

## Research Article

# Infection Waves in Pandemics and Risk Prediction: Physical Diffusion Theory and Data Comparisons

Romney B. Duffey\*

Idaho Falls, ID 83404, USA

**ARTICLE INFO***Article History*

Received 15 November 2020

Accepted 08 June 2021

*Keywords*COVID-19  
infection waves  
community spreading  
diffusion theory  
data  
predictions  
policies**ABSTRACT**

We predict the magnitude and estimate the uncertainties of the spread, growth and maximum expected long-term infection rates that affect emergency policies and plans. For the COVID-19 and 1918 viral pandemics, large second or successive peaks, waves or plateaux of increased infections occur long after the initial rapid onset. The key question is what physical model can explain and predict their occurrence trends and timing? We establish the principal that the timing and magnitude of such increases can be based on the well-known physics of classical diffusion theory, so is fundamentally different from the commonly used multi-parameter epidemiological methods. This physical model illuminates our understanding of the societal viral progress, providing quantitative predictions, estimates and uncertainties supporting risk decision-making and resilient medical planning. We obtain an approximate relation for predicting the risk of the observed magnitudes, timing and uncertainties of second and more waves, as needed for proactive emergency pandemic planning, bed count and decision-making purposes. The dynamic results and characteristics are compared and fitted to data using just two physical parameters for a number of countries and regions, and the concept shown to apply for both entire national and local regional populations. The present analysis *quantitatively* shows how the timing and magnitude are reduced by more learning and effective countermeasures. The medical system and health policy must recognize and pro-actively plan for such inexorable diffusive spread and large residual infection waves.

© 2021 The Author. Published by Atlantis Press B.V.

This is an open access article distributed under the CC BY-NC 4.0 license (<http://creativecommons.org/licenses/by-nc/4.0/>).

## 1. INTRODUCTION: TYPICAL INFECTION PATTERNS AND REPEATING “WAVE” PHASES

Health crisis policy and decisions are based on infection rates so need quantitative predictions of risk trends, while medical system pandemic planning and response often depends on projected emergency bed count capacity which lag infection rates. Since viral pandemics exhibit successive waves or peaks of infections, the key question is what physical model can explain and predict the risk of their occurrence, size trends and timing? The present note expands the numerical study idea by Aciola [1] by examining if the dynamic community-wide spread can be described and understood using classical diffusion theory.

To avoid confusion, the relevant fundamental physics is completely distinct from the use of the term “diffusion” by sociologists and political scientists to qualitatively explain the varying implementation of pandemic countermeasures and policies between nations [2].

The onset and subsequent progress of pandemics such as H1N1 in 1918 and COVID-19 (aka SARS-CoV-2) in 2020, are known to be characterized by multiple increases or “waves” of infections, peaking, declining and returning over many 100’s of days. We

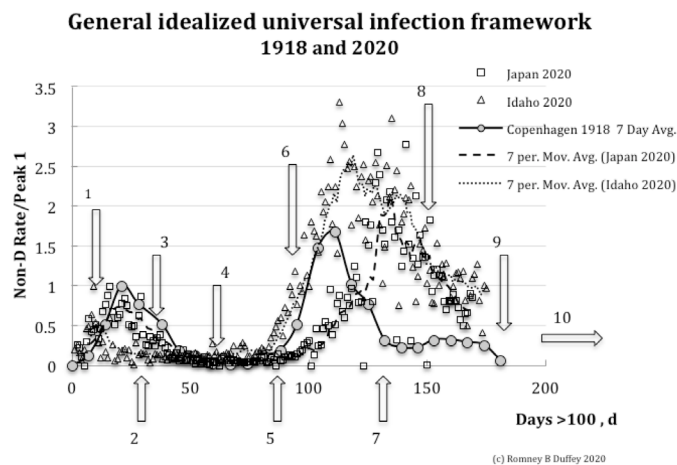
do not re-iterate all the well-documented features of pandemics and their behavior [see, e.g. 3–6], or the many and various person-to-person infection pathways and mechanisms [see, e.g. 7,8]. We distinguish between: the dominant mechanisms of initial (direct and local transmission) rapid spreading for the first wave; and subsequent slower community spreading and transmission for producing the second and more waves (within region/country). We postulate the first peak or “wave” is the result of rapid but mainly *externally introduced* infections as the virus opportunistically attacks the first susceptible and unaware hosts, and increases limited according to initially rapid (few days) transmission and incubation timescales. Infections grow exponentially by unconstrained random person-to-person transmission in an initial sample population or local neighborhood without prior resistance or effective countermeasures. The initial peak (which we term here Peak 1) is reached in about 30 days until social and Non-pharmaceutical Intervention (NPI) countermeasures (e.g. improved hygiene, social distancing, improvised quarantines, public awareness, etc...) effectively counterbalance the growth in the infection rate. The Peak 1 wave subsequently declines universally according to learning theory from adopting such simple countermeasures [9,10].

However, zero risk or rates are not achieved or achievable (as known for perennial influenza) due to hidden residual infections, asymptomatic cases, and/or undetected importations or “clusters”, which ensure the virus continues to be slowly, inexorably and

\*Email: [duffeyrb@gmail.com](mailto:duffeyrb@gmail.com)

randomly distributed. The virus progressively becomes distributed throughout the community, and largely undetected or hidden progressively continues to spread. This initially low level of mainly *internally introduced* or enabled community-wide spread *inside* countries is widely observed, even in island states where complete entry and viral import restrictions, mandates or controls are possible (like Hawaii, Australia and New Zealand). Since the initial infective decline is competing against diffusive increase, the second wave onset only starts to be discernable or noticeable many weeks after the Peak 1. Eventually, after about 100 days, a second distinct “wave” or increase toward a plateau (which we term here Peak 2) is more clearly observed, which is also reached before some balance is achieved between increased diffusive transmission and effective countermeasures.

