medRxiv preprint doi: https://doi.org/10.1101/2020.12.01.20241570; this version posted December 2, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity.
It is made available under a CC-BY-NC-ND 4.0 International license .

The Acceleration Index as a Test-Controlled Reproduction Number: Application to COVID-19 in France*

Christelle Baunez† Mickael Degoulet‡ Stéphane Luchini,§
Matteo L. Pintus¶ Patrick A. Pintus‖ Miriam Teschl**

December 1, 2020

Abstract: We show that the acceleration index, a novel indicator that measures acceleration and deceleration of viral spread (Baunez et al. 2020a,b), is essentially a test-controlled version of the reproduction number. As such it is a more accurate indicator to track the dynamics of an infectious disease outbreak in real time. We indicate a discrepancy between the acceleration index and the reproduction number, based on the infectivity and test rates and we provide a formal decomposition of this difference. When applied to French data for the ongoing COVID-19 pandemic, our decomposition shows that the reproduction number consistenly underestimates the resurgence of the pandemic since the summer of 2020, compared to the acceleration index which accounts for the time-varying volume of tests. From the comparison that we present here follows that the acceleration index is a sufficient statistic to track the pandemic's propagation, as it captures in real time the sizeable time variation featured by viral circulation.

JEL Classification Numbers: I18; H12

Keywords: COVID-19; Reproduction Number; Lock-down; Acceleration Index; Real-time Analysis; France

*This work was supported by French National Research Agency Grants ANR-17-EURE-0020.

†Senior Research Fellow, Aix-Marseille Univ, CNRS, Institut Neurosciences Timone. Email:christelle.baunez@univ-amu.fr

‡Research Fellow, Aix-Marseille Univ, CNRS, Institut Neurosciences Timone. Email:mickael.degoulet@univ-amu.fr

§Research Fellow, Aix-Marseille Univ, CNRS, Aix-Marseille School of Economics. Email:stephane.luchini@univ-amu.fr

¶AgroParisTech. Email:matteo.pintus@agroparistech.fr

‖Professor, Aix-Marseille Univ, CNRS, Aix-Marseille School of Economics. Email:patrick.pintus@univ-amu.fr

**Associate Professor, Aix-Marseille Univ, CNRS, EHESS, Aix-Marseille School of Economics. Email:miriam.teschl@ehess.fr

NOTE: This preprint reports new research that has not been certified by peer review and should not be used to guide clinical practice.

medRxiv preprint doi: https://doi.org/10.1101/2020.12.01.20241570; this version posted December 2, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity.
It is made available under a CC-BY-NC-ND 4.0 International license .

1 Introduction

The reproduction number is a now classic and widely used measure of how fast a pathogen propagates both at the outset and during an infectious disease outbreak (see for example May and Anderson [12]). But it has one major shortcoming: it does not control for the amount of tests (or any other form of diagnostics) in real time. Doing so is of crucial importance for two reasons. One is the fact that accurate empirical estimates of reproduction numbers are time-varying in nature (see e.g. Fraser [8] among many others). A considerable source of time variation comes from the fact that the amount of tests varies substantially across time and hence affects the number of known cases. Second, testing acts as a magnifying lens on viral activity at least on the part of the population that has effectively been tested. The reproduction number however does not rely on that information but rather makes assumptions on the infectivity of positive cases based on the observation of onset of symptoms and transmission in closed systems such as households (see e.g. Cori et al. [7]). But this information, again, depends on tests (or any diagnostics more generally). Hence inferring infectivity from (assumed) transmissions is only secondary information based on the availability of tests. The logical consequence of these observations is that any measure of virus reproduction should take tests (or any available diagnostics) into account and this is exactly what the acceleration index does in a transparent way.

In Baunez et al. [3, 4], we have introduced an alternative and novel measure of viral activity in the context of COVID-19 - the acceleration index. This measure considers the variation of cases *relative to* the variation of tests and thus it avoids the shortcomings mentioned above. The purpose of this article to discuss the reproduction number in the light of our acceleration index, and to show that the former is actually a special case of the latter and in fact a less accurate metric of the pandemic's time-varying speed.

We examine this important issue in two steps. In Section 2, we start from the very definition of the reproduction number as a gross rate of growth of infected people, traditionally denoted R, and derive a general formula that connects it to our acceleration index that we denote ε. In particular, we present an explicit measure of the ratio between R and ε, the interpretation of which is further discussed in terms of the infectivity and test rates. Our theoretical inquiry stresses that while the acceleration index is a ratio of growth rates - that of cases divided by that of tests - the reproduction tracks only the growth rate of cases. In Section 3, we apply such an analysis to French data and we show that there is a sizeable difference between both measures, in the sense that the reproduction number R largely *under-estimates* the spread of the virus, compared to our test-controlled measure of acceleration. It is in this sense that we say that the reproduction number is biased. This has obviously important consequences if the reproduction number is used as the basis for which public health decisions such as entering or exiting a lock-down. We also look at the effects of the second lock-down period in France, which started October 30, 2020, through the lens of both indicators, as a further example that illustrates the bias unavoidably implied by not adjusting for the volume of

medRxiv preprint doi: https://doi.org/10.1101/2020.12.01.20241570; this version posted December 2, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity.
It is made available under a CC-BY-NC-ND 4.0 International license .

tests over time when measuring the pandemic's acceleration.

2 Controlling the Reproduction Number for the Quantity of Tests: Theory

About a century ago, a series of seminal articles by Kermack and McKendrick [9, 10, 11] have laid the foundations for a mathematical theory of epidemics. More specifically, their compartmental (that is, Susceptible-Infected and Removed or SIR-type) and time-since-infection models have been extensively used and refined in the academic literature about infectious and emerging diseases. A core concept in this paradigm is the reproduction number, usually noted R, which roughly captures how many secondary cases originate, on average, from a pool of primary cases who is still currently infectious (see May and Anderson [12]).

