



Estimation du nombre de cas de covid-19 en France et dans différents pays: homogénéisation basée sur la mortalité

Marc Dhenain

► To cite this version:

Marc Dhenain. Estimation du nombre de cas de covid-19 en France et dans différents pays: homogénéisation basée sur la mortalité. Bulletin de l'Académie Vétérinaire de France, 2020, 10.4267/2042/70840 . hal-03053239

HAL Id: hal-03053239

<https://hal.science/hal-03053239>

Submitted on 10 Dec 2020

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

ESTIMATION DU NOMBRE DE CAS DE COVID-19 EN FRANCE ET DANS DIFFÉRENTS PAYS : HOMOGENÉISATION BASÉE SUR LA MORTALITÉ

ESTIMATION OF COVID-19 CASES IN FRANCE AND IN DIFFERENT COUNTRIES: HOMOGENEISATION BASED ON MORTALITY

Par Marc DHENAIN^(1, 2, 3)

(Note soumise le 4 Avril 2020,

Acceptée le 6 Avril 2020, Mise à jour le 1^{er} Mai 2020)

RÉSUMÉ

Chaque jour une estimation du nombre de personnes touchées par le Covid-19 et un décompte de la mortalité sont effectués par les autorités de différents pays. Nous proposons que la mortalité rapportée dans chaque pays puisse être utilisée pour créer un indice du nombre de cas réels à un temps t_0 . Cet indice est:

$$C_{t_0\text{-estimé}} = (M_{t_0} / M_{t\text{-est}}) * (M_{t_0} / M_{t_0-3j})^6$$

Avec M_{t_0} : nombre de morts rapportés dans un pays au temps t_0 ; M_{t_0-3j} : nombre de morts rapportés dans un pays au temps t_0 moins 3 jours; $M_{t\text{-est}}$: Taux de mortalité estimé. En utilisant un taux de mortalité estimé de 2%, nous avons évalué le nombre de cas le 10 avril 2020 en Allemagne, Belgique, Chine, Corée du Sud, Espagne, France, Iran, Italie, Pays-Bas, Royaume-Uni et aux USA. Ce nombre atteignait 2 872 097 en France et 924 892 personnes en Allemagne. Ce travail suggère une très forte sous-estimation du nombre de cas de personnes touchées, avec un indice de notification souvent inférieur à 5%. La formule proposée permet également d'évaluer l'impact de politiques de prévention de la dissémination du virus.

Mots-clés : Covid-19, estimation du nombre de cas, mortalité, prévalence

ABSTRACT

Every day authorities of different countries provide an estimate of the number of persons affected by Covid-19 and a count of fatality. We propose to use the fatality reported in each country to provide a better estimate ($C_{t_0\text{-estimated}}$) of the number of cases at a given time t_0 .

$$C_{t_0\text{-estimated}} = (M_{t_0} / M_{t\text{-est}}) * (M_{t_0} / M_{t_0-3j})^6$$

With M_{t_0} : number of fatalities reported in a country at time t_0 ; M_{t_0-3j} : number of fatalities reported in a country at time t_0 minus 3 days; $M_{t\text{-est}}$: estimated fatality rate. Based on a fatality rate of 2%, we assessed the number of cases April 10th 2020 in Belgium, China, France, Germany, Iran, Italy, South Korea, Netherlands, Spain, United Kingdom and USA. This number reached 2,872,097 in France and 924,892 persons in Germany. This work suggests a very strong underestimation of the number of cases of people affected, with a notification index often less than 5%. The proposed formula also makes it possible to evaluate the impact of policies to prevent the spread of epidemic.

Key words: Covid-19, estimated number of cases, fatality, prevalence.

(1) Académie Vétérinaire de France, Académie Nationale de Médecine.

(2) Centre National de la Recherche Scientifique (CNRS), Université Paris-Sud, Université Paris-Saclay UMR 9199, Laboratoire des Maladies Neurodégénératives, 18 Route du Panorama, F-92265 Fontenay-aux-Roses, France.

(3) Commissariat à l'Énergie Atomique et aux Énergies Alternatives (CEA), Institut François Jacob, Molecular Imaging Research Center (MIRcen), 18 Route du Panorama, F-92265 Fontenay-aux-Roses, France.
Courriel : marc.dhenain@cea.fr

INTRODUCTION

L'infection par le coronavirus Sars-CoV-2 qui entraîne le Covid-19 s'est propagée à l'ensemble du monde et entraîne un nombre important de décès (European Centre for Disease Prevention and Control 2020). Chaque jour une estimation du nombre de personnes touchées et un décompte de la mortalité sont effectués par les autorités de différents pays (Dong et al. 2020, https://github.com/CSSEGISandData/COVID-19/tree/master/csse_covid_19_data/csse_covid_19_time_series). La connaissance du nombre de personnes touchées est un paramètre important pour développer des stratégies de protection des populations et de sortie de crise. Ce nombre révèle des différences flagrantes entre les pays et ne reflète que partiellement la réalité (Tableau 1). Par exemple, le jour où le nombre de morts avoisinait 3 000 personnes, la France comptait 44 550 personnes touchées contre beaucoup plus (80 537) pour la Chine et 122 171 pour l'Allemagne. Calculer le taux de mortalité (M_t) un jour (t_0) donné est une autre façon d'objectiver les différences entre pays. De prime abord,

$$M_t = M_{t_0} / C_{t_0} \quad (\text{Eq. 1})$$

Avec M_{t_0} = nombre de morts le jour t_0 ; C_{t_0} = nombre de cas rapportés le jour t_0 . Lorsqu'on observe les données fournies par chaque pays le jour le plus proche du 10 avril (date de cet article) où le nombre de morts avoisinait 3000 personnes (ou le dernier chiffre disponible lorsque les 3000 morts n'étaient pas atteints ce jour), trois pays (Allemagne, Corée du Sud, États-Unis) avaient des taux de mortalité (M_t) proches de 2% ; sept pays (Belgique, Espagne, France, Iran, Italie, Pays-Bas, et Royaume-Uni) avaient

des taux entre 6% et 12%, et la Chine présentait une valeur intermédiaire de 3,7% (Tableau 1).