While not exactly followed by all regions, these overall data patterns, trends and surges are compelling and apparently independent of continent, culture, society or century, as shown in Figure 1 (the rates per day are normalized to the first peak and the temporal evolution measured in days exceeding 100 initial infections as a nominal start or pandemic threshold). We hypothesize that the second and subsequent waves have a fundamentally different physical spreading character, having moved from the early rapid infection phase of initial local populations dominated by rapid personal incubation timescales (3–5 days) and contacts to slower and more extensive overall community spreading (50–100 days) dominated by slower societal interactions, more cautious behaviors



**Figure 1** | General idealized trends of pandemic infection waves over the last century, with daily infection numbers normalized to the first peak. (1) Initial random infections,  $n_0$ , from external introduction with rapid rise (incubation limited timescale,  $G$ ). (2) Exponential rise to first wave peak,  $n_{M1}$ , in circa 20–30 days,  $d_{M1}$ . (3) Decline due to learning (countermeasure like awareness, hygiene, and social distancing...),  $k$ . (4) Minimum achievable infection rate (or detection threshold),  $n_m$ . (5) Second “wave” apparent onset,  $n_{02}$ , at circa 50–100 days (community spread limited timescale). (6) Second rise,  $n(d)$ , from random inter-community spread (societally embedded). (7) Second peak or plateau,  $n_{M2}$ , at about,  $d_{M2}$ , circa 100+ days often larger than first peak. (8) Plateau or decline due to learning, countermeasures plus human host spreading limits. (9) Minimum “acceptable” or achievable rate achieved again. (10) Additional waves possible (annual, seasonal, social...).

and residual infections.<sup>1</sup> These phases may of course overlap or not be entirely distinct everywhere, so we look to data for guidance, quantification and verification.

The usual Susceptibility-Exposed-Infected-Recovery (SEIR) and  $R_0$  models for the infection rate [11–13] have been fitted using some seven parameters to first peak rise and decline data [e.g. 14,15] where the infection increase rate is limited by personal incubation timescales. Previous work on second waves [16] solved the standard SEIR model plus a further two equations, giving six first-order differential equations with eight variable parameters, four representing various assumed socio-economic countermeasures (e.g. “lockdown”, school closure, perception and learning). Despite the SEIR approach being known to be inaccurate and not predictive [13], three parameters were first fitted or “calibrated” to data, while varying the five remaining of the eight parameters gave differing Peak 2 to Peak 1 ratios. The results showed oscillations, with Peak 2 at 240 days and then a decline, and that “In some cases, the second peak was higher than the first peak, while for other parameter combinations it was lower.” It was concluded “a second wave of COVID-19 on account of the coupled behavior-disease feedbacks... will characterize many populations” [16]. There were no comparisons given to actual data as in our new approach.

To model the societal medico-physics, and avoid so many empirical parameters, we postulate that after the initial peak this subsequent community-wide spreading is dominated and governed by random diffusion, with infections seeping steadily, systematically and *inevitably* throughout the population, and largely independent of countermeasures. As first pointed out by Aciola [1] for the first peak using purely numerical solutions: “a simple diffusion model treats each individual in a population as a Brownian particle ... added to this model is the incubation period of the virus and a probability of transmission of the virus if individuals are closer than a certain distance.” This fundamental idea reflects the reality that individual infection transfers and opportunities are statistically random; and of not being able to identify, track, eliminate or monitor all the encounters and social interactions of every infected person and prevent such transfers. For a whole society, we are not looking at individuals but entire populations and overall societal behaviors and the resulting infection trends.

Therefore, for second or other successive waves the problem and equation to be solved is essentially Fick’s or Fourier’s Law, a second order differential equation, where the net rate of change of infections at any location is proportional to the incremental infection gradient. For any infection number,  $n$ , at any time,  $t$ , conventionally, without additional internal or imported infection sources:<sup>2</sup>

$$\frac{\partial n}{\partial t} = D \nabla^2 n \tag{1}$$

The community-wide diffusion coefficient or diffusivity,  $D$ , physically represents all random person-to-person and intra-societal cross infections and is the key parameter to be deduced

<sup>1</sup>We utilize infection rates as a leading indicator of spread, while public health officials usually focus on death numbers and rates which lag infections; and which fraction of infections is variable being highly dependent on the propensity, vulnerability, medical treatment and ages of the population so cannot be solely diffusivity dominated.

<sup>2</sup>For a source add to the RHS a term, say  $n'_t$ , being the number per unit time.

from observational data. Importantly, Equation (1) is a second order differential equation, so: (a) is fundamentally more general than the first order differential equations used in classic  $R_0$  and SEIR epidemiological spread models with multiple adjustable parameters; (b) contains a very limited number of variable but physically-based coefficients that can be directly tested against or fitted to data; and (c) therefore able to draw on the foundations and prior knowledge of classical physics and methods for heat and concentration diffusion. To illuminate the physics, we seek an analytical solution to Equation (1) that is testable against publically available data for the second and/or subsequent waves.

