



Quatre questions clés sur les indicateurs de suivi de l'épidémie de COVID-19 en France

Octobre 2020

Valérie Golaz*, Bénédicte Gastineau**, Marie-Laurence Flahaux**, Stéphanie Dos Santos**

* Chercheur, démographe, Institut national d'études démographiques, Laboratoire Population-Environnement-Développement (IRD, Aix Marseille Université)

** Chercheur, démographe, Institut de recherche pour le développement, Laboratoire Population-Environnement-Développement (IRD, Aix Marseille Université)

« Le taux de positivité dépasse 10% » « Explosion de l'incidence » « 17% de tests positifs en île de France, un niveau jamais atteint »

Dans le cadre du suivi épidémiologique de la COVID-19, des indicateurs concernant les tests virologiques sont publiés chaque semaine en France depuis le mois de juin dernier. Il s'agit en particulier du nombre et de la proportion de tests positifs, ainsi que du nombre de nouveaux cas et de l'incidence de la COVID-19. Ces chiffres sont repris par les médias et comparés à ceux du printemps 2020 et aux résultats des pays voisins. Ces comparaisons sont parfois utilisées pour appuyer les mesures mises en place ou les remettre en question. Quelle est la source de ces indicateurs ? Les résultats des tests reflètent-ils l'emprise réelle de l'épidémie en France ? Que représentent ces quatre indicateurs ? Pourquoi sont-ils produits sur une semaine et par département ?

1 - Quelle est la source des indicateurs de suivi de l'épidémie de la COVID-19 ?

Une partie des indicateurs de suivi de l'épidémie de COVID-19 en France sont issus des tests de dépistage.

Ces chiffres officiels sont fournis par Santé publique France, l'agence nationale de santé publique. La surveillance épidémiologique s'appuie actuellement sur le « système d'informations de dépistage » (SI-Dep). Ce système, opérationnel depuis le 13 mai 2020, vise au suivi de l'ensemble des patients testés en France dans les laboratoires privés et dans les laboratoires hospitaliers. Les tests pris en compte dans ce système d'information sont exclusivement les tests virologiques « RT-PCR »¹. Officiellement, tous les tests RT-PCR effectués sont enregistrés : il s'agit d'un suivi exhaustif. Avant le 13 mai, la surveillance virologique s'appuyait sur les données partielles, transmises à Santé publique France par un réseau de laboratoires privés et par les laboratoires hospitaliers. Ces transformations

¹ L'acronyme RT-PCR signifie *Reverse Transcription Polymerase Chain Reaction* en référence à la procédure suivie pour identifier l'ARN du virus et en mesurer la quantité : transformation en ADN puis amplification par réaction en chaîne.

de l'enregistrement des résultats des tests sont décrites chaque semaine dans le point épidémiologique publié par Santé publique France, disponible en ligne².

En France, les tests RT-PCR s'effectuent par prélèvement nosocomial (au fond du nez), car c'est là que la charge virale est la plus importante. Ils permettent de déterminer si, au moment du prélèvement, l'individu concerné porte le virus et le porte en suffisamment grande quantité pour être considéré comme 'positif'³. Il existe à l'heure actuelle d'autres familles de tests sur le marché français, qui ne sont pas comptabilisés dans les résultats rendus publics (tests sérologiques, tests rapides d'« orientation diagnostique » sérologiques ou virologiques). En cas de résultat positif, il est toujours préconisé de confirmer le diagnostic par un test RT-PCR. Seuls les tests RT-PCR sont utilisés pour produire des indicateurs et suivre l'épidémie.

2 – Les résultats des tests reflètent-ils l'emprise réelle de l'épidémie sur la population de la France ?