Les patients qui meurent un jour donné ont été infectés beaucoup plus tôt. Le dénominateur du taux de mortalité devrait donc être le nombre total de patients infectés en même temps que ceux qui sont décédés (Baud et al. 2020). Cela est d'autant plus vrai que les vitesses d'évolution de la pandémie évoluent différemment dans divers pays: en mars 2020, le nombre de personnes touchées augmentait fortement de jour en jour en France, alors qu'il était stabilisé en Chine.

Une meilleure estimation du taux de mortalité est donc:

$$M_{t-x\text{jours}} = M_{t_0} / C_{t_0-x\text{jours}} \quad (\text{Eq. 2})$$

Avec $C_{t_0-x\text{jours}}$ = nombre de cas signalés le jour t_0 moins x jours, avec x = délai moyen entre l'apparition des symptômes et le décès. Une durée moyenne de 18 jours est rapportée entre l'apparition des symptômes du Covid-19 et le décès des patients (Ruan et al. 2020 ; Verity et al. 2020 ; Zhou et al. 2020). Ainsi, le taux de mortalité ajusté (M_{t-18j}) qui prend en compte ce retard moyen est (Flaxman et al. 2020).

$$M_{t-18j} = M_{t_0} / C_{t_0-18j} \quad (\text{Eq. 3, Tableau 1})$$

Avec C_{t_0-18j} = nombre de cas rapportés le jour t_0 moins 18 jours. Le calcul de M_{t-18j} révèle des écarts qui se creusent davantage entre les pays par rapport à M_t avec des variations qui vont de 2,3% (Corée du Sud) à plus de 700% pour l'Espagne.

En comparant M_{t_0} et C_{t_0-18j} dans différents pays (Figure 1), on voit une relation linéaire entre la mortalité à t_0 et le nombre de

cas à $t_0-18\text{jours}$ pour tous les pays (test de corrélation linéaire de Pearson, $p < 0,05$ sauf pour la Belgique ($p = 0,07$) en raison du petit nombre de points ($n = 3$)). Les pentes des droites de régression correspondant aux données varient considérablement d'un pays à l'autre, ce qui est cohérent avec les variations rapportées de M_{t-18j} . Les valeurs du taux de mortalité M_{t-18j} basées sur les cas signalés par différents pays ne sont donc pas fiables, en partie parce que le nombre de cas signalés dans différents pays n'est pas fiable (stratégies de dépistage différentes selon les pays).

Comment évaluer plus précisément le nombre de personnes touchées en utilisant une méthode similaire pour tous les pays. Nous proposons une méthode simple utilisant le nombre de décès signalés par chaque pays pour estimer et comparer le taux de personnes touchées par Covid-19. Cette méthode repose sur trois premières hypothèses :

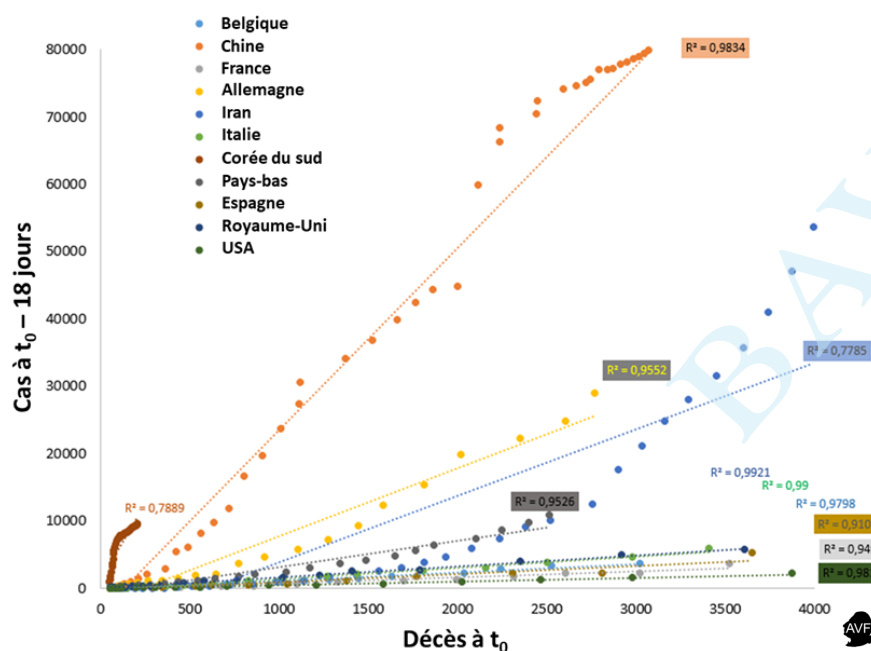


Figure 1: Relations entre les décès un jour donné (t_0) et le nombre de cas dix-huit jours avant (t_0-18 jours) dans différents pays. Le chiffre ne comprend que des valeurs comprises entre 50 et 4 000 décès (ou moins si le nombre de décès était inférieur dans le pays le 10 avril).