## 2. MATERIALS AND METHODS

### 2.1. Basic Theory and Postulates

Exact solutions of Equation (1) are numerical [1] whereas we only seek a working correlation to aid our understanding and enable rapid estimates for risk decision-making. For the second and successive “waves”, we make simplifying assumptions and reasonable approximations to derive a basic analytical solution form for the dynamic infection trends and the limiting physical societal transmission mechanism, the accuracy and applicability of which can be determined by comparisons to data. Hence, even if and as countermeasures are deployed, without complete isolation or elimination the virus spreads:

- (a) by diffusion in any local region(s) or cities, becoming inevitably embedded in the wider community (regions of higher infection numbers naturally infect regions with lower);
- (b) throughout any national or regional location, all the population is equally able to be randomly infected so the distribution of the virus is to first order homogenous and any and all infections are equally possible;
- (c) with fundamental physics limiting the extent the rate of subsequent community spreading is described by the classic Fick’s Law with diffusivity parameters averaged over the population/region/society; and,
- (d) being equally possible so affording a simple homogenous and one-dimensional approximation, with the number of infections,  $n$ , some fraction of the total possible,  $N$ , where usually  $n \ll \ll N$ , depending on the overall community transmission mechanisms, societal behaviors, and countermeasure effectiveness.

For some overall societal characteristic effective transmission scale,  $L$ , after an elapsed time measured in days,  $d$ , the rate of change of infections or number counted on any day,  $n(d, L)$ , in one dimension Equation (1) becomes,

$$\frac{\partial n}{\partial d} = D \frac{\partial^2 n}{\partial L^2} \tag{2}$$

Using the usual non-dimensional formulation,  $\theta = \frac{n - n_m}{n_{02} - n_m}$ , with nominal minimum achievable and second wave initial numbers,  $n_m$ , and  $n_{02}$ , respectively we have,

$$\frac{\partial \theta}{\partial d} = D \frac{\partial^2 \theta}{\partial L^2} \tag{3}$$

Prior physics and classic texts provides the analogous Fourier flux solution satisfying Equation (3) [1,17,18] and is the usual error function,

$$\theta(d) = 1 - \operatorname{erf} \left\{ \frac{L}{(4Dd)^{1/2}} \right\} \tag{4}$$

where, the term,  $L_D = (4Dd)^{1/2}$ , is an effective overall community viral diffusive penetration distance or societal transmission “scale”. In contrast to the classic multi-parameter SEIR models [13,16] there are now only two governing and physically-based parameters to be determined.

### 2.2. Approximate Correlation Derivation for Wave Timing and Magnitude

To guide our thinking for present correlation and risk trending purposes, for the large-scale ratios,  $L/(4Dd)^{1/2}$ , relevant to whole communities and regions during diffusion, we can retain just the first term in the series expansion of Equation (4) [see 18 #586], so:

$$\theta(d) = \frac{n(d) - n_m}{n_{02} - n_m} \approx L^{-1} \left( \frac{8Dd}{\pi} \right)^{1/2} e^{-\left( \frac{L^2}{8Dd} \right)} \tag{5}$$

Usually, the minimum rate number,  $n_m \ll n(d)$  and  $n_{02}$ , so  $\theta(d) \approx \frac{n(d)}{n_{02}}$ , and increases initially as  $(Dd)^{1/2}$ . The validity, accuracy and limitations of this (or any other) approximate analytical solution must be demonstrated by data. For any chosen region this simple solution Equation (5) has the sensible limiting conditions,

$n \rightarrow n_m$  as  $L \rightarrow \infty$ ,  $n \rightarrow n_{02}$  as  $d \rightarrow 0$ , and  $n \rightarrow \infty$  as  $d \rightarrow \infty$  for any given diffusivity,  $D$ . After the decrease from Peak 1, we can expect to observe an initially slow square root increase in infection numbers with time, followed by an inevitable exponential increase at longer elapsed times which is a prediction consistent with the observations (Figure 1). Writing Equation (5) in a convenient non-dimensional form, with  $L^* = L_D/L$ ,

$$\theta(L^*) = \left( \frac{2}{\pi} \right)^{1/2} L^* e^{-\left( \frac{1}{2L^*} \right)^2} \tag{6}$$

Using Equations (5) and (6) for describing the second wave requires determining just two adjustable but physically based parameters: the diffusion coefficient,  $D$  ( $m^2/day$ ), and the assumed characteristic or effective length scale,  $L$  ( $m$ ), assuming the Peak 1 parameters ( $G$  and  $k$ ) still prevail. Both of course depend on some integration/aggregation of all the exact overall societal viral transmission mechanisms (person-to-person, aerosols, crowds, contamination, random exposure...), which we simply assume to exist.

We can then also estimate the relative size and timing,  $d_{M2}$ , for attaining Peak 2 when the diffusive growth is eventually balanced, assuming the same NPI learning countermeasures are used (e.g. social distancing, improved hygiene, public awareness etc...). For Peak 1, the decrease in the increasing rate decreases when the daily transmission and incubation increase number,  $n_{M1}$ , is comparable to or balanced by countermeasures and public awareness. For some

learning constant,  $k$ , representing the overall countermeasures effectiveness and resilience during and after Peak 1 [10],

$$n_{M1}(d) \sim n_{01} e^{-(G-k)d_{M1}} \quad (7)$$

Here, the parameters for incubation growth for any initial viral variant,  $G$ , learning recovery,  $k$ , and peak day,  $d_{M1}$ , are known from the Peak 1 prior data trends (Figure 1), and allows quantification of countermeasures by evaluating,  $k$ . This second wave or “curve flattening” semi-plateau has an asymptote given by  $\frac{dn_2}{dd} \rightarrow 0$ , so from

Equation (5), Peak 2 tends to flatten as,  $d_{M2} \rightarrow \left(\frac{L^2}{8kD}\right)^{\frac{1}{2}}$ , i.e. gov-

erned by the ratio of the learning countermeasure reduction to the diffusion increase  $e$ -folding rates. The more effective the countermeasures ( $k$  increasing) the earlier the infection curve “flattens”. So, with countermeasures, the ratio,  $R^*$ , of the Peak 1 and Peak 2 infection rate numbers is the key measure, and for the nominal initial base or threshold daily rate numbers,  $n_{01}$  and  $n_{02}$ , respectively, from Equations (5) and (7),

$$R^* = \frac{n_{M2}}{n_{M1}} \approx \frac{n_{02}}{n_{01}} \left\{ \left( \frac{8Dd_{M2}}{L^2\pi} \right)^{\frac{1}{2}} e^{-\left( \frac{L^2}{8Dd_{M2}} - (G-k)d_{M1} \right)} \right\} \quad (8)$$