As evident from publications by health agencies around the world since the outset of the COVID-19 outbreak, much of the guidance for designing policy measures to curb the pandemic relies prominently on estimates of R, among other things. The reproduction number is initially a theoretical concept, conceived to understand the transmissibility of an epidemic. Many efforts have been put into defining ways to empirically estimate it. Broadly speaking, estimation strategies fall into two broad categories. The first one rests on the basic SIR model and its variants, which predicts that the reproduction R is the product of four parameters: the duration of infection, the number of contacts per case and the fraction of contacts who are in turn infected, on average, and finally the fraction of total population susceptible to infection. Although each of these parameters could be estimated in real-time, this turns out to be a gigantic task, in particular when a novel pathogen like SARS-Cov-2 emerges. A short-cut to avoid such a demanding procedure is to fit a SIR model using the number of cases, so as to estimate R directly, given the infection duration (see, among many others, Althaus [1] for a recent example related to Ebola using maximum likelihood estimation). This is feasible, even in real-time, provided that enough data is available to ensure precision and structural assumptions about the time-dependency of R are made.[1] The second estimation strategy addresses more directly the time-varying dimension of R, which is more in line with epidemiological and clinical data. Many health agencies rely on such estimates of time-since-infection transmission models rather than SIR-type models. Here the basic idea is that R is essentially (1+) the growth rate of infected, which is the ratio between the number of new (that is, secondary) cases arising, say, within 24 hours, and the number of primary cases (see Fraser [8]).[2]

Our task here is to relate the acceleration index defined in Baunez et al. [3, 4] and the reproduction

[1] An additional issue arising from estimation based on compartmental models is the sizeable range of estimates. See Chris et al. [2] for SARS, and Viceconte and Petrosillo [13] for the early stages of COVID-19.

[2] For example, the French agency in charge of health statistics uses the Cori method, after Cori et al. [7]. See https://www.santepubliquefrance.fr/content/download/266456/2671953. Other European health agencies are also using this method, e.g. Austria and Germany.

medRxiv preprint doi: https://doi.org/10.1101/2020.12.01.20241570; this version posted December 2, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity.
It is made available under a CC-BY-NC-ND 4.0 International license .

number that is estimated using the time-since-infection approach just described. The main purpose of this section is to derive a theoretical relationship between both concepts, which helps both to explain why they are different, to give a sense of the magnitude of their difference, and to state the conditions under which they are equivalent. We then turn, in the next section, to data to gauge whether the difference between the two matters to track the COVID-19 pandemic.

Suppose that data is available about the number of tested and positive persons, up to end date T. Denote $\{p_1, \ldots, p_T\}$ the historical times series of the new (per period) number of positive persons from date $t = 1$ to end date $t = T$. Similarly, $\{d_1, \ldots, d_T\}$ is the historical times series of new (per period) diagnosed/tested persons. Denote $P_t = \sum_{\tau=1}^{t} p_\tau$ and $D_t = \sum_{\tau=1}^{t} d_\tau$ the cumulative numbers of positive and diagnosed persons up to date t.

As stressed in Baunez et al. [3], accurate information about the dynamics of a pandemic rests on both the number of cases and the number of tests, and the former cannot be properly understood without the latter. In that paper, we introduce an acceleration index that is an elasticity, which is a scale-free measure of responsiveness of cases to tests. This acceleration index, denoted ε_T at date T, gives the percentage change of cases divided by the percentage change of testing.

$$\varepsilon_T = \left[\frac{P_T - P_{T-1}}{P_T}\right] \div \left[\frac{D_T - D_{T-1}}{D_T}\right] \tag{1}$$

Rearranging the terms of the latter equation, we see that the acceleration index relates to the daily and average positivity rates, in the following way:

$$\underbrace{\frac{P_T - P_{T-1}}{D_T - D_{T-1}}}_{\text{daily positivity rate}} = \underbrace{\frac{P_T}{D_T}}_{\text{average positivity rate}} \times \underbrace{\varepsilon_T}_{\text{acceleration index}} \tag{2}$$

Our rather general assumption is, in accordance with the mathematical literature on epidemics, that the reproduction number is essentially a gross rate of growth and, as such, can be written at date t as:

$$R_t = \frac{p_t}{f_t[p_t, p_{t-1}, \ldots, p_{t-n}]} \tag{3}$$

where f_t is a function of new cases from date t to date $t-n$, which can be thought of as the infectious potential, that is, the average number of people who have been infected at t and before, and who can infect people at t. The parameter n is related to infection duration. Specifications for f_t have been used in the literature. For example, Fraser's [8] equation (9), on page 3 of his paper, defines the time-varying effective reproduction number as follows:

$$R_t = \frac{p_t}{\sum_{j=0}^{n} w_j p_{t-j}} \tag{4}$$

where the weights w's capture the generation time distribution, with $\sum_{j=0}^{n} w_j = 1$. This means that the time-independent function f that follows from the denominator in equation (3) is, in that case, assumed to be linear in the number of cases.[3]

[3]Such an assumption implies that, given R, the dynamics of new cases follow an autoregressive process $AR(n)$.

medRxiv preprint doi: https://doi.org/10.1101/2020.12.01.20241570; this version posted December 2, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity.
It is made available under a CC-BY-NC-ND 4.0 International license .

Even though it might go unnoticed a first sight, we should stress that a major difference between the rather general definition of R in equation (3) and the usual definition in equation (4) is that the function f_t implicitly depends on calendar time t to reflect the fact that positives are detected by biological tests or, more generally, other technologies to diagnose infected cases.[4] This difference turns out to be important to understand the connection between the acceleration index and the reproduction number, as we now show.

