Understanding Elementary Statistics on COVID-19: What do they measure? & How to use them?

Roger R. Betancourt, Professor of Economics, Emeritus, UMD* (December 25, Major revision)

The Statistics and What They Measure

1) IR= Infection rate = total or cumulative # of Covid-19 cases per million persons in the population

IR measures how many persons have contracted the virus relative to its total population at a point in time.

Usefulness, comparing any two countries (or areas) at similar times after identification of the first few cases. For instance, 30 days after the first case. It measures the rate of spread of the virus at a given distance in time from its arrival.

2) TR = Testing rate= cumulative # of COVID -19 tests performed per million persons in the population

TR measures the number of COVID-19 tests performed at any point in time relative to the number of persons in the population.

Usefulness, not much by itself. Essential in conjunction with the infection rate to generate.

3) POR = Positivity Rate = IR/TR= total # of COVID-19 cases relative to the number of tests given.

POR measures how fast COVID-19 is spreading in a country at a point in time, especially accurate when multiple tests to same person excluded from TR. It is still useful as an estimate even if multiple tests to same person not excluded, because we know it biases positivity rate downwards.

Comments based on use experience since August 5. First, it is a very useful leading indicator of the state of the pandemic with respect to the transmission of the infection within any spatial unit. Second, by using the cumulative figures for both infections and testing evaluated every two weeks, it is a very robust indicator of the actual state of the pandemic spread at that point in time, in contrast to daily comparisons or even weekly moving average comparisons. Third, the results show that keeping this rate substantially below 1% as well as stable or decreasing was the most effective tool in preventing the worst welfare outcome of the pandemic, namely deaths per capita. This was true for the analysis focusing on Australia and New Zealand relative to the US, the UK and Canada between September 4 and December 11 and corroborated by other experiences such as China relative to Europe and, to a lesser extent, even relative to Japan and South Korea.

4) DR= Death Rate = cumulative number of COVID-19 deaths per million persons in the population.

DR measures how many people have died from the virus relative to a country's total population at a point in time.

Usefulness, comparing most damaging unsuccessful outcome between countries or, more generally, spatial units at any point in time.

Comments based on use experience since August 5. First, it is a very useful indicator of the most damaging direct welfare consequence of the pandemic at any point in time for any spatial unit. Second, it is a robust indicator of this welfare consequence of the pandemic across spatial units at the same points in time relative to the beginning of the pandemic. For instance, at the end of the pandemic or at the start and end of first or second waves, however defined, as long as the definitions keep the time-period of evaluation fixed such as weekly or biweekly. Third, it generates observable characteristics identifying population subgroups experiencing different incidence of this important welfare consequence in the population. The latter is useful for both health, social and economic policy purposes. Finally, it is a very useful lagging indicator of the state of the pandemic.

5) $\%\Delta DR = [(DR_t - DR_{t-1})/DR]100$ = percentage change in the cumulative death rate per capita over the last two- week period.

% ΔDR measures the state of the pandemic in the last two- week period by looking at the rate of change in its main lagging indicator.

Usefulness, it provides a dynamic measure (every two weeks) of the evolution of the state of the pandemic in a spatial unit at a point in time as a result of prior spread, behavior and policy actions wrt to implementation of both Non-Pharmaceutical Interventions (NPI's) and/or Pharmaceutical interventions (PI's).

Comments based on use experience since August 5. I found the earlier statistic (HSPR = DR/IR) mechanically useful but interpretatively almost useless because it was affected by too many variables. First, it would depend on the basic determinants of the rate of spread such as time of introduction into the spatial unit, population size and density, as well as spatial distribution of virus mutations within the spatial unit (Ives & Bozzuto 2020, p.19 and p. 28; https://doi.org/10.1101/2020.06.18.20134700). Incidentally, the latter reference relies on the death rate per capita across US counties as its basic data in estimating the rate of spread of the virus in its original state, i.e., prior to any intervention by either the population or policy makers. Second, it would also depend on policies implemented to deal with the pandemic, quality of medical facilities and personnel as well as on habits and attitudes of the population wrt to masks, social distancing, testing and contract tracing. Explaining variations in this statistic would be far more appropriate for a professional research project than for an educational one.

Starting on November 11, I began to notice that variations in the percentage change of two- week per capita death (% ΔDR) were useful in understanding the state of the pandemic at a point in time in different countries, especially in conjunction with the change in the two- week positivity rates. Thus, I decided to change to this indicator as a far more useful elementary statistic in the context of an educational memo and in view of the looming introduction of experimental vaccines and treatments as pharmaceutical interventions in the pandemic.