Pays	Méthodes 1&2					Méthode 1		Méthode 2					
	Population (million)	Cas signalés à t0 (Ct0)	Décès signalés à t0 (Mt0)	Cas signalés à t0-18j (Ct0-18j)	Cas estimés à t0-18j (23 mars) C(est-18j)	Taux de progression de t-18 à t0 P18j = Ct0 / Ct-18j (10 avril) P18j	Cas estimés (10 avril) Ct0-estimé	R0	Cas estimés (10 avril) Ct0-estimé	Cas estimés (10 avril) Ct0-estimé	Cas estimés (10 avril) Ct0-estimé	Cas estimés (10 avril) Ct0-estimé	Cas estimés (10 avril) Ct0-estimé
Délai							t0-18d	t0-18d	t0-18d	t0-18d	t0-18d	t0-12d	
Mt-est							2%	2%	2%	4%	15%	2%	
Belgique	11,476	26 667	3019	3 743	150 950	7,12	1 075 443	2,86	1 609 257	804 628	214 568	683 894	
Chine	1,384,688	82 941	3340	81 498	167 000	1,02	169 957	1,00	168 508	84 254	22 468	168 003	
France	67,795	124 869	13197	19 856	659 850	6,29	4 149 618	1,95	2 872 097	1 436 049	382 946	1 645 302	
Allemagne	83,073	122 171	2767	29 056	138 350	4,20	581 717	2,35	924 892	462 446	123 319	482 669	
Iran	82,022	68 192	4232	23 049	211 600	2,96	626 033	1,28	360 726	180 363	48 097	294 241	
Italie	60,360	147 577	18849	63 927	942 450	2,31	2 175 668	1,31	1 674 559	837 279	223 275	1 364 771	
Corée du Sud	51,709	10 450	208	8 961	10 400	1,17	12 128	1,25	16 811	8 406	2 242	14 306	
Pays-bas	17,282	23 097	2511	4 749	125 550	4,86	610 619	1,64	365 882	182 941	48 784	240 570	
Espagne	46,935	158 273	16081	35 136	804 050	4,50	3 621 909	1,46	1 811 469	905 735	241 529	1 345 788	
Royaume-Uni	65,761	73 758	8958	6 650	447 900	11,09	4 967 851	2,72	4 240 198	2 120 099	565 360	2 041 524	
Etats-Unis	328,240	496 535	18586	43 847	929 300	11,32	10 523 638	2,75	9 035 229	4 517 615	1 204 697	4 086 903	

Tableau 1 : Taux de mortalité dans différents pays lorsque le nombre de morts avoisinait 3 000 personnes (ou si ce nombre n'a pas été atteint, le dernier chiffre disponible le 10 avril 2020) (https://github.com/CSSEGISandData/COVID-19/tree/master/csse_covid_19_data/csse_covid_19_time_series).

1. Le nombre de décès signalés par chaque pays est fiable;
2. Le taux de mortalité (M_{t-est}) est connu et similaire dans différents pays;
3. Le délai moyen entre l'apparition des symptômes et la mort est connu (ici considéré comme 18 jours). Sur la base de ces hypothèses, on peut calculer le nombre de cas présentés dix-huit jours avant un jour donné (t_0). Deux méthodes sont alors proposées pour déduire le nombre de cas, dix-huit jours plus tard, à l'instant t_0 . La première s'appuie sur l'augmentation, en fonction du temps, du nombre de cas signalés dans les bases de données pendant ces 18 jours. La seconde modélise l'évolution pendant 18 jours du nombre de cas en s'appuyant sur le taux de variation quotidien du nombre de cas estimés 18 jours avant t_0 .

MÉTHODES

Estimation du nombre de cas au temps t_0 moins 18 jours sur la base du taux de mortalité

Une façon d'estimer le nombre de cas de Covid-19 est de déduire ce nombre de cas à partir du nombre de décès et du taux de mortalité rapporté par des études bien contrôlées en utilisant la formule suivante

$$C_{(est-18j)} = M_{t_0} / M_{t-est} \quad (\text{Eq.4, Tableau 2})$$

Avec $C_{(est-18j)}$: nombre de cas estimés 18 jours avant t_0 ; M_{t_0} = nombre de décès signalés au jour t_0 ; M_{t-est} = taux de mortalité estimé à partir d'études bien contrôlées. M_{t-est} peut être évalué à partir d'études bien contrôlées basées sur des résidents de Chine

continentale, des voyageurs revenant de Chine continentale, rapatriés de Chine, des passagers du bateau de croisière *Diamond Princess* (valeurs de 0,7 à 3,6% (Verity *et al.* 2020)). Ici, sur la base de cette dernière étude, nous proposons d'utiliser $M_{t-est} = 2\%$.

Connaissant $C_{(est-18j)}$ dix-huit jours avant t_0 , il reste à évaluer le taux de progression des cas au cours des 18 derniers jours pour estimer le nombre de cas au jour t_0 ($C_{t_0-estimé}$). Nous avons testé deux méthodes pour l'évaluer.