Using equations (2) and (3), we can relate our acceleration index and the reproduction number in the following way:

$$\varepsilon_T = R_T \times \frac{A_T}{B_T} \quad \text{with} \quad A_T = \frac{f_T[p_T, p_{T-1}, \ldots, p_{T-n}]}{P_T} \quad \text{and} \quad B_T = \frac{d_T}{D_T} \tag{5}$$

Equation (5) shows that the difference between the acceleration index and the basic reproduction number can be thought of itself as a ratio. The numerator of this latter ratio, A, is simply the infectivity rate, that is, the ratio of the number of primary cases up to period T who can originate infections in T as a fraction of the total number of persons who have been infected since the outset of the pandemic. The denominator, B, on the other hand, represents the fraction of tests in period T out of cumulated tests, that is, the current test rate. To sum up, the ratio of the acceleration index to the reproduction number is, in any period, the ratio of the infectivity rate to the test rate.

From Equation (5) we see that both indicators are equivalent, that is, $\varepsilon = R$, if and only if at all dates t:

$$f_t[p_t, p_{t-1}, \ldots, p_{t-n}] = d_t \times \frac{P_t}{D_t} \tag{6}$$

Equation (6) is very important to conceptualize the core idea of this paper: in order to properly control for the (time-varying) volume of tests/diagnostics, one needs to use the appropriate function f_t, that is, one which depends on calendar time because tests do. Said differently, the function f_t should be specified in such a way that it takes account of the fact that cases are produced by tests or any other diagnostics. The linear form with no time dependence which appears in the denominator of equation (4) is therefore problematic, as it assumes away tests which are however key to measure the pandemic's dynamics. In this sense, the acceleration index ε nests the basic reproduction number R: if the function f_t is specified as in equation (6), R is equivalent to ε as it takes account of testing; in any other case, ε is more general than R as in equation (4) that takes account of cases only.

To better understand the relation between R and ε as indicated in equation (5), a more theoretical analysis and its implications may be helpful. First of all, as it also becomes clear from equation (5), when $A = B$ - that is when equation (6) holds - then $\varepsilon = R$. This basically means that if the test rate B follows the dynamics of the infectivity rate A, there is enough testing to capture viral

[4] The fact that the diagnostics dimension is largely ignored in the literature about SIR-type models surfaces, for instance, in Wallinga and Lipsitch [14], who relate the epidemic growth rate to incidence and generation time interval only. But it seems reasonable to assume that infectious and emerging diseases involve a diversity of pathogens, and have a variety of technologies to diagnose. In the context of COVID-19, PCR tests are of course key.

medRxiv preprint doi: https://doi.org/10.1101/2020.12.01.20241570; this version posted December 2, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity.
It is made available under a CC-BY-NC-ND 4.0 International license .

activity. In fact, seen from this perspective, we have a clear testing strategy: the daily tests d_t need to offset the assumed infectiousness captured by function f_t, and more specifically equation (6). The smaller the total number of cases, the easier it will be to match that testing requirement in particular through contact tracing. As total cases go up, contact tracing and sufficient testing may come to its structural and systemic limits. This in itself is a sign that additional health policies will need to be promoted.

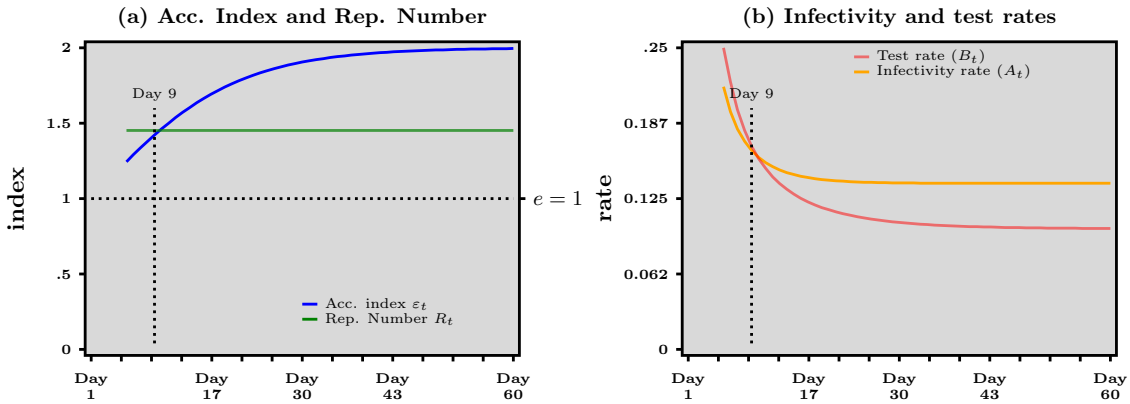
If $A > B$, then $\varepsilon > R$, whilst when $A < B$, $\varepsilon < R$. In the former case, the infectivity rate A of the pandemic cannot be captured by the testing rate B in place. That is, R does not give the appropriate picture of the infectiousness of the pandemic, in fact it underestimates it. To alleviate this bias, either testing would need to be increased, or viral spread would need to be cut by establishing policies that reduce contacts or a mixture of both. In any case, it shows that equation (3) that composes R does depend on more than past and current cases, because they themselves depend on tests and other factors that may favour or not transmissibility. Equivalently in the latter case, R will overestimate the speed of the pandemic if the test rate B is greater than the infectivity rate A. In such a situation, greater testing than underlying infectivity will necessarily find more cases, actually too many to reflect the correct transmissibility. To capture the correct picture, either testing would need to be reduced, which however seems counterproductive at least to the extent that testing is a way to look at the underlying viral dynamics, or the infectivity function f of equation (3) needs to be adapted to reflect reduced transmissibility.