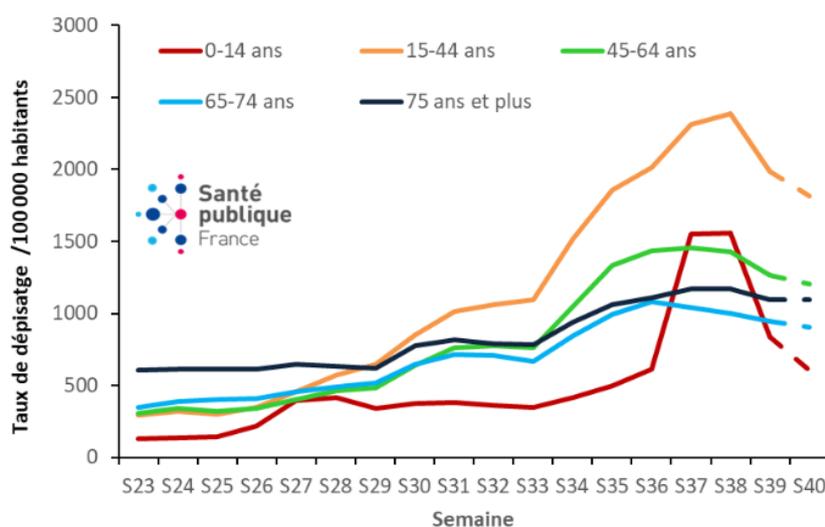
Les résultats des tests effectués ne reflètent pas l'emprise exacte de la maladie dans la population. Il n'y a pas d'enquête représentative régulière qui permette d'avoir une évaluation de la circulation du virus dans la population française au fil de l'année 2020⁴. Dans le cas de la COVID-19, savoir qui est atteint ou ne l'est pas est d'autant plus difficile qu'une grande partie des personnes qui contractent le virus ne développent pas de symptômes. Pour que les indicateurs représentent l'ampleur réelle de l'épidémie dans la population, il faudrait que toutes les catégories de la population soient testées chaque semaine de la même façon, en tenant compte du fait que la sensibilité à la maladie peut varier en fonction de nombreux critères, notamment de l'âge. Or, les dépistages, tels qu'ils sont réalisés, ne permettent pas d'en tenir compte puisqu'ils privilégient les personnes qui présentent des symptômes ou qui ont une santé à risque. Par exemple, le taux de dépistage n'est pas proportionnel à la population de chaque groupe d'âge : la figure 1 nous montre entre autres que les 45-64 ans (courbe verte) ont d'abord été moins testés que les 75 ans et plus (en bleu marine), puis autant (semaines 32 et 33, les deux premières semaines d'août), puis davantage. Le groupe d'âge des 45-64 ans est pourtant plus de deux fois plus nombreuse que celle des 75 ans et plus dans la population française. Une bonne représentation dans les tests impliquerait donc qu'on ait plus de deux fois plus de tests chez les 45-64 ans que chez les 75 ans et plus.

² Point épidémiologique national accessible depuis la page d'accueil COVID-19 de Santé publique France <https://www.santepubliquefrance.fr/dossiers/coronavirus-covid-19/coronavirus-chiffres-cles-et-evolution-de-la-covid-19-en-france-et-dans-le-monde>

³ Ces tests sont parmi les plus fiables mais comportent une proportion importante de faux négatifs (30%), en particulier lorsque le prélèvement a été mal fait ou que le virus n'était pas encore ou plus assez développé. En France, la recommandation actuelle pour les cas contacts est d'effectuer un test 7 jours après le dernier contact. Aux États-Unis, le *Center for Disease Control* insiste sur le fait qu'un test négatif ne présage pas du fait que le virus ne se développera pas dans les jours qui suivent et conseille de maintenir la quarantaine en cas de suspicion de contamination (<https://www.cdc.gov/coronavirus/2019-ncov/hcp/testing-overview.html>). Comme soulevé dans le *New York Times*, un débat existe à l'heure actuelle à l'international : <https://www.nytimes.com/2020/08/29/health/coronavirus-testing.html>, et développé ici : <https://www.industrie-techno.com/article/covid-19-trois-resultats-scientifiques-cles-pour-repenser-la-doctrine-de-depistage-par-rt-pcr.61429>) sur le fait qu'il faudrait plutôt fonder le diagnostic sur la rapidité des cycles de reproduction du virus dans le sang plus que sur ce chiffre lui-même pour arrêter la pandémie.

⁴ Notons que Santé publique France annonce sur internet une collecte représentative dont deux premières vagues auraient eu lieu en mars et avril 2020. Mais aucun résultat n'est annoncé, ni aucune suite. <https://www.santepubliquefrance.fr/etudes-et-enquetes/covid-19-une-etude-pour-connaître-la-part-de-la-population-infectée-par-le-coronavirus-en-france>.