Additional Comments on How to Use Them.

The earlier biweekly comments on how to use the positivity rates and the death rates are still valid, applicable and available on the development section of my home page at (http://econweb.umd.edu/~betancourt/development.htm) in the file titled "Understanding Elementary Pandemic Statistics.Final Form.Sept/Dec.2020". The data source continues to be the same, Worldometer; the basic countries analyzed in detail continue to be the same (US, UK, Canada, Australia and New Zealand). The additional 13 countries added for robustness to check the analysis as well as the two US states (MD and FLA) continue to be the same. The main change henceforth is that the third statistic reported every biweekly period in the table is the percentage change in the last two weeks of cumulative deaths per million persons in the population of the spatial unit, % ΔDR as defined above. In order to preserve continuity, I have replaced the HSPR measure with the %ΔDR in the first two tables of the earlier file (starting whenever feasible from the prior data) and included them here. The detailed new analysis here starts with the two- week period ending on December 25. The data for that date starts on the third table (which is to be included on or after December 25, of course).

Main Insights of August to December Analysis Useful for the Design of Future Exercises or Projects.

1. The use of cumulative infection and testing rates to construct positivity rates provides a robust elementary statistic to evaluate the effectiveness of policies during the evolution of the pandemic. In the relative early stages covered by the period March-December, it shows the superiority of a vigorous approach to prevent spread by keeping the positivity rate substantially below unity. Moreover, the main policies to prevent spread during these early stages are NPI's.
2. The use of cumulative deaths per capita provides a robust elementary statistic to evaluate the main direct welfare outcome of the pandemic across spatial units during the evolution of the pandemic. In the relatively early stages covered by this period, it shows that positivity rates above unity generate a variety of possible trends for this important outcome variable that include reversals of fortune across countries with respect to its level.

Potential Additional Insights of Future Analysis

1. We expect similar insights to arise over time from the positivity rates and the death rates per capita due to more widespread use of recently introduced new pharmaceutical treatments, e.g., such as Regeneron, and the introduction of recently developed vaccines over the next few months such as Pfizer, Moderna and Astro Zeneca.
2. Furthermore, the use of percentage changes in cumulative deaths per capita levels over the last two- week period provides a dynamic elementary statistic to evaluate the effectiveness of policies during the evolution of the pandemic that should yield new insights in two different dimensions. By itself, it shows the recent impact on the main direct outcome of the pandemic of prior policies of all types and a population's behavioral responses to these policies at a point in time. In conjunction with the trend of prior positivity rates, it suggests acceleration or deceleration in the state of the pandemic through its own trends due to its role as a lagging indicator of the state of the pandemic. Moreover, in this second dimension, it might yield insights on the direct impact of treatments and vaccines (through death rates) as well as on their indirect impact (through positivity rates) with respect to the evolution of the pandemic given the behavioral responses of a population to these PI policies, including moral hazard.

What do we learn from the December 25 updates in the context of this major revision and recent pandemic related developments?

First, something we already knew that is worth reiterating more explicitly. The statistics presented here based on cumulative data are complementary not substitutes for the ones presented in the media based on current new infections, testing and deaths per capita. The cumulative ones are relevant for policy evaluation over the evolution of the pandemic; the current ones are relevant for crisis management at a particular point in time.

*Exchanges with Juan Belt, Julio Betancourt, Chris Clague, Allan Drazen, Beatriz Hardy, Tanjim Hossain, Alicia Juarrero, Luis R. Luis, Paul Meyer and José Ramón de la Torre have substantially increased my confidence in the robustness and educational usefulness of the comparison exercises explicitly illustrated here. I am grateful for these comments. Any remaining issues that may arise with the data and analysis above are the sole responsibility of the author.

Country*/State September 04- October 16, 2020, December 25 Rev.