A further analysis of equation (5) requires structural assumptions, in particular to generate predictions about how both ε and R, rather than their ratio, move over time. Just for the sake of illustration, we now resort to the handy example of deterministic exponential growth. Time is assumed to be continuous, to ease derivation of results, and the number of cases grows exponentially over time, as usually assumed in epidemiological models, of SIR type and related for example. In such a case, R as defined in equation (4) is constant and any difference between R and ε is due to differences between A and B. For ε, we also have to introduce tests and we assume that they also grow exponentially.

Under those assumptions, we show in Appendix A that while the reproduction number is constant over time, the acceleration index is not, as it features different regimes depending on how the growth rate of daily cases compares with the growth rate of daily tests. For example, when the former is larger than the latter, the acceleration index first rises and then approaches a plateau, where it equals the ratio of growth rates, which is larger than 1 in that case. In contrast, the reproduction number stays constant over time. We can visualize this more easily in the simple setting of exponential growth (see Appendix A), but it also holds more generally that the difference between both indicators is essentially due to the fact that *while the acceleration index is the ratio of two growth rates, that of cases divided by that of tests, the reproduction number tracks only the former, thus ignoring the latter*. It is also for this reason that we say the the acceleration index ε nests the basis reproduction number R, which is simply a special case of ε. Different configurations

medRxiv preprint doi: https://doi.org/10.1101/2020.12.01.20241570; this version posted December 2, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity.
It is made available under a CC-BY-NC-ND 4.0 International license .

Figure 1: Numerical example for exponential growth of both daily cases and daily tests. Panel (a) Acceleration index (blue curve) and Reproduction number (green curve). Panel (b) Infectivity rate (orange curve) and Test rate (red curve)



may in principle occur, therefore, over time, depending on how fast cases grow compared to the growth of tests.

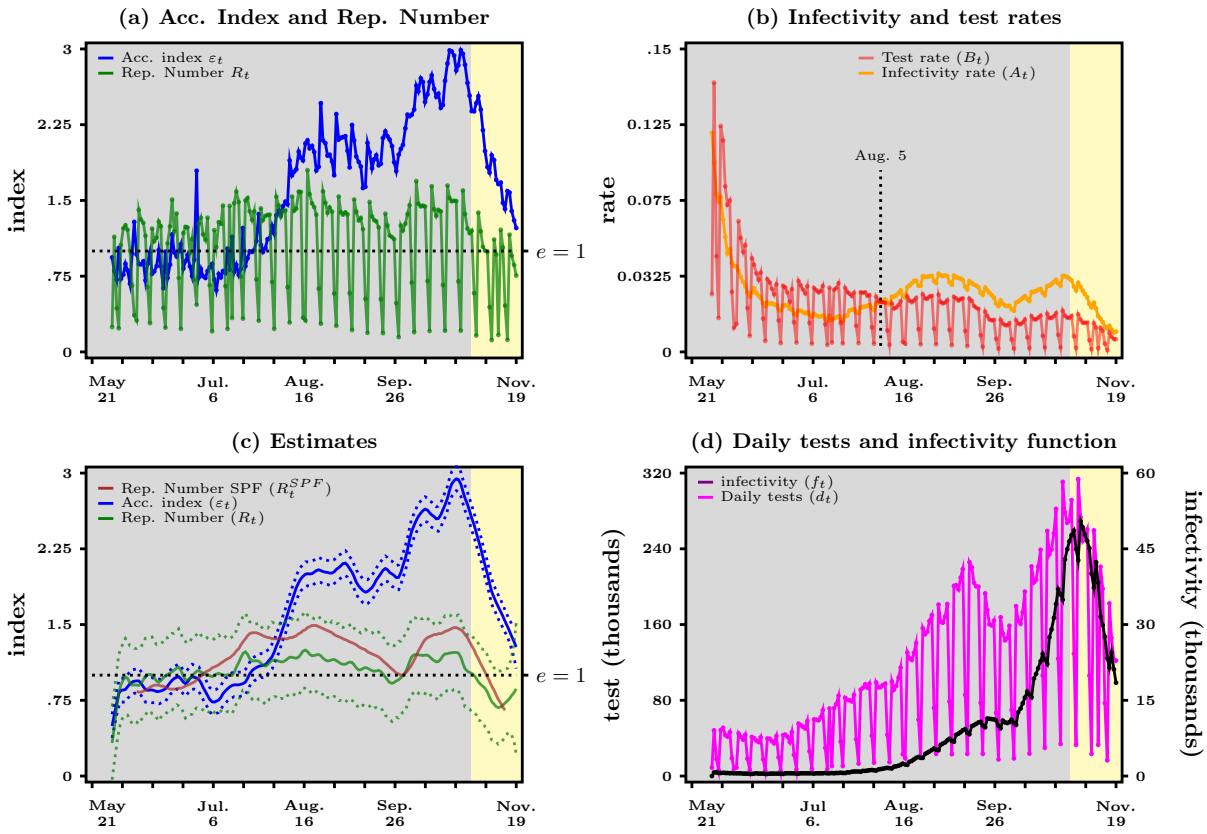
To further illustrate what happens in the case of exponential growth outlined above and studied in more details in Appendix A, we now provide an illustration such that the growth rate of daily cases is twice as large as the growth rate of daily tests. Figure 1 illustrates how the acceleration index and the reproduction number, as well as the infectivity and test rates, evolve over time in this particular example.