Given this approximate second plateau, from Equation (8) the predicted Peak 2 to Peak 1 rate size ratio is of order,  $R^* \approx \frac{n_{02}}{n_{01}} \left\{ (8Dd_{M2} / L^2\pi)^{1/2} e^{((G-k)d_{M1}-1)} \right\}$ , and is usually  $>1$  depending on effective initial learning and the ratio of the low base infection numbers,  $\frac{n_{02}}{n_{01}}$ . Using the ratio removes any dependency on uncertainties in actual case counts, testing variations and reporting protocols. The magnitude ratio for any additional successive (third plus) waves, e.g.  $\frac{n_{M3}}{n_{M2}}$ , and timing,  $d_{M3}$ , will now depend on the subsequent prevailing *virus variant* growth rate,  $G$ , and *effective* countermeasure,  $k$ , values after Peak 2, which latter factor can now include vaccination administration extent and effectiveness.

### 3. RESULTS

#### 3.1. Comparisons of Theory to Pandemic Second Wave Onset Data Trends

Excellent data for daily infection numbers are available in downloadable spreadsheet format [19] from January 2020, and from WHO, Johns Hopkins University and other reliable sources, including local state data from Departments of Public Health websites. As a reference origin, we adopt the initial detection/reportable threshold of  $n > 100$  total cases as a datum for the definite onset of significant infections,  $d = 0$ , and varying this threshold has only a minor impact as Peak 1 usually occurs within 20–30 days.

The origin of the minimum at the second wave onset,  $\theta(d_{02})$ , is the lowest daily infection count continuously achieved following the exponential decline from Peak 1, and used to normalize all subsequently increasing infection rates,  $R^* = \theta(d)/\theta(d_{02})$ . To encompass widely separated and differing societies, we performed test calculations for ongoing and continuing infection data showing distinct

**Table 1** | Country data with first peak and distinct second wave

Country/ region	Start $d = 0$ , 2020	Peak 1 rate	$\theta(d_0)$ day	$D$ (m <sup>2</sup> /day)	$L$ (m)
Australia	18 March	715	81	0.013	10
Canada	12 March	2760	53	0.012	7
France	5 March	4610	107	0.014	10
Italy	25 February	6560	100	0.014	12
UK	7 March	5290	120	0.014	9

second waves as listed in Table 1 for UK, Italy, France, Canada, and Australia (as of October 15, 2020). The latter is specifically chosen as being a distinct “island state” with comparatively low (but non-zero) rates but second waves (or “clusters”) despite extensive internal controls and border/entry restrictions.

Comparisons to the theory are shown in Figure 2, where the fits were derived using the range of  $0.012 < D < 0.014$  m<sup>2</sup>/s and  $7 < L < 12$  m-values as shown in Table 1 and are largely independent of country or societal culture. These fitted parameter values in Table 1 indeed are consistent with a whole society reaching a peak or plateau at  $d_{M2} = \left(\frac{L^2}{8kD}\right)^{1/2} \sim 200$  days using  $k \sim 0.02$ /day as we previously derived for the recovery from Peak 1 in Italy [9] and also applicable for UK and Turkey. The values for the long-term diffusion coefficient in Table 1 and Figure 2 are all  $O(10^{-2})$  m<sup>2</sup>/day, and for  $L \sim 10$  m and  $d \sim 100$  days we have,  $L/(4Dd)^{1/2} \sim 5$ , so the first term approximation for Equation (5) is reasonable.

The overall agreements with such disparate data are not perfect (Canada and UK exhibit over predictions) but the salient overall slow growing trend and 100–200 days timescales are reasonable (especially for Australia, France and Italy) given the approximations in the theory. Capturing the steep exponential onset is also inexact at present. The Australia case with literally different boundary conditions on infection control shows clearly that Peak 2 can also be reached, while terminating the second wave growth as also observed in the classic 1918 pandemic [4,20,21].

#### 3.2. Estimating Second Wave Magnitude and Uncertainties

When recovery from Peak 1 is not always complete,  $n_{02} > n_{01}$ , ensuring that Peak 2 daily rate is larger and what is actually observed, affecting emergency pandemic planning and decision-making. For countries already known to exhibit second wave peaks with recovery/decline, Australia actually had a Peak 1 to Peak 2 infection rate magnitude ratio,  $R^* = \frac{n_{M2}}{n_{M1}}$ , of 2.2 and Japan of 3. The still evolving

USA data had a first peak of circa 30,000 at 40 days which, after an initial decline at 150 days merged into a second “wave”, almost trending to a wavy persistent plateau and a still increasing rate of over some 2,000,000 per day, giving a peak ratio of at least  $\frac{n_{M2}}{n_{M1}} \sim 6$ .