Figure 1. Les taux de dépistage par groupe d'âge en France depuis la première semaine de juin (semaine 23)



Source : Santé publique France, point épidémiologique du 8 octobre 2020, p.7

Ainsi, les indicateurs issus des tests correspondent encore aujourd'hui à des groupes particuliers, non représentatifs de la population. Certaines sous-populations sont plus susceptibles d'être testées que d'autres : les personnes à risque et notamment les personnes de plus de 75 ans (figure 1), le personnel médical et hospitalier, les personnes qui voyagent à l'étranger (certains pays imposent le test). Ce périmètre de la population testée chaque semaine se transforme au cours du temps, selon les règles ou les préconisations des services sanitaires et des pouvoirs publics.

Les indicateurs produits à partir des tests dépendent également de l'accès aux tests qui a varié depuis le début de l'épidémie. Parce que les tests n'étaient pas disponibles en grand nombre, les pouvoirs publics en ont d'abord limité l'accès. Par exemple, en mars dernier, les personnes vivant en EHPAD n'étaient pas systématiquement testées, même avec des symptômes⁵. En revanche, les personnes hospitalisées l'étaient systématiquement, y compris sans symptômes. Les tests ne sont réalisables pour tous sans ordonnance ni avance de frais que depuis le 25 juillet 2020⁶. Au cours des dernières semaines, la proportion de la population qui effectue un test RT-PCR est au plus haut depuis le début de l'épidémie : ce qui représente environ 1,5% des français chaque semaine⁷.

Les tests effectués, et par extension les indicateurs qui en découlent, ne concernent donc pas l'ensemble de la population, ils n'en sont pas représentatifs et ils ne peuvent pas être généralisés. Ils concernent les personnes qui ont effectué un test RT-PCR. Ils dépendent par conséquent de la disponibilité de ces tests, de leur accessibilité et des mesures en place qui incitent ou découragent de les effectuer. Ces trois aspects varient selon le lieu (en particulier le pays), et au cours du temps. Les résultats des tests sont donc difficilement comparables.

⁵ Ministère des solidarités et de la santé, Guide méthodologique Phase épidémique Covid-19 (version du 16 mars 2020), p.7. <https://solidarites-sante.gouv.fr/IMG/pdf/guide-covid-19-phase-epidemique-v15-16032020.pdf>.

⁶ Voir par exemple <https://www.ameli.fr/bouches-du-rhone/assure/covid-19/les-tests-de-depistage-du-covid-19/ce-quel-faut-savoir-sur-les-tests-de-depistage-du-covid-19>. L'infographie proposée par Marius Gilbert au sujet de la Belgique est tout à fait éclairante à ce sujet : <https://twitter.com/mariusgilbert/status/1298517274282860544/photo/1>.

⁷ Il est compris entre 1,2% et 1,8% de la fin août à la mi-octobre (semaines 35-41). Santé publique France, Covid-19. Point hebdomadaire du 15 octobre 2020.

3 - Que représentent et ne représentent pas le nombre et la proportion de tests positifs, le nombre de nouveaux cas et l'incidence ?

Les quatre indicateurs fondés sur les tests virologiques hebdomadaires sont le nombre de tests, le nombre de nouveaux cas, l'incidence et la proportion de tests positifs.

Le **nombre hebdomadaire de tests positifs** est le nombre de tests RT-PCR ayant détectés une charge virale à la COVID-19 une semaine donnée. Ce nombre ne correspond ni au nombre de porteurs du virus, ni aux personnes malades (c'est-à-dire avec des symptômes). En effet, certaines personnes ont pu être testées plusieurs fois dans la même semaine. Par exemple, en semaine 38 (14-20 septembre), 1 300 711 tests par RT-PCR ont été réalisés en France mais seulement 1 074 309 personnes ont été testées, il y a donc plus de 200 000 individus (environ 20% des personnes testées) qui ont été testés deux fois la même semaine⁸. De plus, beaucoup de personnes malades ou porteuses du virus n'ont pas été testées cette semaine-là (elles n'ont pas été testées du tout ou ont été testées une semaine précédente).