In Figure 1, panel (a), we report the evolution over time of the time-varying acceleration index $\varepsilon(t)$ and the constant reproduction number R that follow from the numerical example. In Appendix A, we show that while R is constant, ε tends to the ratio of growth rates, which is equal to 2 in the example. As a consequence, a first regime with the reproduction number exceeding the acceleration number happens, followed by a second regime that features the reverse configuration. Not surprisingly, panel (b) in Figure 1 shows that the first regime materializes when $B(t) > A(t)$ - that is the test rate exceeds the infectivity rate - while the second regime is associated with $A(t) > B(t)$. Panel (b) reveals in particular that the plateau for the acceleration index that is featured in panel (a) comes from the fact that both $A(t)$ and $B(t)$ themselves plateau, with the infectivity rate exceeding the test rate. Overall, therefore, the acceleration tracks the ratio of grow rates - which equals 2 in our example - while the reproduction number underestimates that ratio because it roughly reflects only its numerator.

In the next section, we illustrate the theoretical decomposition outlined above to capture how the reproduction number and acceleration index differ in the context of the current COVID-19 epidemic in France.

medRxiv preprint doi: https://doi.org/10.1101/2020.12.01.20241570; this version posted December 2, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity.
It is made available under a CC-BY-NC-ND 4.0 International license .

Figure 2: Panel (a) Acceleration index (blue curve) vs reproduction number (green curve). Panel (b) Infectivity rate (orange curve) vs test rate (red curve). Panel (c) Kernel estimates with confidence bands (dashed lines). Panel (d) Daily tests (purple line) and infectivity function (black curve). Beige area depicts the lock-down period. Source: Agence Santé Publique France and authors' computations.



3 Test-Controlling the Reproduction Number in Practice: An Illustration using French Data

In Figure 2, panel (a), we report both the acceleration index (blue curve) and the reproduction number (green curve) over time, using data for France. The acceleration index ε is computed using equation (1) while the reproduction number R is computed from equation (4) with $n = 7$ and equal weights w.[5] The lower spikes of the reproduction function R are due to lower amount of testing during week-ends. This can clearly be seen in the panel (d) of Figure 2 that presents the number of daily tests (in pink). We represent here the raw data rather any smoothed estimates in order to avoid any layer of interpretation.

[5]The infection kernel could be adapted to account for sub-exponential growth as in Chowell et al. [6].

medRxiv preprint doi: https://doi.org/10.1101/2020.12.01.20241570; this version posted December 2, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity.
It is made available under a CC-BY-NC-ND 4.0 International license .

In panel (c) of Figure 2, we present local polynomial regressions of R and ε of panel (a), that use the Savitzky-Golay filter also known as a locally estimated scatter-plot smoothing method in modern statistics (see Cleveland and Devlin [5]). The blue line is again our acceleration indicator. In red, we show the reproduction number as used by Santé Publique France, whereas the green line represents our own estimation. As can be noted, the reproduction number estimated by Santé Publique France falls within the confidence bands of our own estimate which are the dotted green lines.[6] In addition, as can be seen from the confidence bands, the acceleration index is estimated more precisely than the reproduction number, because we take account of variations of tests and thus cases due to the week-end effect.[7]

If we concentrate on panel (c) of Figure 2, we see that right after the end of the first lock-down, both indicators are roughly "plateauing" just under 1, with R being slightly above ε. We concentrate here on the green line, i.e. our estimation of R. This estimated R then rises quickly to a higher plateau in the first half of July to indicate greater transmissibility, and stays at a level of about 1.2 until mid-August. At that same time ε first remains put at a level smaller than 1 and becomes greater than 1 a few days later, effectively crossing R at the beginning of August and accelerating all along until about mid-August.

The difference in dynamics of both indicators can easily be explained by looking at panel (b) in Figure 2. Here, we report the two terms that appear in equation (5), that is, A, the infectivity rate (orange curve), and B, the test rate (red curve). The latter graph exhibits negative autocorrelation, due to the fact that much less tests, if any, are performed during week-ends. The test rate follows a downward trend that simply reveals the fact that tests being done in a given period constitute, over time, a smaller and smaller fraction of the cumulated amount of diagnostics. What we see in panel (b) in particular is that before August 5, the test rate B is greater than the infectivity rate A that has, at first, also a downward trend. A greater testing rate implies that more cases will be found. This corresponds to the period when R is greater than ε. But R basically overestimates viral activity because it does not consider tests and focuses on cases only, while ε takes account of this because it looks at the ratio of both infectivity and test rates. The opposite is true for the period after August 5, when the infectivity rate A becomes greater than the testing rate B and ε greater than R. Despite growing daily tests until the end of August, as we can see in panel (d) of Figure 2, R remains first at a plateau but then steadily declines until the second half of September.

This clearly shows that R has been unable to represent the viral dynamics basically during the both summer months because, as we can see from panel (b), the infectivity rate is rising more quickly than the test rate. Therefore, R overlooks the testing dimension and only captures the number of cases, but those are undervalued given a lower test rate B. Even worse, testing then declines at the end of August while the infectivity function f, indicated in black in panel (d) starts going up.

[6] Even though Santé Publique France refers to the Cori method, we have not found public information about the weights attached to past values for the number of cases in computing infectivity.

[7] In effect, being a ratio of growth rates, the acceleration index is smoother than the growth rate of cases.

medRxiv preprint doi: https://doi.org/10.1101/2020.12.01.20241570; this version posted December 2, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity.
It is made available under a CC-BY-NC-ND 4.0 International license .

This affects R that declines to reach a level of about 1 by the end of September. This is in very stark contrast to ε that accelerates from early July onwards and then hovers at a plateau of about 2 up to end of September. It takes appropriately into account the relationship between the changing growth rates of testing and infectivity.

Both indicators go up again from the end of September onwards as testing rises again. But while R reaches a plateau again as the first curfew measures where put into place to cut transmission, ε further indicates acceleration. Both indicators then start declining when, at the end of October, the second lock-down was put into place.

