

Annexe 7

Juin 2022

Évaluation de la stratégie nationale
de santé 2018-2022

Étude de cas sur l'utilisation des données dans la gestion de la crise sanitaire Covid-19

Table des matières

INTRODUCTION ET APPROCHE MÉTHODOLOGIQUE	3
A. RAPPEL DU CONTEXTE DE L'ÉTUDE ET DES ENJEUX ASSOCIÉS	3
B. UNE ÉTUDE DE CAS MENÉE SUR L'UTILISATION DES DONNÉES DANS LA GESTION DE LA CRISE SANITAIRE, AFIN D'IDENTIFIER DES PISTES DE RÉFLEXION POUR L'AVENIR	4
C. CADRE D'ANALYSE.....	5
I. UN CONSENSUS FORT AUTOUR DE L'UTILISATION DES DONNÉES POUR LE PILOTAGE DE LA GESTION DE CRISE, ET UN EFFORT EXCEPTIONNEL POUR METTRE EN PLACE LES REMONTÉES ET L'EXPLOITATION DES DONNÉES AU FIL DU TEMPS	6
A. L'IRRUPTION DE LA CRISE A ENTRAÎNÉ UN BESOIN MAJEUR ET SOUDAIN D'UN NOMBRE EXCEPTIONNEL DE DONNÉES.....	6
B. UNE VOLONTÉ POLITIQUE FORTE DE METTRE EN PLACE UNE GESTION DE CRISE S'APPUYANT SUR LES DONNÉES.....	7
C. LA STRUCTURATION DES REMONTÉES DE DONNÉES EN LIEN AVEC LA CRISE S'EST FAITE PROGRESSIVEMENT À PARTIR DE L'ACTIVATION DU CENTRE DE CRISE SANITAIRE FIN JANVIER 2020, MARQUANT LE DÉBUT DU PILOTAGE DE LA CRISE SANITAIRE	8
1. Phase 1, fin janvier 2020 – mai 2020 : une structuration dans l'urgence des premières remontées de données pour permettre la compréhension de la situation et appuyer les décisions opérationnelles.....	9
2. Phase 2, mai 2020 – printemps 2021 : une stabilisation des remontées de données pour fiabiliser le suivi et piloter la réponse à la crise (restrictions, vaccination...)	12
3. Phase 3, printemps 2021 – septembre 2021 : une mobilisation des données dans un fonctionnement permettant la vie avec le Covid-19, et des réflexions sur la pérennisation des dispositifs.....	14
II. LA MOBILISATION SANS PRÉCÉDENT DES ACTEURS AUTOUR DE LA DATA DANS LE CONTEXTE DE CRISE A MIS EN LUMIÈRE DES DÉFIS À TOUTES LES ÉTAPES DE LA COLLECTE ET DE L'UTILISATION DES DONNÉES	15
A. DES DIFFICULTÉS DE FORMULATION ET DE PRIORISATION DES BESOINS EN DONNÉES, ET UN MANQUE DE COORDINATION ENTRE ACTEURS.....	16
B. LA MISE EN PLACE DE LA COLLECTE ET DES REMONTÉES DE DONNÉES AUX DIFFÉRENTES PHASES DE LA CRISE S'EST HEURTÉE À DES DIFFICULTÉS TRANSVERSES	17
C. LE TRAITEMENT, L'ANALYSE ET LES TRAVAUX DE VISUALISATION DES DONNÉES ONT ÉTÉ PORTÉS DE MANIÈRE TRÈS FRAGMENTÉE PAR UNE MULTIPLICITÉ DE PARTIES PRENANTES.....	19
D. UN EFFORT MAJEUR DE DIFFUSION DES DONNÉES PENDANT LA CRISE, MAIS DES INITIATIVES MULTIPLES ET PEU COORDONNÉES, N'AYANT PAS RÉPONDU À L'ENSEMBLE DES BESOINS	21
E. LES DONNÉES ONT ÉTÉ FORTEMENT MOBILISÉES POUR L'ANALYSE DE LA SITUATION,	