Le **nombre de nouveaux cas** (ou nombre de personnes testées positives) correspond, sur une période d'une semaine, au nombre de personnes ayant effectué un test RT-PCR positif. En semaine 38, 66 671 personnes sur les 1 074 309 testées ont été testées positives. Elles sont considérées comme des nouveaux cas. Notons que sont incluses dans ce résultat des personnes qui avaient déjà été testées la semaine précédente. Par ailleurs, comme une grande partie des personnes touchées par la COVID-19 n'ont pas de symptômes et ne sont donc pas amenées à être testées, cet indicateur ne représente pas le nombre de cas réels dans l'ensemble de la population. Il n'indique que ce que l'on sait du développement de la maladie à partir des personnes testées. Il ne correspond pas non plus au nombre de malades, puisqu'une grande partie des personnes testées ne présentent aucun symptôme (60% en semaine 38).

L'**incidence** (ou taux d'incidence) de l'épidémie est le rapport du nombre de personnes positives au nombre d'habitants du territoire concerné, et est exprimée en nombre de cas pour 100 000 habitants. Même si le mot 'test' n'apparaît pas dans son nom, cet indicateur est lui aussi estimé sur la base des tests effectués. Les résultats donnés par Santé publique France concernent les personnes testées positives, que ce soit le premier test ou non. Il ne s'agit véritablement de ce qu'on appelle normalement l'« incidence » d'une maladie.

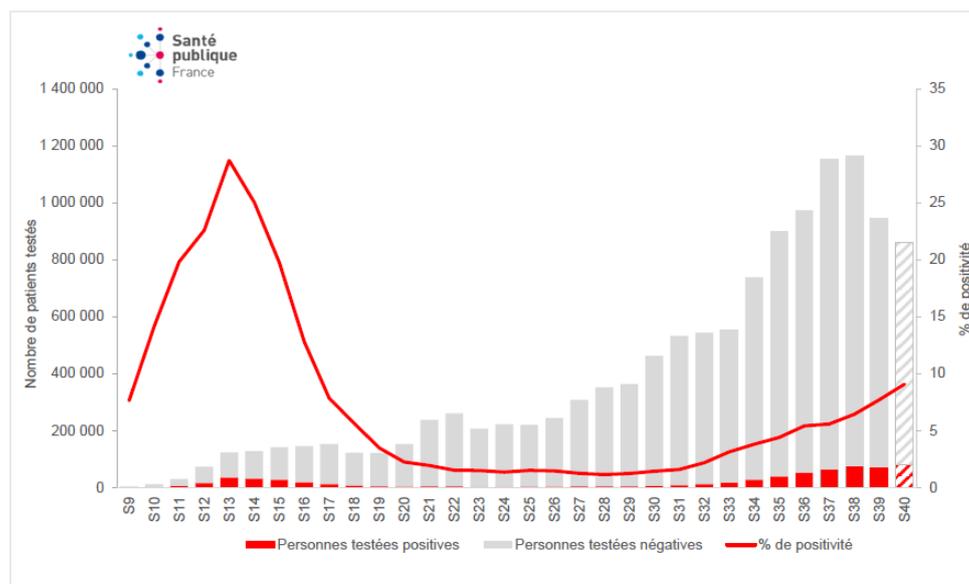
La **proportion de tests positifs** (ou taux de positivité) correspond au rapport entre le nombre hebdomadaire de tests positifs et le nombre total de tests effectués sur la semaine. Cette proportion ne correspond pas à la proportion de la population testée qui est positive, puisqu'il s'agit de tests et non de personnes. Elle correspond encore moins à la proportion de la population qui est positive, puisque les personnes testées ne sont pas représentatives de la population générale. La proportion de tests positifs est encore plus sensible que les autres indicateurs aux variations de la population testée. Plus les stratégies de dépistage touchent un spectre large de la population, plus faible est cette proportion. Ainsi, la croissance ou la décroissance de la proportion de tests positifs n'est pas forcément signe d'une croissance ou d'une décroissance du nombre de cas. En guise d'illustration, fin septembre (semaine 39), le taux de positivité des tests était de 7,6%, en augmentation par rapport au taux de la semaine précédente (6,4% en semaine 38) mais seulement 1 115 946 tests avaient été réalisés (contre 1 300 711 tests en semaine 38)⁹. Cela correspond à 70 458 personnes positives en semaine 39 pour 73 965 en semaine 38, donc à une baisse. La proportion croissante de tests positifs

⁸ Tous les chiffres de cette section concernant la semaine 38 sont issus de Santé publique France, Covid-19. Point hebdomadaire du 24 septembre 2020.