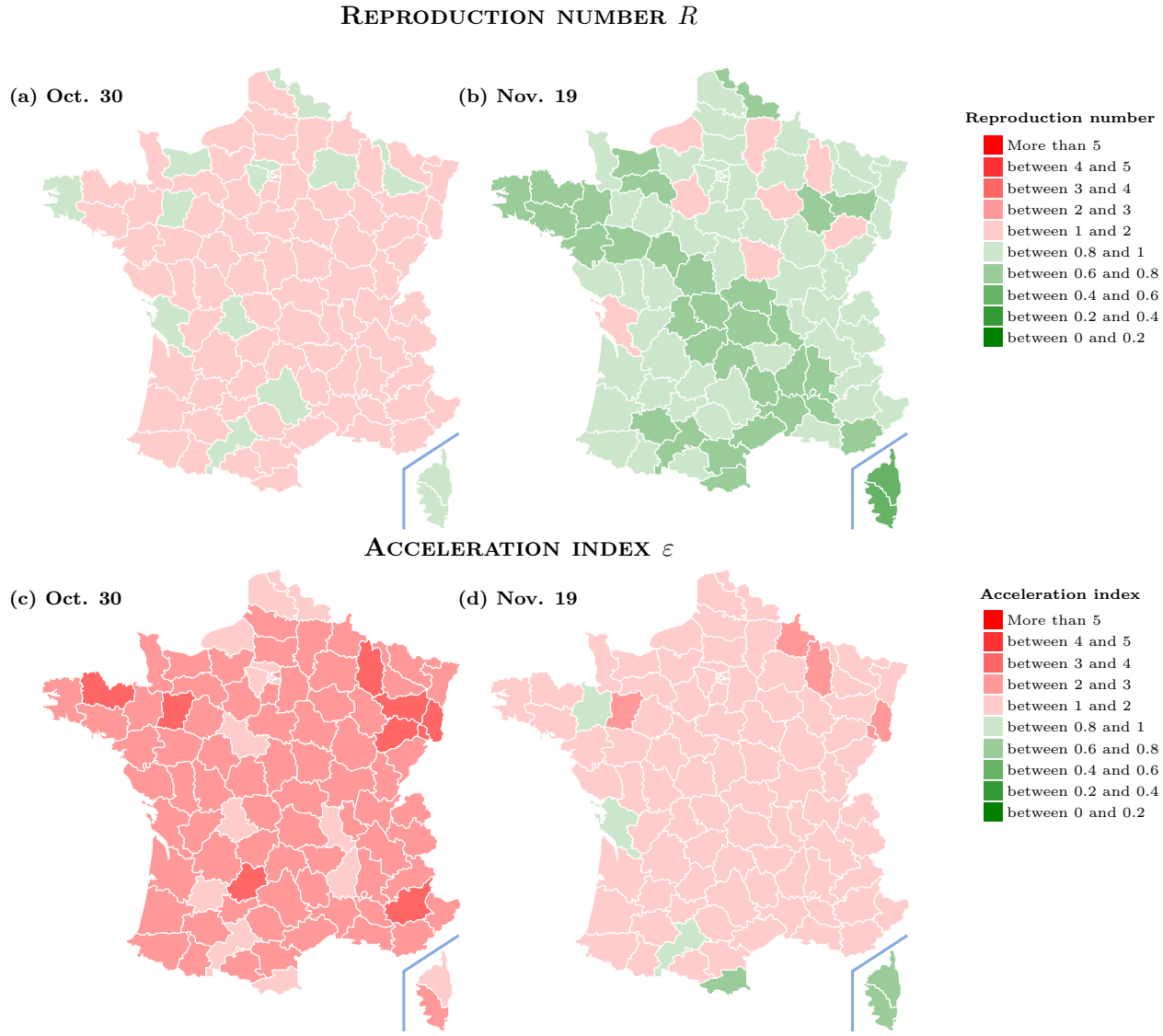
However, $R < 1$ since the beginning of November, whilst at the same time, our indicator ε still shows an ongoing acceleration, although a reduced one with respect to the time before the lock-down. What we see very clearly from panel (d) is that lock-down coincides with a great reduction of testing. Obviously, lock-down is aimed at reducing contacts and thus viral spread. This will necessarily reduce cases and hence R declines. But if at the same time testing is reduced as well, which is the only way to get a clearer picture of the viral activity this necessarily influences R more dramatically and explains the under-evaluation of the viral spread than ε that continues to indicate acceleration. More specifically, ε indicates what happens to cases when we reduce testing by some percentage change. The fact that the percentage change of cases goes in the same direction as the percentage change of testing, i.e. that both decrease, is a good sign and indicates that lock-down measures have their effect. But looking at ε does not yet allow to give an all-clear such as R does. As a consequence, ε captures more accurately the considerable time variation of virus propagation than R.

To sum up, two main differences between the reproduction number and the index appear in panel (c). The first being that R crosses unity earlier than the acceleration index, which starts to increase around July 6. This is most likely due to the infectivity rate reaching its lowest point around that date (as seen in panel (b)). Passing that date, the infectivity rate begins to increase, while the test rate continues downwards. This explains why the acceleration index could not start growing before July 6. Overall, therefore, R is larger than ε before August 5. The second difference is a sudden decrease of the reproduction number in the second half of September, while the acceleration index stays at a plateau. This can be explained by panel (d), in which we see a sharp plummet in the number of daily tests around that period. Less tests equals less detected cases, which R relies heavily on for its calculation. Seeing R rising sharply from about 1 in early October is all the more surprising when seen in isolation. In contrast, the acceleration index, which accounts for variations of both cases and tests, consistently shows a succession of periods of steep rise followed by plateaus over the summer and until the second lock-down.