Within any given society there is wide variation in *local* infection numbers and rates, so we made a more detailed “within country” analysis for the 10 US states which have exhibited a distinct or discernable early Peak 1 plus evidence of already reaching Peak 2 or some semi-plateau in daily rates after about 100+ days (and as of



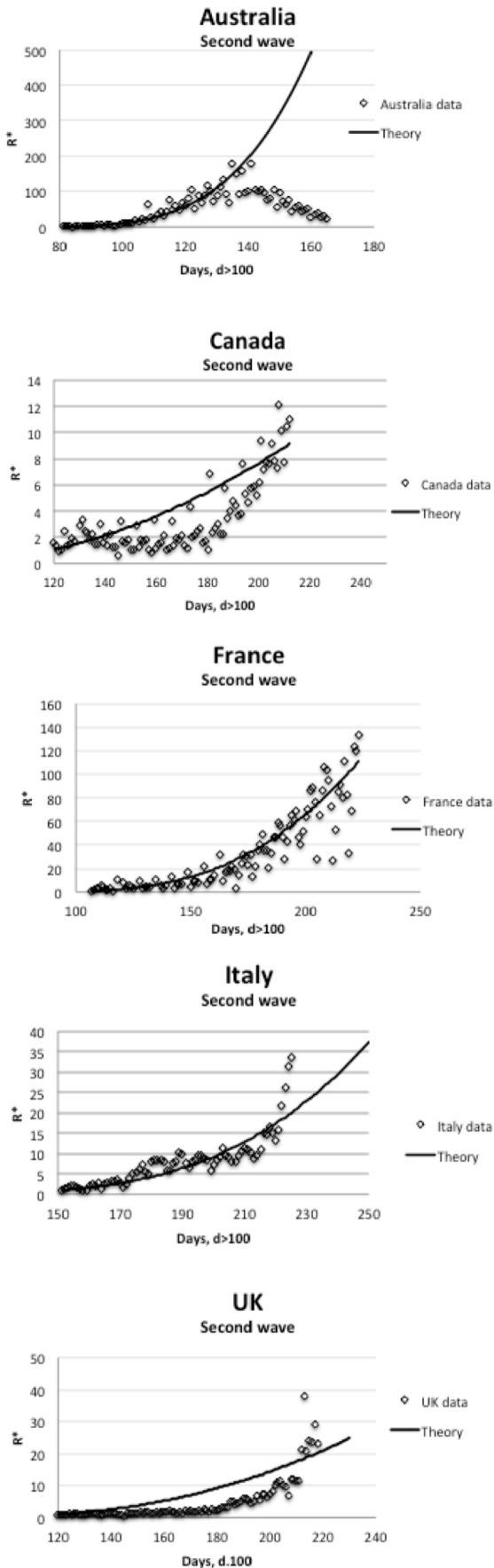


Figure 2 | Second wave trends compared to theory.

June 22, 2021, available state-by-state at <https://coronavirus.jhu.edu/By Region>”), As shown in Table 2, there is a wide range of ratios with an average of  $\frac{n_{M2}}{n_{M1}} = 4.1$ , comparable to the above national average. These observations can be compared to the predicted Peak 1 to Peak 2 number ratio from the diffusion model [Equation (8)]. Using the above observed and fitted values of  $dM_1 \sim 30$  days,  $dM_2 \sim 100$  days, with  $(G-k) \sim 0.15$  per day, gives  $R^* = \frac{n_{M2}}{n_{M1}} \approx 8 \left( \frac{n_{02}}{n_{01}} \right)$ , for  $D \sim 0.01$  m<sup>2</sup>/day,  $L \sim 10$  m, a second Peak 2 or semi-plateau of nearly an order of magnitude larger than the Peak 1 long preceding it. The uncertainty in this simple estimate depends on parameters which are location and policy specific, namely the balance,  $(G-k)$ , between infection growth and societal countermeasure effectiveness.

Following these predictions and estimates, we examined the apparently unique case of China<sup>3</sup> as: (a) being the original source of the pandemic reaching Peak 1 in 26 days of 4000 cases per day by February 13, 2020; (b) exhibiting a rapid decline from Peak 1 which followed learning theory and presaged global recovery trends [9]; (c) thereafter reporting low daily infection rates (claimed to be mainly importations), with a barely discernable second wave Peak 2 reported as 276 cases per day on July 31; and (d) maintaining almost complete control of population movement and mobility, with enforced NPI mandates and personal contact tracking. With low and fluctuating numbers, it is difficult to uniquely pinpoint the second wave onset day,  $d_{n_e}$ , and rate value,  $n_{02}$ , taken here as nominally day  $d = 100$  with just two reported cases per day, suggesting  $n_{02}/n_{01} \sim 2/100 = 0.02$ , which would give  $n_{M2}/n_{M1} \sim 0.16$ , compared to that reported or observed of  $276/4000 \sim 0.07$ .

The comparison of the data to the fitted theory [Equation (5)] is shown in Figure 3 using the same  $D = 0.012$  m<sup>2</sup>/day and  $L = 10$  m values found for the Table 1 countries. This similarity suggests the fundamental physical diffusion and societal transmission processes are globally the same; and that China exhibits the same long-term embedding in the community as everywhere else. In addition since the Peak 1 initial recovery in China is similar to Italy and others [10], the second peak should be  $R^* \approx 8 \left( \frac{n_{02}}{n_{01}} \right) \sim 8$ , close to that observed.

As further confirmation, currently, a wide range of second defined waves or plateaux after Peak 1 are on-going and have not yet

Table 2 | US state data with first peak and distinct second wave

US state	Peak 1 rate, $n_{M1}$	Peak 2 rate, $n_{M2}$	Ratio, $R^*$
Alabama	346	2143	6.2
Arkansas	323	992	3.1
Colorado	526	1139	2.2
Florida	1269	13965	11
Georgia	1333	4813	3.6
Idaho	131	878	6.7
Kansas	345	1040	3
Louisiana	1857	3840	2.1
Nebraska	645	1286	2
Washington	584	959	1.6

<sup>3</sup>This case study was suggested by Professor Francesco D'Auria as being of particular importance in terms of the total population at risk.