Estimation de la progression des cas pendant 18 jours sur la base du nombre de malades signalés

On peut supposer que la progression des cas estimés (P_{18j}) reflète l'augmentation en fonction du temps du nombre de malades déclarés au cours de la même période. Dans ce cas,

$$P_{18j} = C_{t_0} / C_{t-18j} \quad (\text{Eq. 5})$$

Avec C_{t_0} : nombre de cas notifiés dans un pays au temps t_0 ; C_{t-18j} : nombre de cas notifiés dans le pays au moment t_0 moins 18 jours. Donc,

$$C_{t_0-estimé} = C_{(est-18j)} * P_{18j} \quad (\text{Eq. 6, Tableau 2})$$

Estimation de la progression des cas pendant 18 jours sur la base du taux de variation sur 3 jours du nombre de cas estimés $C_{(est-18j)}$

Une autre option pour évaluer la progression des cas pendant 18 jours est de calculer le taux de variation quotidien (R_d) du nombre de cas estimés ou alternativement le taux de variation sur 3 jours (R_{3d}) du nombre de cas estimés

Pays	Jour (t_0)	Nombre morts à t_0 (M_{t_0})	Nombre cas à t_0 (C_{t_0})	Taux mortalité M_t	Nombre cas à t_0-18j (C_{t_0-18j})	Taux mortalité M_{t-18j}
Belgique	10 avril	3 019	26 667	11,3%	3743	80,7%
Chine	5 mars	3 015	80 537	3,7%	70 513	4,3%
France	30 mars	3 024	44 550	6,8%	2 281	132,6%
Allemagne	10 avril	2 767	122 171	2,3%	29 056	9,5%
Iran	1 avril	3 036	47 593	6,4%	12 729	23,9%
Italie	18 mars	2 978	35 713	8,3%	1 128	264,0%
Corée du Sud	10 avril	208	10 450	2,0%	8 961	2,3%
Pays-bas	10 avril	2 511	23 097	10,9%	4 749	52,9%
Espagne	24 mars	2 808	39 885	7,0%	400	702,0%
Royaume-Uni	2 avril	2 921	33 718	8,6%	1 140	256,2%
États-Unis	30 mars	2 978	161 807	1,8%	1 163	256,1%

Tableau 2 : Estimation du nombre de cas dans différents pays le 10 avril 2020 (t_0) à l'aide de différentes méthodes et d'un taux de mortalité (M_{t-est}) estimé à 2%. Le nombre de cas estimés avec différentes méthodes est fourni en utilisant des délais de 18 ou 12 jours entre l'apparition des symptômes et le décès.

$$R_{3d} = (C_t / C_{t-3j}) - 1 \quad (\text{i.e. } (C_t - C_{t-3j}) / C_{t-3j}) \quad (\text{Eq. 7})$$

Avec C_t : nombre de cas notifiés au temps t ; C_{t-3j} : nombre de cas notifiés au temps t_0 moins $3j$.

Le dernier jour où le calcul de R_{3d} est possible est 18 jours avant t_0 .

$$R_{3d-18j} = (C_{(est-18j)} / C_{(est-21j)}) - 1 \quad (\text{i.e. } (C_{(est-18j)} - C_{(est-21j)}) / C_{(est-21j)}) \quad (\text{Eq. 8})$$

Avec $C_{(est-18j)}$: nombre de cas estimés 18 jours avant t_0 (cf. Eq. 4); $C_{(est-21j)}$: nombre de cas estimé trois jours avant. En supposant que la progression des cas estimés suit un modèle exponentiel,

$$C_{t0-estimé} = C_{(est-18j)} * (1 + R_{3d-18j})^6 = C_{(est-18j)} * (M_{t0} / M_{t0-3j})^6 \quad (\text{Eq. 9})$$

Avec M_{t0-3j} : nombre de morts rapportés dans un pays au temps $t_0 - 3$ jours. L'exposant 6 représente la période du modèle car $6 * 3$ jours = 18 jours. Cette estimation suppose que R_{3d-18j} soit stable au cours des 18 derniers jours de l'estimation.

$$\text{Soit } C_{t0-estimé} = (M_{t0} / M_{t-est}) * (M_{t0} / M_{t0-3j})^6 \quad (\text{Eq. 10})$$

Comparaison avec le taux de reproduction de base du pathogène (R_0)

Le taux de reproduction de base d'un pathogène (R_0) est un indice important en épidémiologie. Il est défini comme le nombre moyen de cas secondaires résultant d'un cas primaire dans une population entièrement sensible. Un autre paramètre important est le temps de génération moyen (T_g), c'est-à-dire le temps entre l'infection d'un cas primaire et l'infection d'un cas secondaire (Flaxman *et al.* 2020; Keeling & Rohani, 2008). En supposant que la progression de Covid-19 suit un modèle SEIR basé sur quatre compartiments (sujets sensibles, exposés, infectieux et guéris (Keeling & Rohani, 2008)), l'augmentation de la prévalence pendant la phase d'invasion de la maladie est estimée à

$$C_{SEIR(t)} \approx C_{(t0)} \exp \{[(\sqrt{R_0} - 1) * \gamma] * t\} \quad (\text{Eq. 11, Keeling \& Rohani, 2008})$$

Avec C : nombre de cas infectés à un instant t ou t_0 ; R_0 : taux de reproduction de base du pathogène; γ : le taux de récupération γ avec $1 / \gamma$: période contagieuse (D_i).

Cette équation permet d'estimer le temps de doublement pour les cas (T_2) si $C_{SEIR}(t) / C(t_0) = 2$.

$$T_2 = \ln(2) / [(\sqrt{R_0} - 1) * \gamma] \quad (\text{Eq. 12})$$

$$\text{donc } R_0 = [(\ln(2) / (T_2 * \gamma)) + 1]^2 \quad (\text{Eq. 13})$$

$$T_g \approx D_e + D_i / 2 \quad (\text{Eq. 14})$$

Si le Covid-19 est modélisé avec un modèle SEIR (avec D_e = temps d'exposition pendant lequel un sujet est exposé mais pas contagieux), et $D_e \approx D_i$ (Li *et al.*, 2020).