To further illustrate the differences between R and ε and their consequences for public health decisions, we consider Figure 3. This Figure 3 gives an overview of how the second lock-down, which started on October 30, 2020, has contributed to reduce virus circulation across French départements.

medRxiv preprint doi: https://doi.org/10.1101/2020.12.01.20241570; this version posted December 2, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity.
It is made available under a CC-BY-NC-ND 4.0 International license .

Figure 3: Acceleration index and reproduction number for French départements at different dates. October 30 dates the start of the second lock-down in France. Data source: Agence Santé Publique France and authors' computations.



Looking at the two bottom maps shows that the acceleration index has been reduced everywhere during the period from October 30 to November 19. While a similar improvement is indicated by the reproduction number, as one concludes from the top maps, that measure of virus spread tells an altogether different story. Three weeks after the beginning of the second lock-down in France, the acceleration index suggests that the acceleration regime still prevails, except for 6 départements which happen to be the happy few, but with values still close to unity. However, one concludes rather wrongly from the reproduction number that, at the same date, deceleration is underway in most départements in green. In view of the discussion about the ε/R ratio in Section 3, this comes

medRxiv preprint doi: https://doi.org/10.1101/2020.12.01.20241570; this version posted December 2, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity.
It is made available under a CC-BY-NC-ND 4.0 International license .

as no surprise since the reproduction number under-estimates virus circulation, due to the fact that it does not take tests into account.

In practice, many public health agencies report (daily or weekly) positivity rates, to complement the information contained in the reproduction rate R. In light of the connection with ε that we have highlighted in this note, both formally and empirically, we argue that the acceleration index is closer to a sufficient statistic that helps tracking the rapidly changing dynamics of any pandemic, because it explicitly takes into account the dynamics of diagnostics. In the context of COVID-19, diagnostics equal tests, but our claim is valid more generally when this is not the case. This means that the index can potentially be applied to any effort designed at detecting infected people, no matter what the pathogen agent triggering the infectious disease turns out to be. In real time, this is quite valuable, we believe, to guide health policies and to assess containment measures, especially in the context of a new pathogen appearing (such as SARS-Cov-2), with unforeseeable pandemic dynamics.

4 Conclusion

We show in this paper that the reproduction number R is a special case of the the acceleration index proposed in Baunez et al. [3]. While the former only looks at the growth rate of cases, the latter considers variations of cases in relation to tests. The acceleration index is a sufficient statistic of viral spread, in the sense that it aggregates information that is usually looked at separately, like positivity or prevalence rates. As such, a test-controlled reproduction number like the acceleration index should be part of any data dashboard to track an epidemic, and especially to guide public policy in the design of the most efficient methods to curb it. For example, we have shown in Baunez et al [3] that an accurate measure of virus circulation is a key input to feed algorithms that are designed to efficiently allocate the diagnostic effort across space.

This makes the acceleration index a much more useful indicator to track a pandemic in real-time, as it is context-dependent: the acceleration index takes into account the effort to diagnose people who have been infected by the pathogen. In the case of COVID-19, diagnostics equal PCR (and other types of biological) tests, but this might not be the case for other diseases where diagnostics require even greater effort. However, our analysis makes a strong case for incorporating in any measure of pathogen circulation the observed effort to diagnose the agent that makes people sick.

Even though there is admittedly a variety of infectious (and emerging) diseases, with different pathogens and various ways to diagnose them, we claim that our conceptual approach is general enough to shed light, not only on the current pandemic, but also on any future ones which may come.

medRxiv preprint doi: https://doi.org/10.1101/2020.12.01.20241570; this version posted December 2, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity.
It is made available under a CC-BY-NC-ND 4.0 International license .

References

[1] Althaus C.L. (2014): Estimating the Reproduction Number of Ebola Virus (EVOB) during the 2014 Outbreak in West Africa. *PLOS Currents*, Sep 2:6. 3

[2] Bauch C.T., Lloyd-Smith J.O., Coffee M.P., Galvani A.P. (2005): Dynamically Modeling SARS and Other Newly Emerging Respiratory Illnesses: Past, Present, and Future. *Epidemiology*, 16:791-801. 3

[3] Baunez C., Degoulet M., Luchini S., Pintus P., Teschl M. (2020): Tracking the Dynamics and Allocating Tests for COVID-19 in Real-Time: an Acceleration Index with an Application to French Age Groups and Départements. MedRχiv preprint 11.05.20226597 available at https://doi.org/10.1101/2020.11.05.20226597. 2, 3, 4, 12

[4] Baunez C., Degoulet M., Luchini S., Pintus P., Teschl M. (2020): An Early Assessment of Curfew and Second COVID-19 Lock-down on Virus Propagationin France. MedRχiv preprint 11.11.20230243 available at https://doi.org/10.1101/2020.11.11.20230243. 2, 3

[5] Cleveland, W.S., Devlin, S.J. (1988): Locally-Weighted Regression: An Approach to Regression Analysis by Local Fitting. *Journal of the American Statistical Association*. 83:596-610. 9

[6] Chowell G., Viboud C., Simonsen L., Moghadas S.M. (2016): Characterizing the Reproduction Number of Epidemics with Early Subexponential Growth Dynamics. *Journal of the Royal Society Interface*, 13(123):20160659. 8

[7] Cori A., Ferguson N.M., Fraser C., Cauchemez S. (2013): A New Framework and Software to Estimate Time-Varying Reproduction Numbers during Epidemics. *American Journal of Epidemiology*, 178:1505-1512. 2, 3

[8] Fraser C. (2007). Estimating Individual and Household Reproduction Numbers in an Emerging Epidemic. *PLoS One*, 2:e758. 2, 3, 4

[9] Kermack W.O., McKendrick A. G. (1927): A Contribution to the Mathematical Theory of Epidemics. *Proceedings of the Royal Society of London. Series A*, 115:700-721. 3

[10] Kermack W.O., McKendrick A. G. (1932): A Contribution to the Mathematical Theory of Epidemics II. The Problem of Endemicity. *Proceedings of the Royal Society of London. Series A*, 138:55-83. 3

[11] Kermack W.O., McKendrick A. G. (1933): A Contribution to the Mathematical Theory of Epidemics. III. Further Studies of the Problem of Endemicity. *Proceedings of the Royal Society of London Series A*, 141: 94-122. 3

[12] May R.M., Anderson R.M. (1991). Infectious Diseases of Humans: Dynamics and Control. Oxford University Press. ISBN $0-19-854040-X$. 2, 3

medRxiv preprint doi: https://doi.org/10.1101/2020.12.01.20241570; this version posted December 2, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity.
It is made available under a CC-BY-NC-ND 4.0 International license .