⁹ Santé publique France, Covid-19. Point hebdomadaire du 1er octobre 2020.

(+1,2 point), visible sur la figure 2, correspond donc à un nombre décroissant de nouvelles personnes infectées (-5%). Cet indicateur indique une augmentation alors que le nombre de nouveaux malades connus diminue. La figure 2 illustre également très bien le fait que plus les tests sont sélectifs et centrés sur les malades (comme en mars-avril dernier, semaines 10 à 17), plus le taux de positivité est élevé.

Figure 2. Nombre de personnes testées, nombre de personnes testées positives pour le SARS-CoV-2 et taux de positivité dans les laboratoires, par semaine, France, (source S9-S19 : 3 Labo et laboratoires hospitaliers ; depuis S20 : SI-DEP, données au 7 octobre 2020)



S40 : Données provisoires , en cours de consolidation

Source : Santé publique France, point hebdomadaire du 8 octobre 2020

Ces quatre indicateurs –nombre de tests positifs, nombre de nouveaux cas, incidence et positivité des tests– donnent une information sur la diffusion de l'épidémie sur le territoire, qui est particulièrement utile en début d'épidémie. Quand Santé publique France rend publics ces indicateurs qui montrent ce que l'on sait de l'épidémie, c'est en ayant en tête que les populations testées diffèrent selon la période et le lieu et que ces tendances ne peuvent être interprétées simplement. En revanche, quand les médias et la sphère politique parlent de taux de positivité croissants, c'est parfois sur un ton alarmiste qui ne tient pas compte du fait que les cas de COVID-19 n'ont pas été mesurés de la même façon au cours du temps, ce qui prêche à confusion.

Les comparaisons internationales souffrent de ce même problème : les résultats sont liés aux politiques de dépistage nationales, aux conditions de prise en charge du dépistage et de la maladie sur place, et à la construction des indicateurs de suivi du pays. De nombreux sites proposent des illustrations graphiques des tendances nationales et des comparaisons internationales. Sur le site Statista, se trouve une carte comparant les taux de positivité à travers l'Europe¹⁰. D'autres sites, proposent de comparer les pays et les continents entre eux, comme celui de l'Organisation mondiale pour la santé (<https://covid19.who.int/>) ou www.worldometers.info. La plupart des infographies ne donne aucune information claire sur le fait que dans certains pays, les tests sont payants et

¹⁰ <https://fr.statista.com/infographie/21970/covid-19-coronavirus-part-des-tests-de-depistage-positifs-en-europe/>

inabondables pour 90% de la population, alors que dans d'autres, ils sont pris en charge par l'Etat, voire imposés à certaines catégories de population¹¹. Ces comparaisons sont par conséquent trompeuses, voire dangereuses.

4 - Pourquoi des données hebdomadaires et pas journalières, pourquoi des données départementales et pas plus précises ?

Au moment du prélèvement RT-PCR, l'hôpital ou le laboratoire documente un certain nombre d'informations : la date du prélèvement, son lieu, et le lieu de résidence du patient. Il est ainsi en théorie possible d'attribuer à chaque test un code postal et une date. Néanmoins, il est problématique de considérer des tendances journalières, ou de descendre le niveau d'analyse à l'échelle de la commune... pourquoi ?