On peut donc estimer que

$$R_0 = [2 * T_g * \ln(2) / (3 * T_2)] + 1 \quad (\text{Eq. 15})$$

En utilisant le taux de variation quotidien (R_d) (ou $R_{3d} / 3$), et la «règle des 70», on peut estimer

$$T_2 = 0,7 / R_d \quad (\text{Eq. 16})$$

$$\text{Ainsi } R_0 = [2 * T_g * \ln(2) / (3 * 3 * 0,7 / R_{3d})] + 1 \quad (\text{Eq. 17})$$

T_g est estimé égal à 6,5 d'après (Flaxman *et al.* 2020).

RÉSULTATS

Estimations reposant sur le nombre de cas au moment t_0 moins 18 jours

En utilisant l'équation 4 et les bases de données internationales (Dong *et al.* 2020, https://github.com/CSSEGISandData/COVID-19/tree/master/csse_covid_19_data/csse_covid_19_time_series), nous avons estimé le nombre de personnes touchées par le Covid-19 dix-huit jours avant le 10 avril 2020 dans différents pays (**Tableau 2**). Cette estimation était de 659 850 personnes en France et de 138 350 personnes en Allemagne. Nous avons ensuite proposé deux méthodes pour déduire le nombre de cas 18 jours plus tard (**Figure 2**). La méthode basée sur l'évaluation P_{18j} utilisant le nombre de cas déclarés par chaque pays au cours de la même période a suggéré une diminution du nombre de cas en fonction du temps dans certains pays (par exemple en Allemagne (**Figure 2C**) ou aux États-Unis (**Figure 2D**)), ce qui n'est pas cohérent. La méthode basée sur l'évaluation de R_{3d} a fourni une meilleure correspondance avec les cas estimés dans tous les pays testés (**Figure 2**). Nous avons donc conservé les résultats de la méthode R_{3d} pour la suite de nos analyses. Le nombre de cas estimés au 10 avril était de 2 872 097 en France, 924 892 en Allemagne, 1 811 469 en Espagne, 4 240 198 au Royaume-Uni et 9 035 229 aux États-Unis (**Tableau 2, Figure 2, Figure 3A**). L'évaluation R_{3d} peut en outre être utilisée pour estimer le taux de reproduction de base du pathogène (R_0) en utilisant l'équation 17. Nous avons trouvé des valeurs R_0 de 1,0 (Chine) à 2,86 (Belgique) (**Tableau 2**).

Cette analyse est basée sur un taux de mortalité estimé à partir d'études bien contrôlées (M_{t-est}) de 2%. Le nombre estimé de cas doit être divisé par deux si le taux de mortalité utilisé passe de 2 à 4% (**Tableau 2**). Il doit être doublé si le taux de mortalité utilisé passe de 2 à 1%. Certains auteurs suggèrent que le taux de mortalité réel pour Covid-19 pourrait être de 5,6 à 15,6% (Baud *et al.* 2020). Si le calcul utilise un taux de mortalité de 15%, le nombre estimé de cas tombe à 382 946 pour la France, mais il devient inférieur au nombre de cas effectivement signalés pour certains pays (par exemple 2 242 contre 10 450 pour la Corée du Sud), ce qui n'est pas cohérent (**Tableau 2**).

Dans notre étude, nous avons fixé le délai entre l'apparition des symptômes et la mort à 18 jours sur la base de données solides issues de la littérature (Ruan *et al.* 2020; Verity *et al.*

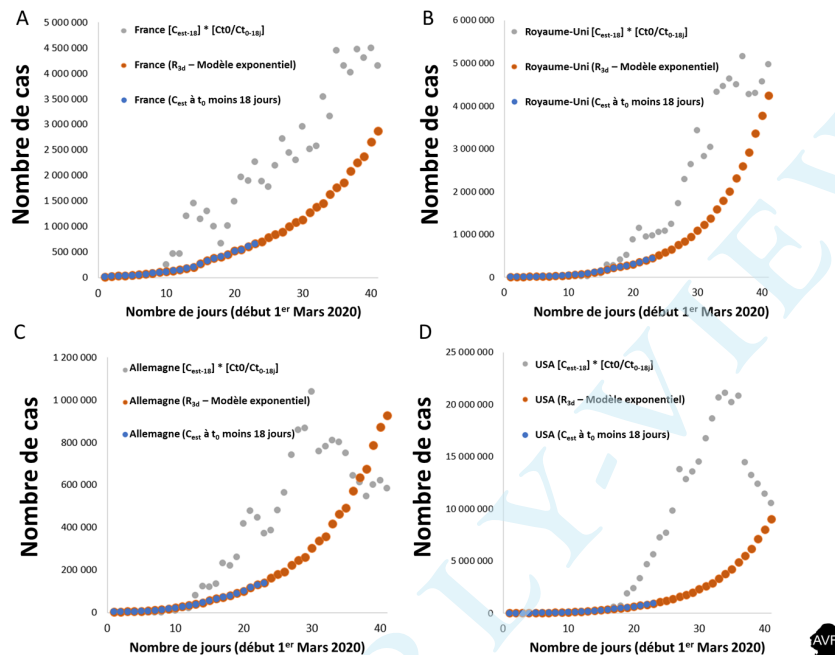


Figure 2 : Comparaison de l'évolution des cas estimés de Covid-19 en France (A), au Royaume-Uni (B), en Allemagne (C) et aux États-Unis (D) du 1er mars au 10 avril 2020. Les cas estimés 18 jours avant t_0 (Cest-18j), basés sur le nombre des décès à t_0 sont affichées en bleu. La progression de ces cas en prenant en compte un facteur multiplicatif ($P18j$) qui reflète l'augmentation en fonction du temps du nombre de cas déclarés au cours de la même période est notée en gris. Cette méthode de calcul conduit à un grand nombre de cas et à des variations quotidiennes. Les marques oranges correspondent à un modèle basé sur R_{3d} ($(1+Cest-3j) * R_{3d}$). Il fournit des courbes qui suivent les valeurs correspondant à Cest à partir de décès estimés 18 jours avant un jour donné (marques bleues).