[13] Viceconte G., Petrosillo N. (2020): COVID-19 $R0$: Magic Number or Conundrum? *Infectious Disease Reports*, 12:8516. 3

[14] Wallinga J., Lipsitch M. (2007): How Generation Intervals Shape the Relationship between Growth Rates and Reproductive Numbers. *Proceedings of the Royal Society of London Series B*, 274:599-604. 5

A The Case of Exponential Growth

We now illustrate the relationship between the acceleration index and the reproduction number when time is assumed, to ease derivation of results, to be continuous and when the number of cases grows exponentially over time, as usually assumed in epidemiological models, of SIR type and related for example. Although typically absent in the latter strand of literature, we have to introduce tests and we assume that they also grow exponentially. More formally, using the notation in the previous section, suppose that the number of cases per unit of time is denoted by $p(t) = \alpha e^{\beta t}$ while the number of tests per unit of time is $d(t) = \gamma e^{\nu t}$, where the growth rates β and ν are assumed to be positive for the sake of illustration. The reproduction number is then constant over time, as we now show when the infection kernel is uniform, that is, if the weights w are constant and equal to $1/\delta$ to ensure that $w \int_0^\delta ds = 1$.[8] The analog of equation (4) is as follows:

$$R(t) = \frac{p(t)}{w \int_0^\delta p(t-s)ds} \tag{7}$$

which implies that, given $p(t) = \alpha e^{\beta t}$, one has:

$$R(t) = R = \frac{\beta\delta}{1 - e^{-\beta\delta}} \tag{8}$$

It is easily seen that the reproduction is the growth rate of daily cases (adjusted for the delay δ) only, since it obviously does not take into account tests. In fact, R is the growth rate of the denominator in equation (7), that is, what we note f in the main text.

Cumulated cases and tests are then noted $P(t) = \int_0^t p(\tau)\, d\tau$ and $D(t) = \int_0^t d(\tau)\, d\tau$, respectively. It is easy to derive, by straight integration, the expressions:

$$P(t) = \frac{\alpha}{\beta}\left(e^{\beta t} - 1\right), \; D(t) = \frac{\gamma}{\nu}\left(e^{\nu t} - 1\right) \tag{9}$$

It follows that our acceleration index is given, as function of time, by:

$$\varepsilon(t) = \frac{p(t)/P(t)}{d(t)/D(t)} = \frac{\beta}{\nu}\left(\frac{1 - e^{-\nu t}}{1 - e^{-\beta t}}\right) \tag{10}$$

From equation (10), one concludes that the acceleration index is essentially the *ratio of the growth rate of cumulated cases divided by that of cumulated tests*, since $p(t) = dP(t)/dt$ and $d(t) = dD(t)/dt$.

[8] It is not difficult to show that the constancy of R holds for other infection distributions as well.

medRxiv preprint doi: https://doi.org/10.1101/2020.12.01.20241570; this version posted December 2, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity.
It is made available under a CC-BY-NC-ND 4.0 International license .

In addition, ε tends to the ratio of the growth rate of daily cases to that of daily cases β/ν. This means that the acceleration tracks that ratio over time and converges to it eventually.

It follows that three cases occur. When $\beta = \nu$, that is, when both daily cases and daily tests grow at the exact same rate, then our acceleration index equals 1 at all dates. When the two growth rates differ, however, $\varepsilon(t)$ converges, when t goes to infinity, to the ratio of growth rates β/ν, independently of the scale parameters α and γ. As an illustrative example, suppose that $\beta > \nu$, so that positives grow faster than tests. Then the pattern of our acceleration index $\varepsilon(t)$ over time will have two regimes: it first grows almost linearly and eventually reaches the upper bound $\beta/\nu > 1$. Obviously, in that case both the daily positivity rate $p(t)/d(t)$ and the average positivity $P(t)/D(t)$ grow over time, and the latter quantity exceeds the former all the time so that acceleration prevails. The symmetric case when $\beta < \nu$ is easily adapted.

The property that the reproduction number is constant under our assumptions essentially means that the dynamics of the acceleration index is driven by that of $A(t)/B(t)$, which we now decompose to help understand how R and ε compare over time. It follows from the definition of A and B in equation (5) that:

$$A(t) = \frac{e^{\beta t} - e^{\beta(t-\delta)}}{\delta(e^{\beta t} - 1)} \quad \text{and} \quad B(t) = \nu \frac{e^{\nu t}}{e^{\nu t} - 1} \tag{11}$$

with both functions $A(t)$ and $B(t)$ decreasing with time t.

To further illustrate what happens in the theoretical case outlined above with $\beta > \nu$, let us take a numerical example. Suppose that the growth rate of cases β equals 20% while the growth rate of tests $\nu = 10\%$. In addition suppose that the delay parameter $\delta = 4$ so that the weight $w = 1/4$. Figure 1 illustrates how the acceleration index and the reproduction number, as well as the infectivity and test rates, evolve over time in this particular example.