Country*/State	POR=IR/TR				DR				%ΔDR (L2wks)				% POP>65^
	9/04	9/18	10/02	10/16	9/04	9/18	10/02	10/16	9/04	9/18	10/02	10/16	'19
USA	.075	.072	.069	.067	577	610	642	672	n.a.	5.7	5.2	4.7	15
UK	.019	.019	.019	.023	611	614	621	637	n.a.	0.5	1.1	2.6	17
Canada	.070	.022	.022	.023	242	243	246	256	n.a.	0.4	1.2	4.1	16
Australia	.004	.004	.004	.003	029	033	035	035	n.a.	13.8	6.1	0.0	15
N. Zealand	.002	.002	.002	.002	005	005	005	005	n.a.	0.0	0.0	0.0	15
Spain	.053	.061	.061	.064	625	650	684	718	n.a.	0.4	5.2	5.0	17
Italy	.031	.021	.028	.029	587	590	594	602	n.a.	0.5	6.8	1.3	21
France	.040	.042	.052	.064	470	476	490	507	n.a.	1.3	2.9	3.5	18
Germany	.020	.019	.017	.019	112	113	114	117	n.a.	0.9	0.9	2.6	19
Denmark	.007	.007	.007	.008	108	110	112	117	n.a.	1.9	1.8	4.5	19
Japan	.045	.043	.040	.038	010	012	012	013	n.a.	0.2	0.0	8.3	25
S. Korea	.010	.010	.010	.010	006	007	008	009	n.a.	16.6	14.3	12.5	13
China**	.001	.001	.001	.001	003	003	003	003	n.a.	0.0	0.0	0.0	10
Cuba	.010	.010	.009	.008	009	010	011	011	n.a.	11.1	10.0	0.0	14
Jamaica	.048	.062	.083	.093	010	017	037	054	n.a.	70. 0	118.	45.9	9
Brazil	.281	.305	.271	.289	586	634	680	716	n.a.	8.2	7.3	5.3	8
Chile	.166	.149	.138	.128	597	637	669	701	n.a	6.7	5.0	4.8	10
E. Salvador	.080	.076	.073	.071	115	124	131	140	n.a.	7.8	5.6	6.9	8
Maryland	9/26→	.049	.048	.045	→	648	653	666	→	n.a.	7.7	2.0	16^
Florida	9/26→	.133	.133	.131!	→	648	673	733	→	n.a.	3.9	8.9	21^

*All country/state data taken from Worldometer; ^ World Bank data for 2019; Census Estimates 2020.

**The actual POR rates are .00053 for all dates; rounded up to.001 for ease of presentation. ! figure corrected on 10/29/2020 (from .171)

Country*/State October 30 – December 11, December 25 Rev.

Country*/State	POR=IR/TR				DR				%ΔDR (L2wks)				%pop.> 65^
	10/30	11/13	11/27	12/11	10/30	11/13	11/27	12/11	10/30	11/13	11/27	12/11	‘19
USA	.060	.067	.071	.075	706	749	813	903	5.1	6.1	8.5	11.1	15
UK	.029	.034	.037	.038	676	749	838	927	8.9	10.8	11.9	10.6	17
Canada	.023	.028	.032	.036	266	284	312	346	3.9	6.8	9.9	10.9	16
Australia	.003	.003	.003	.003	035	035	035	035	0.0	0.0	0.0	0.0	15
N. Zealand	.002	.002	.002	.002	005	005	005	005	0.0	0.0	0.0	0.0	15
Spain	.074	.077	.075	.072	762	865	949	1,012	6.1	13.5	9.7	6.6	17
Italy	.040	.059	.071	.075	631	721	875	1,036	4.8	14.3	21.4	18.4	21
France	.082	.102	.108	.083	551	658	780	871	8.7	19.4	18.5	20.8	18
Germany	.023	.030	.036	.042	124	146	188	253	6.0	17.7	28.8	34.6	19
Denmark	.009	.010	.011	.012	123	130	140	158	5.1	5.7	7.7	12.9	19
Japan	.037	.038	.041	.043	014	015	016	020	7.7	7.1	6.7	25.0	25
S. Korea	.010	.010	.011	.012	009	009	010	011	0.0	0.0	11.1	10.0	13
China**	.001	.001	.001	.001	003	003	003	003	0.0	0.0	0.0	0.0	10
Cuba	.008	.008	.008	.008	011	012	012	012	0.0	9.1	0.0	0.0	14
Jamaica	.095	.094	.094	.093	068	076	084	091	25.9	11.8	10.5	8.3	09
Brazil	.251	.215	.283#	.264	746	771	804	843	4.2	3.4	4.3	4.9	08
Chile	.120	.112	.105	.100	736	767	794	822	5.0	4.2	3.5	3.5	10
El Salvador	.071	.071	.071	.071	149	158	170	170	6.4	6.0	7.6	7.1	08
Maryland	.043	.043	.044	.047	683	705	752	829	2.6	3.2	6.7	10.2	16^
Florida	.080	.080	.080	.081	776	809	850	912	5.9	4.3	5.1	7.3	21^

*All country/state data taken from Worldometer; ^Word Bank estimates, 2019; Census estimates, 2020.

**The actual POR rates are .00053 for all dates; rounded up to.001 for ease of presentation; #, the testing rate for Brazil in this two- week period decreases, which is peculiar and unexplained.