Une question d'échelle

En ce qui concerne l'échelle spatiale, dans les données, lorsque le code postal de la personne testée est disponible, les données sont attribuées par Santé publique France au département et à la région de résidence. Lorsqu'il ne l'est pas, les données ne sont prises en compte que dans les analyses nationales¹². Santé publique France a ainsi mis en place une information régulière par département, qui permet de déterminer le niveau d'alerte du département. Une cartographie des indicateurs est proposée à cette échelle sur leur site web et dans les bulletins hebdomadaires. Le choix du département est un compromis entre les deux tendances opposées qui est de fournir l'information la plus précise possible et d'utiliser un niveau qui soit le plus pertinent par rapport à l'organisation du système de santé publique (réseau hospitalier, etc.). Néanmoins, on peut de nouveau se demander si la comparaison est pertinente car les mesures et recommandations du gouvernement varient selon le niveau d'alerte du département et parfois au sein des départements. Cela affecte le comportement de la population par rapport au dépistage et donc les résultats de ces indicateurs. Le taux de dépistage varie énormément en fonction du département : en semaine 40, il est plus de six fois plus élevé dans le département qui teste le plus que dans celui qui teste le moins (figure 3). Cet écart est beaucoup trop important pour qu'une comparaison ait du sens.

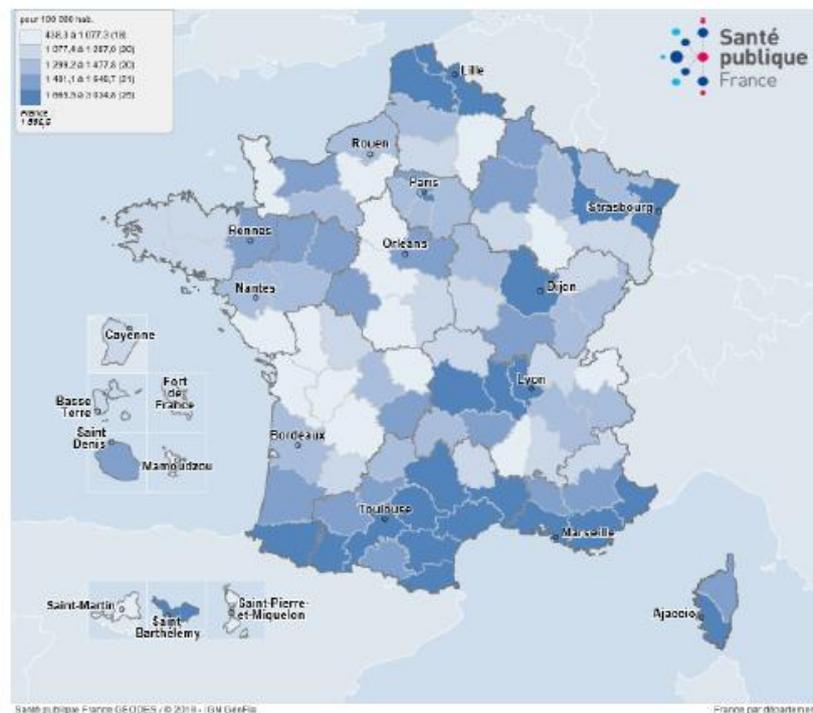
Les données produites à des niveaux inférieurs à celui du département (commune, agglomération) présentent des problèmes supplémentaires. La polémique récente au sujet des mesures prises dans la métropole d'Aix-Marseille et les chiffres extrêmes annoncés par les autorités locales pour le 5^e Arrondissement de Marseille ont fait apparaître un problème d'identification des lieux dans les données. Ainsi, si officiellement c'est le lieu de résidence qui est analysé, il semble qu'une pratique courante dans les laboratoires d'analyse soit de simplement reporter le code postal du laboratoire dans le formulaire de saisie, à l'emplacement prévu pour le lieu de résidence... A l'échelle départementale les erreurs s'annulent peut-être, mais à l'échelle locale, les plus grosses structures, comme l'institut hospitalo-universitaire de la Timone (IHU), situé dans le 5^e arrondissement de Marseille, déséquilibrent fortement les résultats et laissent faussement penser que ce quartier serait particulièrement frappé par l'épidémie¹³.

¹¹ Le lien entre nombre de tests effectués et détermination de cas positifs avait déjà été bien analysé en mars dernier. Voir par exemple https://www.lemonde.fr/les-decodeurs/article/2020/03/20/coronavirus-la-france-pratique-t-elle-assez-de-tests_6033865_4355770.html.

¹² Santé publique France, point épidémiologique du 8 octobre 2020. Voir aussi : <https://www.data.gouv.fr/fr/reuses/statistiques-de-tests-de-depistage-du-coronavirus-covid-19-en-france-par-region-et-par-departement/>.