LA PRISE DE DÉCISION ET LA COMMUNICATION, MAIS INSUFFISAMMENT UTILISÉES POUR L'ANTICIPATION.....	23
F. UNE GOUVERNANCE ET UNE ORGANISATION DE LA GESTION DE CRISE SUR LA PARTIE DATA FRAGMENTÉES, QUI ONT NOURRI LES DIFFICULTÉS RENCONTRÉES TOUT AU LONG DU PROCESSUS D'IDENTIFICATION DES BESOINS, DE COLLECTE ET D'UTILISATION DES DONNÉES.....	25
III. PISTES DE RÉFLEXION.....	27
A. LA CRISE DU COVID-19 A ACTÉ L'IMPORTANCE STRATÉGIQUE DE COLLECTER ET D'UTILISER DES DONNÉES POUR LE PILOTAGE DU SYSTÈME SANITAIRE, ET A DÉMONTRÉ L'IMPORTANCE DE DÉFINIR « À FROID » LEURS MODALITÉS D'UTILISATION EN TEMPS DE CRISE	27
B. L'ÉTUDE A PERMIS DE METTRE EN LUMIÈRE TROIS AXES CLÉS POUR RENFORCER LE PILOTAGE DES CRISES SANITAIRES GRÂCE À L'UTILISATION DES DONNÉES, EN PRÉPARATION D'UNE CRISE FUTURE	29

Cette annexe a été réalisée avec l'appui des cabinets Accenture et McKinsey.

INTRODUCTION ET APPROCHE MÉTHODOLOGIQUE

A. RAPPEL DU CONTEXTE DE L'ÉTUDE ET DES ENJEUX ASSOCIÉS

- **L'utilisation et le partage des données de santé : un sujet transverse à de nombreux domaines d'action prioritaires de la stratégie nationale de santé (SNS), notamment au sein de son Axe 3**

La préparation et la réponse aux alertes et aux crises sanitaires constituent le septième élément de la politique de santé au sens de l'article L. 1411-1 du CSP définissant la responsabilité de l'État en matière de politique de santé.

Un des sept principes d'action posés par la SNS est : « Appuyer les décisions sur l'analyse critique des meilleures connaissances disponibles, notamment celles issues des données concernant la santé des personnes et leur environnement, des évaluations des actions menées, des expertises indépendantes, des études d'impact en santé et de la recherche ». L'exploitation des données de santé dans le cadre d'une crise sanitaire figure notamment dans l'objectif de **l'Axe 3** : adapter le système de veille et de vigilance aux « signaux faibles » ou aux risques émergents et renforcer la résilience du système de santé face aux situations sanitaires exceptionnelles.

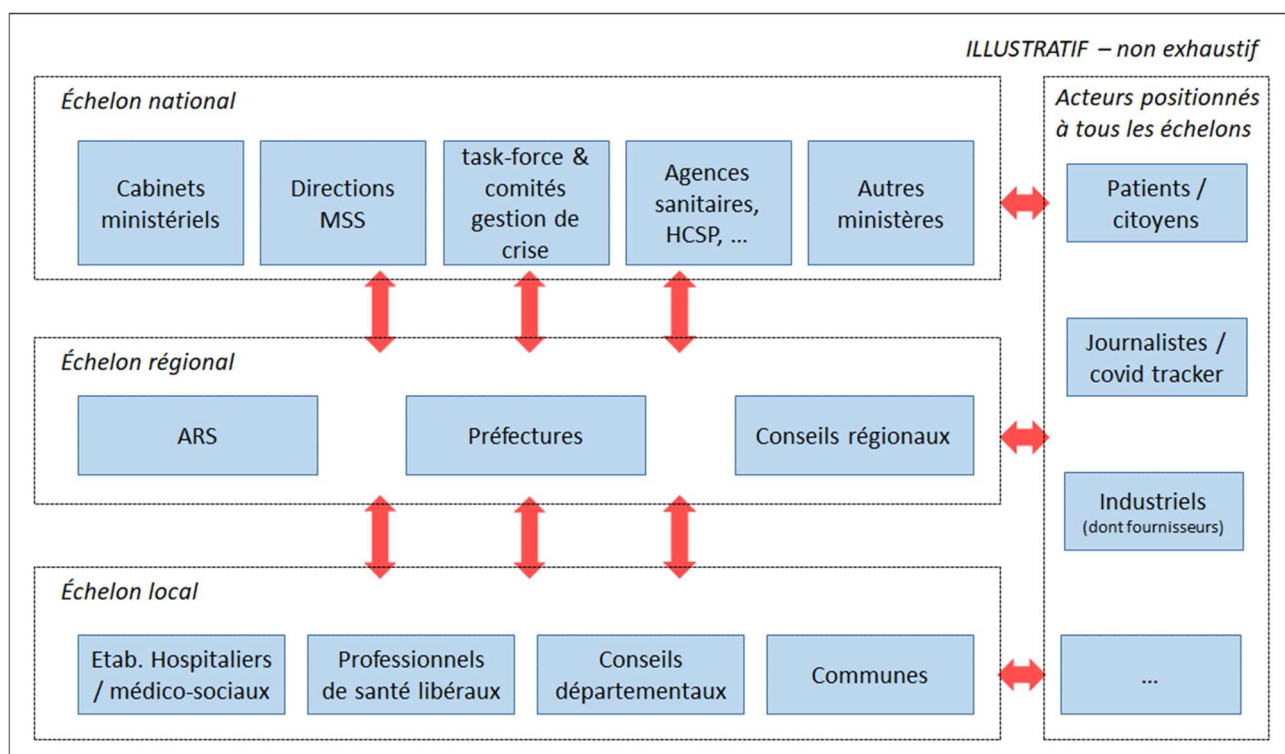
Ce rapport se centre sur **l'utilisation et le partage des données de santé pour la production de statistiques** et le **pilotage des politiques publiques**.