2020; Zhou *et al.* 2020) et des délais utilisés dans d'autres modèles (Flaxman *et al.* 2020). Réduire ce délai, par exemple à 12 jours, diminue fortement le nombre de cas estimés (par exemple 1 645 302 pour la France (**Tableau 2**)) bien qu'il reste élevé par rapport aux chiffres rapportés par la plupart des pays.

En utilisant des estimations basées sur le modèle R_{3d} , avec un délai de 18 jours entre l'apparition des symptômes et la mort ainsi qu'un taux de mortalité de 2%, nous avons pu comparer les cas estimés de Covid-19 dans différents pays (**Figure 3A**), la proportion de cas dans différents pays (**Figure 3B**), ainsi que des indices de notification, c'est-à-dire la capacité de signaler les cas (**Figure 3C-D**). Ces données mettent en évidence de fortes disparités entre les pays. Elles suggèrent une proportion élevée de personnes affectées en Belgique et montre également des indices de notification qui varient de 60 à 80% en Corée mais inférieurs à 5% dans la plupart des pays.

Comparaison des cas estimés avec le nombre de cas rapportés à postériori

Le nombre de cas évalués entre le 24 mars et le 10 avril 2020 à partir du nombre de cas au temps t_0 moins 18 jours et le modèle R_{3d} a été comparé aux cas estimés à partir du nombre de cas au temps t_0 moins 18 jours (sans le modèle R_{3d}) calculé à partir des données de mortalités recueillies entre le 11 et 28 avril 2020 (**Figure 4**). Les cas estimés avec le modèle R_{3d} étaient plus élevés que ceux estimés à postériori. Cela peut être expliqué par la réduction de la propagation de la maladie (et donc de R_{3d}) au cours des 18 derniers jours suite aux mesures de confinement mises en place dans de nombreux pays. Le taux de variation sur trois jours (R_{3d}) du nombre de cas estimés que nous avons utilisé est surévalué.

Les mesures que nous avons effectuées permettent d'évaluer l'impact des politiques de confinement. Par exemple, pour la France, les cas estimés le 10 avril reposaient un R_{3d} de 0,28 (mesuré le 23 mars). Cinq jours avant cette date (le 19 mars) le R_{3d} était de 0,50. L'utilisation d'un R_{3d} de

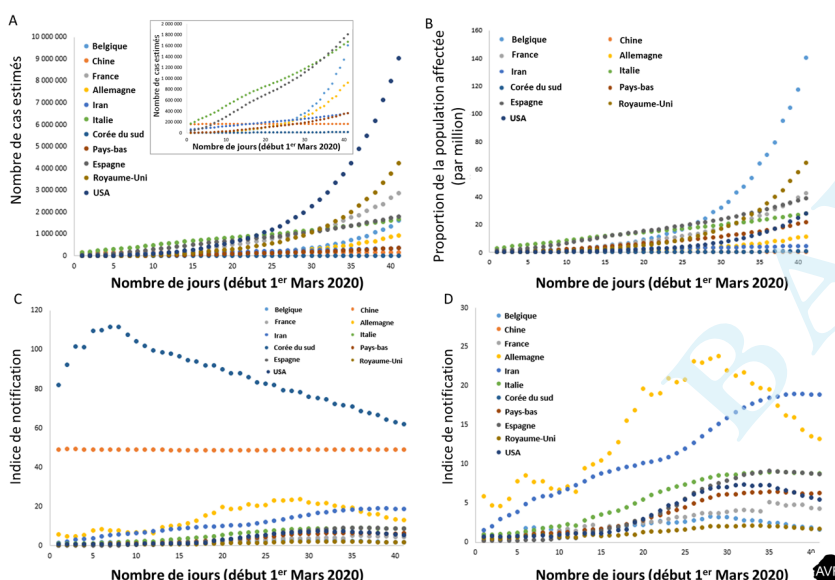


Figure 3 : Comparaison des cas estimés et des paramètres associés dans différents pays. A. cas estimés. B. Proportion de personnes affectées par rapport à la population du pays. C-D. Indices de notifications reflétant le nombre de cas signalés par différents pays par rapport au nombre estimé de cas (pourcentages).