¹³ <https://www.laprovence.com/article/papier/6132411/chiffres-du-covid-19-vous-avez-dit-transparence.html>.

Figure 3. Taux de dépistage pour le SARS-CoV-2 par département (/100 000 habitants), France, du 28 septembre au 04 octobre 2020 (source : SI-DEP, données au 07 octobre 2020)



Source : Santé publique France, point épidémiologique du 8 octobre 2020

Rythme hebdomadaire et délais de publication

Santé publique France met à jour ses données et sa page web de manière journalière. Cependant, les laboratoires ne travaillent pas tous les jours de la semaine, et les informations ne remontent pas de manière régulière depuis tous les laboratoires. Au cours du mois de septembre 2020, du fait de la recrudescence du nombre de tests effectués, les délais d'obtention des résultats – et *a fortiori* de remontée à Santé publique France - vont souvent bien au-delà des 48h officielles¹⁴, c'est encore vrai fin septembre dans certaines régions¹⁵. Ces délais font que les indicateurs concernant les derniers jours sont moins fiables que ceux portant sur les semaines précédentes, pour lesquelles l'ensemble des données a été rassemblé. Ceci explique pourquoi, dans les résultats publiés, les dernières estimations en date sont toujours provisoires. On parle de résultats consolidés pour ceux qui tiennent compte de l'ensemble des sources. Les résultats provisoires, non consolidés, sont voués à changer au fur et à mesure que de nouveaux résultats arrivent. Dans les graphiques, les tendances récentes sont souvent exprimées en pointillés ou avec des hachures plutôt qu'en traits pleins et en aplats de couleur (voir la figure 2 pour la semaine 40 par exemple). Les remontées sont également tributaires du fait que certaines structures ne font pas remonter l'information de manière régulière, mais débloquent certains jours les données concernant les jours précédents. Par ailleurs, la plupart des laboratoires privés étant fermés en fin de semaine, les séries journalières connaissent des creux non liés à l'épidémie elle-même, mais aux week-ends¹⁶.

¹⁴ https://www.lemonde.fr/planete/article/2020/09/05/le-gros-embouteillage-des-tests-du-covid_6051063_3244.html

¹⁵ https://www.lemonde.fr/economie/article/2020/10/01/tests-rt-pcr-les-tensions-d-approvisionnement-demeurent_6054301_3234.html

¹⁶ Une bonne analyse de ces problèmes a été publiée par *Libération* dès juillet 2020 https://www.liberation.fr/checknews/2020/07/09/covid-19-quelle-est-l-evolution-du-nombre-de-contaminations-en-france_1793684.

Conclusion

Beaucoup de mésinterprétations ou de lectures trop rapides amalgament des résultats chiffrés sans nuancer leur portée. Tous les utilisateurs n'ont pas les clés d'interprétation nécessaires pour se pencher sur ces nombres et ces proportions dont les médias et les pouvoirs publics se nourrissent. Du côté des médias, si de bonnes analyses existent depuis le début de l'épidémie, d'autres en revanche manquent cruellement de recul et de rigueur par rapport aux chiffres, prenant les statistiques publiées et leurs tendances au pied de la lettre, sans voir l'ensemble des données et sans relativiser la portée des résultats. Du côté du pouvoir public, ce sont ces mêmes résultats et leurs hypothétiques projections dans le futur qui servent de justification pour la mise en place de mesures dont les effets –prévus et imprévus– peuvent affecter fortement les populations. Une fois de plus, dans notre monde gouverné par les chiffres, il est de la responsabilité de chacun, en particulier des décideurs politiques, de bien comprendre comment les chiffres sont produits et ce qu'ils signifient avant de les utiliser pour justifier des décisions dont le coût social et économique est catastrophique.

Citation

Golaz Valérie, Gastineau Bénédicte, Flahaux Marie-Laurence, Dos Santos Stéphanie, Quatre questions-clés sur les indicateurs de suivi de l'épidémie de COVID-19 en France. Laboratoire Population-Environnement-Développement, Marseille, Octobre 2020, 8p.

Contacts

valerie.golaz@ined.fr

benedicte.gastineau@ird.fr

