7,370		123	
<u>Chile</u>	7,213		80	

<u>Poland</u>	6,674		232	
<u>Romania</u>	6,633	+333	318	+2
<u>Norway</u>	6,525		128	
<u>Australia</u>	6,359	+46	61	+2
<u>Denmark</u>	6,318	+144	285	+12
<u>Czechia</u>	5,991		139	+1
<u>Pakistan</u>	5,374	+144	93	+2
<u>Philippines</u>	4,932	+284	315	+18
<u>Malaysia</u>	4,817	+134	77	+1
<u>Mexico</u>	4,661	+442	296	+23
<u>Indonesia</u>	4,557	+316	399	+26
<u>Saudi Arabia</u>	4,462		59	
<u>UAE</u>	4,123		22	
<u>Serbia</u>	3,630		80	
<u>Panama</u>	3,400		87	
<u>Luxembourg</u>	3,281		66	
<u>Qatar</u>	3,231	+252	7	
<u>Ukraine</u>	3,102	+325	93	+10

290

295

<u>Finland</u>	3,064	+90	59	+3
<u>Dominican Republic</u>	2,967		173	
<u>Belarus</u>	2,919	+341	29	+3
<u>Colombia</u>	2,776		109	
<u>Thailand</u>	2,579	+28	40	+2
<u>Singapore</u>	2,532		8	
<u>Argentina</u>	2,208	+66	95	+5
<u>South Africa</u>	2,173		25	
<u>Greece</u>	2,114		98	
<u>Egypt</u>	2,065		159	
<u>Algeria</u>	1,914		293	
<u>Morocco</u>	1,746	+85	120	+2
<u>Iceland</u>	1,701		8	
<u>Moldova</u>	1,662		35	+4
<u>Croatia</u>	1,650	+50	25	+2
<u>Hungary</u>	1,458	+48	109	+10

Annex 1: list of the 60 countries with the most Covid 19 total cases on 13th April 2020
(see note 3 of this study) from <https://www.worldometers.info/coronavirus/#countries>