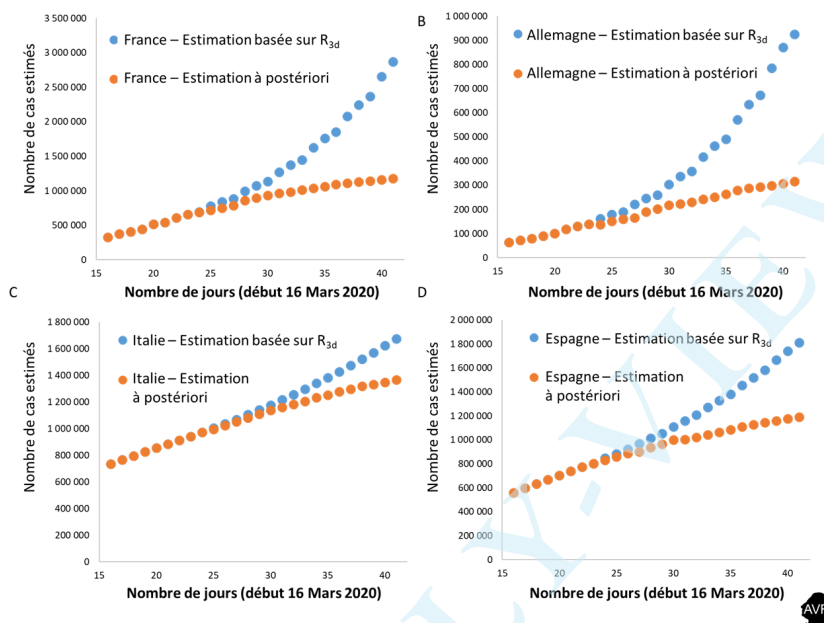


Figure 4 : Comparaison des cas estimés dans différents pays à partir du jour 16 (16 mars) en utilisant une méthode basée sur les taux de mortalité déclarés du 16 mars au 10 avril avec une estimation basée sur R_{3d} (bleu). Les résultats d'une méthode à postériori suivant les taux de mortalité signalés du 3 avril au 28 avril sont représentés en orange.

0,50 (au lieu de 0,28) mène à une estimation de 7 516 104 cas le 10 avril. La réduction du R_{3d} de 0,50 à 0,28 grâce au confinement a donc empêché l'apparition de 4 644 007 nouveaux cas en France (7 514 104 - 2 872 097). Le nombre de cas estimés à partir des décès survenus le 28 avril 2020 était en fait de 1 181 450. Ainsi, le nombre de cas réellement évités est de 7 514 104 - 1 181 450, soit 6 332 654. Cela représente 126 653 décès évités.

DISCUSSION

Il est essentiel d'évaluer le nombre de personnes touchées par le Covid-19 dans tous les pays du monde pour endiguer cette crise. Nous proposons d'utiliser la mortalité rapportée par chaque pays à un temps t_0 pour créer un indice du nombre de cas réels à ce même temps t_0 . La mortalité à t_0 permet d'estimer le nombre de cas 18 jours plus tôt ($C_{(est-18j)}$). Ensuite, le taux de variation au cours du temps des cas estimés ($C_{(est)}$) est évalué sur 3 jours (R_{3d}). Ce taux est utilisé pour moduler $C_{(est-18j)}$ et calculer le nombre de cas au temps t_0 . Ce calcul mène à l'équation suivante

$$C_{t0-estimé} = (M_{t0} / M_{t-est}) * (M_{t0} / M_{t0-3j})^6 \quad (\text{Eq. 10})$$

Avec M_{t0} : nombre de morts rapportés dans un pays au temps t_0 ; M_{t-est} : nombre de morts rapportés dans un pays au temps t_0 moins 3 jours ; M_{t-est} : Taux de mortalité estimé (2%). Cette analyse repose sur quatre hypothèses :

1. Le nombre de décès signalés par chaque pays est fiable,
2. Le taux de mortalité estimé parmi les personnes touchées est connu (M_{t-est} , ici considéré comme 2%),

3. Le temps moyen entre l'apparition des symptômes et la mort sont connues (ici considérées comme 18 jours).

4. Le taux de variation sur trois jours des cas estimés (R_{3d}) ne change pas au cours des 18 derniers jours.

Cette dernière condition n'est pas entièrement exacte car, grâce aux politiques de confinement, le taux de variation sur trois jours a diminué continuellement jusque début Mai 2020. Le $C_{t0-estimé}$ est ainsi surestimé comme le montre la comparaison des valeurs obtenues avec des mesures à posteriori c'est-à-dire sans tenir compte de la modélisation des 18 derniers jours. L'estimation du taux de variation sur 3 jours du nombre de cas estimés ($C_{(est-18j)}$ - Eq.4) peut être utilisée pour évaluer R_0 . Les valeurs de R_0 que nous avons rapportées (de 1,0 (Chine) à 2,86 (Belgique)) sont cohérentes avec les données de la littérature (par exemple $R_0 = 4,0$ dans Flaxman *et al.* 2020, au début de la pandémie). À l'inverse, l'estimation de R_0 sur la base de modèles épidémiologiques (cf. par exemple (Flaxman *et al.* 2020)) pourrait être utilisée pour calculer R_{3d} (cf. Eq. 17) et affiner l'estimation du nombre de cas. Il serait également possible de lisser les risques de variations journalières de R_{3d} par l'utilisation de mesures moyennes de R_{3d} sur une période de temps plus longue.

Nos analyses ont montré que le nombre de cas de Covid-19 dans de nombreux pays dépasse largement le nombre de cas présentés dans les bases de données internationales (2 872 097 (ou 1 181 450 cas avec des mesures à posteriori) contre 124 869 pour la France au 10 avril 2020). Les valeurs très élevées des cas estimés que nous rapportons sont cohérentes avec celles évaluées avec une autre méthode par (Flaxman *et al.* 2020). Par exemple, nous signalons 1,8 million de cas en Espagne tandis que Flaxman rapporte 7,0 millions le 28 mars. Notre calcul repose sur une méthode relativement simple alors que celle de Flaxman utilise des analyses plus complexes (modèle bayésien hiérarchique semi-mécanistique). Notre modèle a utilisé un taux de mortalité de 2% tandis que plusieurs études internationales très bien contrôlées ont rapporté des taux de 0,7 à 3,6% (Verity *et al.* 2020). Des valeurs de 0,5 à 4% pourraient donc être d'autres options raisonnables pour estimer le taux de mortalité. L'une des limites de notre modèle est que les taux de mortalité peuvent changer d'un pays à l'autre, par exemple en fonction de la distribution de la population de différents groupes d'âge qui ont une sensibilité différente au Covid-19. Il est également possible que le taux de mortalité évolue dans le temps dans un pays donné, par exemple en raison de la saturation des hôpitaux ou de la correction des chiffres de mortalité pour inclure des cas non comptabilisés (exemple de la France entre le 1^{er} et le 4 Avril