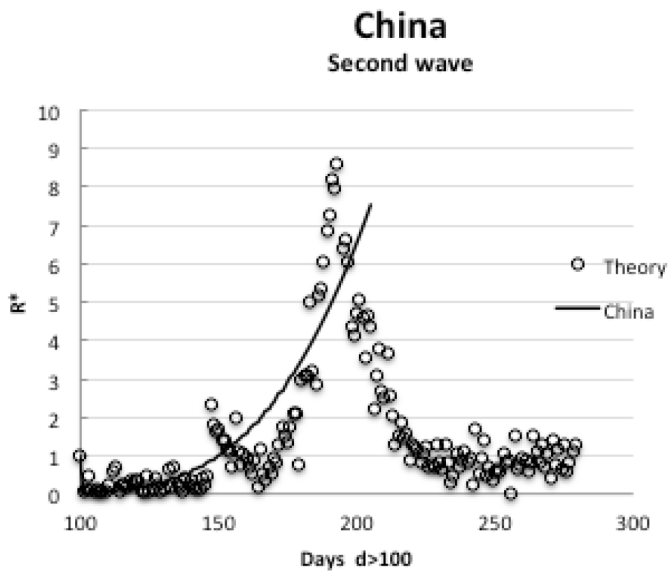


Figure 3 | Comparison of theory to second wave data in China.

declined [22]. Current ratio examples include: Austria (8), Belgium (9), Canada (2), Denmark (4), France (11), Germany (4), Indonesia (4), Iran (4), Israel (5), Italy (6), Nepal (9), Spain (3), Sweden (3), and UK (6). These overall ratios cover a similar range (2–12) to the US internal/regional state ratios in Table 2, and encompass the predicted estimate,  $R^* = \frac{n_{M2}}{n_{M1}} \approx 8 \left( \frac{n_{02}}{n_{01}} \right)$ , which surely is not a coincidence.

### 3.3. Local and Regional Peak Ratio Trends and Predictions

Insignificant initial infections may occur locally outside the initial peak regions (e.g. large cities like New York or Wuhan) but increase only after about 100–150 days solely due to slow progressive internal diffusive spread throughout the region, notably seen for US states like Alabama, Alaska, Wisconsin and some European countries. Being important for local public health risk planning purposes, we examined if an even small regional population internal to the country followed the same trends. Data were obtained directly from the state-run East Idaho Public Health District (EIPHD) covering Bonneville County, Idaho (where the author lives), with a local population only circa 120,000, a rural and urban setting without a first local first peak despite Idaho overall having a Peak 1 from infections cases occurring in other locations (Figure 1). The second wave arrived and emerged after about 100 days after the first onset, a similar timeframe to the second wave onset for the whole state data in Figure 2.

Figure 4 compares the data to Equation (8), with the starting,  $d = 100$  days after first exceeding 100 initial cases in the whole state (Figure 1). The “best” theory fit shown uses the same values for  $D$  (0.012 m<sup>2</sup>/day) and  $L$  (10 m) as for China and all the large countries in Table 1 and encompasses the data. This is a major and quite unexpected confirmation that the diffusion concept represents infection risk growth throughout whole and totally disparate societies with large differences in population and enforced or “mandated” countermeasures.

In passing, note that despite continued controls and countermeasures, New York State is currently exhibiting the symptoms of the onset of a second wave after day 150 (August 7, 2020), with cases slowly and inexorably rising. This is a pure prediction for a possible second wave of similar magnitude to Italy.<sup>4</sup>

For successive (third or more) waves arising in the ever-widening infected community, the assumption is each wave is superimposed on (or builds upon) any base or remnants of predecessor infections from the earlier wave(s). In USA and Japan, despite being totally distinct and distanced societies, distinct Peak 3 occurred after a further 100 days after Peak 2, so an initial scoping analysis was performed using the identical  $L$  and  $D$  values derived for Peak 2 to predict the USA Peak 3 growth, as shown in Figure 5.

Although a simplified and pessimistic over-prediction, the overall trends are reclaimed, with the ratio,  $\left( \frac{n_{M3}}{n_{M2}} \right) \sim 8/3 \sim 2.7$  in accord with Table 2  $\left( \frac{n_{M2}}{n_{M1}} \right)$  values, implying again that diffusion represents the limiting or controlling processes for any and all additional waves independent of societal countermeasures.

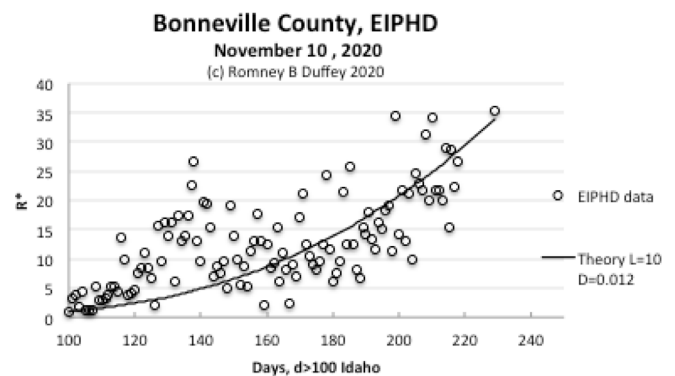


Figure 4 | Internal second wave onset in local region of a USA state.