- **L'importance de disposer de données pour prendre des décisions pertinentes mise en lumière par la crise sanitaire**

Pendant la crise, de nombreux sujets ont supposé l'utilisation de données dans la prise de décision, notamment :

- **l'analyse de l'évolution sanitaire** (incidence, dépistages, nombre d'hospitalisations, de décès, etc.) ;
- **la définition et l'adaptation des mesures visant à limiter la propagation du virus** (restrictions, *contact tracing*...), avec une approche par population et par territoire ;
- **la gestion des capacités du système de santé et l'allocation des moyens**, en fonction des situations sanitaires locales ;
- **la préparation et la gestion de la campagne de vaccination**, en tenant compte des populations prioritaires et des vaccins disponibles (commandes, flux logistiques et modalités de distribution, centres, etc.).

Cela a entraîné de nombreux flux et partages d'information entre acteurs de l'écosystème de la santé en France aux échelons national, régional et local.



B. UNE ÉTUDE DE CAS MENÉE SUR L'UTILISATION DES DONNÉES DANS LA GESTION DE LA CRISE SANITAIRE, AFIN D'IDENTIFIER DES PISTES DE RÉFLEXION POUR L'AVENIR

Dans ce contexte, la DREES a proposé au comité de pilotage de l'évaluation de la Stratégie nationale de santé de conduire une étude de cas sur l'utilisation des données dans la gestion de crise sanitaire, avec l'objectif **d'identifier des pistes de réflexion pour l'avenir visant à améliorer le pilotage du système sanitaire (en temps de crise et hors crise)**.

Pour cela, l'étude a suivi une approche autour de quatre thématiques clés :

Processus - une analyse de l'ensemble du processus, depuis l'identification du besoin de données jusqu'à l'utilisation de ces données pour le pilotage de la crise

-Nature des données - la prise en compte de l'ensemble des données permettant de nourrir le pilotage (données de suivi de la pression épidémique, données de capacité du système de soins, ...)

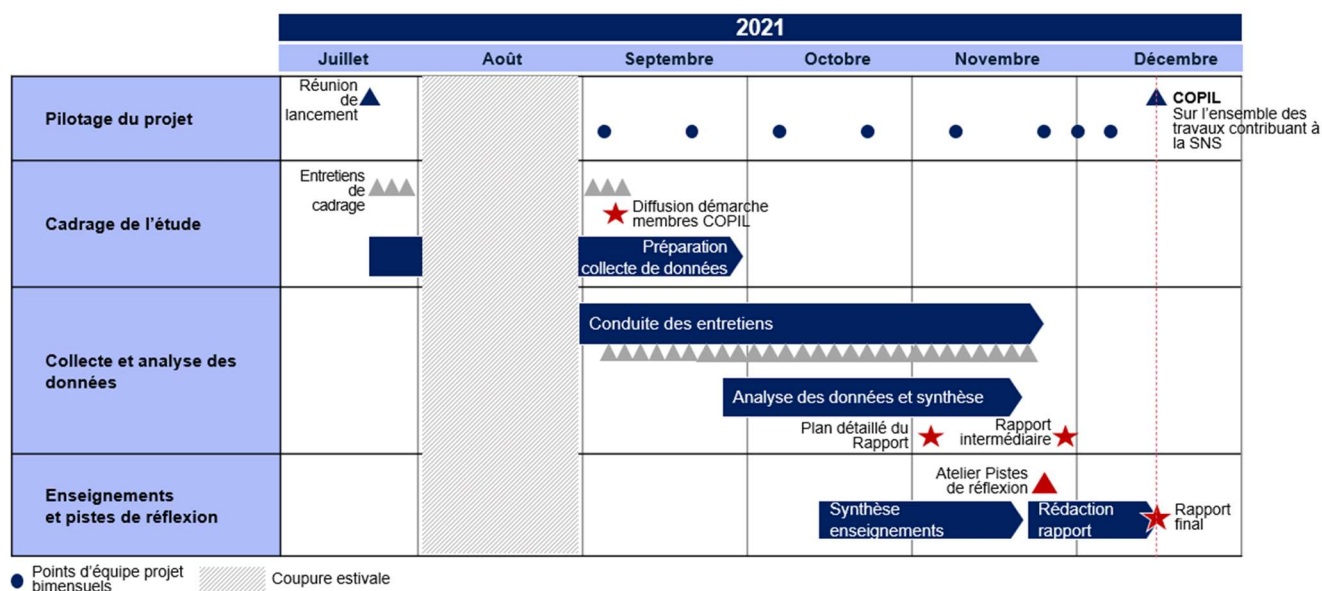
-Chronologie - une approche chronologique, pour prendre en compte l'évolution de la situation, depuis la mise en place du pilotage de la crise sanitaire par l'activation du Centre de crise sanitaire le 27 janvier 2020 jusqu'à fin septembre 2021