2020 pour inclure la mortalité en EHPAD). Nous avons fixé une valeur unique pour le temps entre l'apparition des symptômes et la mort (18 jours). En réalité, ce temps est variable avec un intervalle crédible à 95% de 16,9 à 19,2 ou plus selon (Verity *et al.* 2020). Nous avons cependant considéré que l'utilisation d'un tel intervalle rendrait le modèle plus compliqué sans ajouter de forte fiabilité par rapport à d'autres sources potentielles d'erreurs. Notre analyse repose uniquement sur le nombre de personnes décédées avec des cas confirmés de Covid-19. Il est donc essentiel que tous les pays soient en mesure de fournir des valeurs très fiables de décès liées au Covid-19. Enfin, notons que pour connaître le nombre de cas réels dans un pays à un moment donné, il faut soustraire des estimations présentées ici le nombre de personnes guéries, y compris celles dont la maladie n'a pas été identifiée.

Pour conclure, notre modèle interroge sur le petit nombre de personnes signalées comme touchées par le Covid-19 dans la plupart des pays par rapport aux grands nombres que nous estimons. Cette différence pourrait être expliquée par une forte sous-estimation du «taux de mortalité». Par exemple, en France, il faudrait changer le «taux de mortalité estimé» de 2 à ~ 46% pour diminuer l'estimation du nombre de cas de 2 872 097 à 124 869. De toute évidence, un taux de mortalité de 46% n'est pas observé. Ainsi, la seule explication qui reste est que la plupart des pays sous-estiment fortement les cas de personnes touchées. L'intérêt secondaire de notre modèle est qu'il prend en compte le taux de changement sur 3 jours, au cours des dix-huit derniers jours, il peut être utilisé pour modéliser l'efficacité des politiques de prévention de la propagation du Covid-19.

REMERCIEMENTS

Nous remercions Matthieu Domenech de Cellès pour ses conseils constructifs pendant la rédaction de ce manuscrit.

CONFLIT D'INTÉRÊT

L'auteur ne déclare aucun conflit d'intérêt dans la rédaction de cette note qui n'exprime que son opinion personnelle.

BIBLIOGRAPHIE

- Baud D, Qi X, Nielsen-Saines K, Musso D, Pomar L, Favre G. Real estimates of mortality following COVID-19 infection [published online ahead of print, 2020 Feb 19]. *Lancet Infect Dis.* 2020. doi: 10.1016/S1473-3099(20)30195-X.
- Dong E, Du H, Gardner L. An interactive web-based dashboard to track COVID-19 in real time [published online ahead of print, 2020 Feb 19]. *Lancet Infect Dis.* 2020;S1473-3099(20)30120-1. doi:10.1016/S1473-3099(20)30120-1. Base de données disponibles sur : https://github.com/CSSEGISandData/COVID-19/tree/master/csse_covid_19_data/csse_covid_19_time_series
- Consulté le 10/04/20
- Flaxman S, Mishra S, Gandy A, Unwin JT, Coupland H, Mellan TA *et al.* 2020. Estimating the number of infections and the impact of nonpharmaceutical interventions on COVID-19 in 11 European countries. Imperial College London (30-03-2020). doi: <https://doi.org/10.25561/77731>. Disponible à : <https://www.imperial.ac.uk/media/imperial-college/medicine/sph/ide/gida-fellows-hips/Imperial-College-COVID19-Europe-estimates-and-NPI-impact-30-03-2020.pdf> consulté le 05/04/20.
- European Centre for Disease Prevention and Control. Situation Update-Worldwide. Disponible à : <https://www.ecdc.europa.eu/en/geographical-distribution-2019-ncov-cases>. Consulté le 04/04/20
- Keeling, M.J. & Rohani, P. 2008. Modeling infectious diseases in humans and animals. Chapter 2- Introduction to simple epidemic models. Princeton University Press.
- Li, R., Pei, S., Chen, B., Song, Y., Zhang, T., Yang, W., Shaman, J. 2020. Substantial undocumented infection facilitates the rapid dissemination of novel coronavirus (SARS-CoV2). *Science.* doi: 10.1126/science.abb3221
- Ruan Q, Yang K, Wang W, Jiang L, Song J. Clinical predictors of mortality due to COVID-19 based on an analysis of data of 150 patients from Wuhan, China [published online ahead of print, 2020 Mar 3]. *Intensive Care Med.* 2020;1-3. doi:10.1007/s00134-020-05991-x
- Verity R, Okell LC, Dorigatti I, Winskill P, Whittaker C, Imai N *et al.* Estimates of the severity of coronavirus disease 2019: a model-based analysis [published online ahead of print, 2020 Mar 30]. *Lancet Infect Dis.* 2020;S1473-3099(20)30243-7. doi:10.1016/S1473-3099(20)30243-7
- Zhou F, Yu T, Du R, Fan G, Liu Y, Liu Z *et al.* Clinical course and risk factors for mortality of adult in patients with COVID-19 in Wuhan, China: a retrospective cohort study. *Lancet.* 2020 ; 395: 1054-1062.