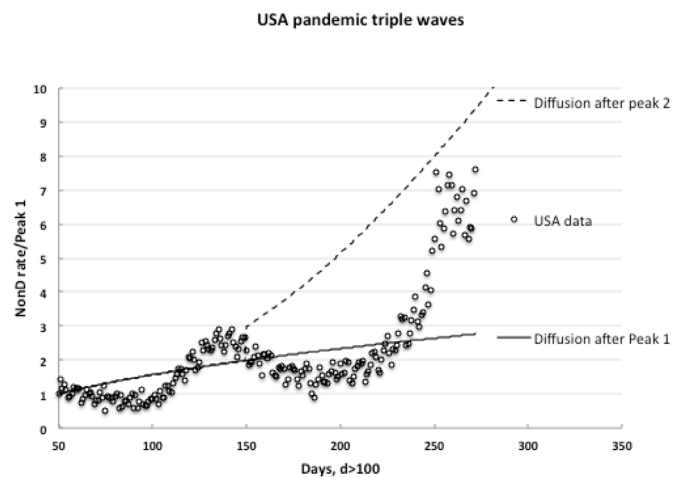


Figure 5 | Initial scoping third wave predictions for the entire USA (data for March 4 to November 10, 2020).

<sup>4</sup>Note added in March 2020: As predicted in October, such a second wave in fact occurred in Italy with Peak 2 in mid-November with rate magnitude ratio  $n_{M2}/n_{M1} \sim 8$ .

#### 4. DISCUSSION: THE RISK AND UNCERTAINTY OF LARGE SUCCESSIVE WAVES AND VACCINE EFFECTIVENESS

We need to pro-actively plan, but as far we know there have been no other second wave size and timing predictions with diffusion adopted as the dominant limiting physics for community spreading. Predicting second wave timing and magnitude is a difficult task with considerable uncertainty so we have tried to find a simple solution to this hard problem while not claiming complete accuracy. Despite significant variability and uncertainty, after a few 100 days second waves (up to a factor of 10 larger than Peak 1) are to be expected, and this elevated risk should be pro-actively planned and managed.

Given diffusion represents the limiting societal transmission phenomena, the present simple one-dimensional theory cannot replicate every detail for more complex scenarios but can plausibly explain as follows:

- (a) no discernable first peak appears because multiple numerous initially rapid infections dominate initial national counts, followed by more widespread cases over the 100's of days of slow diffusive timescale, notably as observed in Argentina and Brazil as well as in local regions.
- (b) only a slow recovery occurred after the initial rapid Peak 1, so many thousands of cases were still occurring during the diffusion masking the second wave onset which is then superimposed after about 150 days, as notably demonstrated by Russia and the USA.

While beyond the scope of the present paper, these detailed trends really require more complex intra-regional analysis. The limiting overall *societal* viral transmission mechanism being diffusion via widespread societal contamination and viral presence replaces the usual paradigm of traditional SEIR and  $C_0$  epidemiological models which assume the new infection rate is directly proportional to the number originally infected [e.g. 11–13,16].

Ascribing physical significance, the observed characteristic diffusive length scale of approximately 10 m is an effective *overall* societal transmission distance not inconsistent with feasible transmission dimensions [21]; and not coincidentally is a typical room, home or store dimension where people and societies interact. The order of magnitude for the coefficient,  $D$ , is a factor of more than 10,000 smaller than the  $10^{-2}$  to  $10^{-3}$  m<sup>2</sup>/s associated or expected for local airborne atmospheric particulate or direct aerosol spreading mechanisms [see, e.g. 23,24] reflecting the slow onset of second and other waves are simply the consequences of normal human and personal interactions and overall societal risk-taking behavior [25], as exemplified in the USA by the contrasting weak and strong policies adopted in Florida and California, respectively.

As to the important question of quantifying effectiveness of vaccines as a pharmaceutical intervention or countermeasure, in principle our learning model [9,10] allows direct estimation of the revised,  $k$ , value in Equation (7) describing the exponential decline from Peak 3, an exercise that is on-going. The new predictions and comparisons illuminate public health policy and risk planning. The present analysis *quantitatively* shows timing and learning and

adopting effective countermeasures reduce magnitudes but cannot eliminate infections completely. Since pandemic second and more waves persist, the only perfect diffusion “barrier” or so-called “circuit breaker” is indefinite isolation for everyone everywhere, which is neither feasible nor desirable.

#### 5. CONCLUSION: BEING PRO-ACTIVE IN RISK PREDICTIONS

We have established the principal that the timing and magnitude of long-term infection increases or successive waves can be based on the well-known physics of classical diffusion theory. This simple concept explains and predicts the successive “waves” or recurrences of higher infection numbers 100's of days after the first peak in case rates, with the subsequent inexorable dynamic community-wide spread and embedding of infections. If the pandemic started in localized centers, this diffusion process represents the inevitable but “imperfect” subsequent transmission of viral infections independent of strain due to prevalent, pervasive and persistent societal behaviors, and is a physical model fundamentally different from the proportional rate assumptions of traditional SEIR and  $C_0$  epidemiological models.

To enable original trend, size and timing predictions and obtain physical insight, we have derived explicit approximate equations for dynamic infection numbers using the simplest homogenous one-dimensional diffusion model. Two physically-based variables, a characteristic societal length scale and an effective community wide diffusion coefficient, have been fitted to the infection rate data for a wide range of countries and local regions of differing societal behaviors and countermeasures, showing that the effective diffusivity and characteristic scale are universal. The timing and relative size of the second and subsequent larger peaks or semi-plateau is also limited by the relative (in)effectiveness of NPI countermeasures and societal learning effects. Medical system planning and emergency bed count capacity projections encompass peak wave magnitude ratios,  $R'$ , of up to about 10 times the initial peak infection rate after about 100 days. We know of no other estimation method or published number for the risk of second and subsequent peaks and for estimating uncertainties. The policy and risk planning implications of such inexorable diffusive spread must be by recognizing potentially larger infection waves will inevitably occur and pro-actively planning for that eventuality.

#### CONFLICTS OF INTEREST

The author declares no conflicts of interest.

#### ACKNOWLEDGMENTS

The author is indebted to all the contributors to infection data collection and reporting that have made this work possible, and Ms. Geri Rackow for supplying the EIPHD data files. He also thanks Dr. E.D. Hughes for his important critique and suggestions; and Professors Enrico Zlo and Francesco D'Auria for their comments, encouragement and extended dialogs on this topic.

## NOMENCLATURE

$D$  – Effective community diffusion coefficient;  
 $d$  – Elapsed time in days;  
 $G$  – Initial characteristic infection rate;  
 $k$  – Learning countermeasure characteristic rate;  
 $L$  – Characteristic transmission societal scale;  
 $L^*$  – Diffusive scale ratio,  $(4Dd/L^2)^{1/2}$ ;  
 NPI – Non-pharmaceutical intervention;  
 $n$  – Number of infections in any day;  
 $R^*$  – Infection number ratio,  $R^* = \theta(d)/\theta(d_0)$ ;  
 $t$  – Time;  
 $\theta$  – Non-dimensional number ratio; [Equation (5)].

### Subscripts

$M1$  – Maximum value at Peak 1;  
 $M2$  – Maximum or asymptotic value of second wave;  
 $m$  – Minimum achieved or attainable value;  
 $0$  – Initial or beginning value;  
 $01$  – Initial value for onset of Peak 1;  
 $02$  – Initial value for onset of second wave.

## REFERENCES

- [1] Aciola PH. Diffusion as a first model of spread of viral infection. 2020. arXiv:2003.11449v2 [physics.ed-ph].
- [2] Sebhatu A, Wennberg K, Arora-Jonsson S, Lindberg SI. Explaining the homogeneous diffusion of COVID-19 nonpharmaceutical interventions across heterogeneous countries. *Proc Natl Acad Sci U S A* 2020;117:21201–8.
- [3] Glezen WP. Emerging infections: pandemic influenza. *Epidemiol Rev* 1996;18:64–76.
- [4] Barry JM. The great influenza: the story of the deadliest pandemic in history. New York: Penguin Random House Books; 2005.
- [5] US Centers for Disease Control (CDC). Coronavirus Disease 2019. US Centers for Disease Control; 2020 (accessed at [www.cdc.gov](http://www.cdc.gov)).
- [6] Binti Hamzah FA, Lau C, Nazri H, Ligot DV, Lee G, Tan CL, et al. CoronaTracker: world-wide COVID-19 outbreak data analysis and prediction. *Bull World Health Organ* 2020 [Preprint].
- [7] National Academies of Sciences, Engineering, and Medicine. Airborne transmission of SARS-CoV-2. In: Proceedings of a Workshop—in Brief. Washington, DC: The National Academies Press; 2020.
- [8] Covello VT, Hyer RN. COVID-19: simple answers to top questions, risk communication guide. Arlington, Virginia: Association of State and Territorial Health Officials (ASTHO); 2020.
- [9] Duffey RB, Zio E. Analysing recovery from pandemics by learning theory: the case of CoVid-19. *IEEE Eng Med Biol Soc Sect* 2020;8:110789–95.
- [10] Duffey RB, Zio E. Prediction of CoVid-19 infection, transmission and recovery rates: a new analysis and global societal comparisons. *Saf Sci* 2020;129:104854.
- [11] Heesterbeek JAP. A brief history of  $R_0$  and a recipe for its calculation. *Acta Biotheor* 2002;50:189–204.
- [12] Holme P, Masuda N. The basic reproduction number as a predictor for epidemic outbreaks in temporal networks. *PLoS One* 2015;10:e0120567.
- [13] Li J, Blakeley D, Smith RJ. The failure of  $R_0$ . *Comput Math Methods Med* 2011;2011:527610.
- [14] Bertuzzo E, Mari L, Pasetto D, Miccoli S, Casagrandi R, Gatto M, et al. The geography of COVID-19 spread in Italy and implications for the relaxation of confinement measures. *Nat Commun* 2020;11:4264.
- [15] Krishna MV. Mathematical modelling on diffusion and control of COVID-19. *Infect Dis Model* 2020;5:588–97.
- [16] Pedro SA, Ndjomatchoua FT, Jentsch P, Tchuenche JM, Anand M, Bauch CT. Conditions for a second wave of COVID-19 due to interactions between disease dynamics and social processes. *Front Phys* 2020;8:574514.
- [17] Carslaw HS, Jaeger JC. Conduction of heat in solids. Second edition. Oxford, UK: Oxford University Press; 1959, p. 310.
- [18] Mills AF. Heat transfer. Chapter 3. Illinois, USA: Irwin; 1992.
- [19] Ritchie H, Ortiz-Ospina E, Beltekian D, Mathieu E, Hasell J, Macdonald B, et al. Coronavirus (COVID-19) cases. 2020. Published online at [OurWorldInData.org](http://OurWorldInData.org). Available from: <https://ourworldindata.org/covid-cases> [Online Resource].
- [20] Andreasen V, Viboud C, Simonsen L. Epidemiologic characterization of the 1918 influenza pandemic summer wave in Copenhagen: implications for pandemic control strategies. *J Infect Dis* 2008;197:270–8.
- [21] Duffey RB, Zio E. CoVid-19 pandemic trend modeling and analysis to support resilience decision-making. *Biology (Basel)* 2020;9:156.
- [22] Dwight HB. Tables of integrals and other mathematical data. New York: MacMillan; 1968.
- [23] Gorbunov B. Aerosol particles laden with COVID-19 travel over 30m distance. *Preprints* 2020;2020040546.
- [24] Korovina NV, Zharova IK, Kudryashova OB, Titov SS. Diffusion coefficient when fine aerosol media propagate in a confined volume. *EPJ Web Conferences* 2016;110:01029.
- [25] Duffey RB. Learning about risk: life, death...and money. 2020.