-Catégories d'acteurs - une approche permettant de couvrir les situations vécues par différentes catégories d'acteurs aux différents échelons du territoire (nationaux, régionaux, locaux).

C. CADRE D'ANALYSE

- Une étude inscrite dans les travaux d'évaluation de la SNS, conduite sur quatre mois

L'étude porte sur la période janvier 2020 – septembre 2021. Elle s'est déroulée en quatre mois selon le calendrier ci-dessous :



- Plus d'une centaine de personnes rencontrées au cours de 61 entretiens, toutes impliquées dans la gestion de la crise sanitaire à tous les niveaux

Afin d'étayer l'étude de cas, 61 entretiens ont été conduits avec plus d'une centaine de personnes de différentes entités impliquées dans la gestion de crise :

- **des acteurs publics nationaux** – des entités de pilotage (CCS, CIC...), les directions du MSS impliquées dans la gestion de crise (DGS, DGOS, DREES, DGCS, SGMAS et SCN...), d'autres acteurs clés (SpF, CNAM, ATIH...) ;
- **des acteurs publics régionaux** – des ARS (pilotage régional et *data scientists* ou statisticiens) et leurs partenaires (GRADEs...) ;
- **des acteurs d'établissements de santé et médico-sociaux** (CHU, Ehpad...) ;
- **d'autres acteurs mobilisés pendant la crise** (citoyens, *data scientists*, chercheurs...).

Dans le contexte de reprise épidémique et d'une charge importante pesant sur les ARS sollicitées pour effectuer les mises en relation, le nombre d'entretiens au niveau régional a été plus limité que ce que l'étude prévoyait initialement.

Acteurs publics nationaux – 34 entretiens réalisés 15 entretiens avec des entités de pilotage de la crise (CCS, CIC, taskforce vaccination, détachement DREES au CCS) 9 entretiens avec les principales directions du MSS impliquées (DGS, DGOS, DREES, DGCS, SCN) 10 entretiens avec d'autres acteurs nationaux clés (SpF, CNAM, Etalab, DNS, ANS, ATIH, HAS, CNS)	Acteurs publics régionaux – 15 entretiens réalisés entretiens réalisés avec 7 ARS 14 Normandie, Centre Val de Loire, Bourgogne Franche Comté, Ile de France, Grand Est, Guadeloupe, Pays de la Loire dont 16 responsables / membres de cellule de crise 10 data scientists / statisticiens 1 entretien avec un GRADeS
Établissements de santé et médico-sociaux et leurs représentants – 7 entretiens réalisés 3 entretiens réalisés avec des établissements de santé 3 entretiens avec des fédérations représentant des établissements de santé 1 Entretien réalisé avec un groupe d'EHPAD	Autres acteurs mobilisés – 5 entretiens réalisés 2 entretiens avec des citoyens data scientists (publication de données en opendata) 3 entretiens avec des chercheurs / acteurs de la recherche et des données (CNRS, HDH, SSA)

CCS : Centre de crise sanitaire au niveau national ; CIC : cellule interministérielle de crise ; SCN : Service à compétence nationale des SI mutualisés des ARS ; SpF : Santé publique France ; CNAM : Caisse nationale de l'assurance maladie ; DNS : Délégation ministérielle au numérique en santé ; ANS : Agence du numérique en santé ; ATIH : Agence technique de l'information hospitalière ; HAS : Haute Autorité de santé ; CNS : Conférence nationale de santé ; GRADeS : Groupement régional d'appui au développement de la e-santé ; CNRS : Centre national de la recherche scientifique ; HDH : Health Data Hub ; SSA : Service de santé des armées.

I. UN CONSENSUS FORT AUTOUR DE L'UTILISATION DES DONNÉES POUR LE PILOTAGE DE LA GESTION DE CRISE, ET UN EFFORT EXCEPTIONNEL POUR METTRE EN PLACE LES REMONTÉES ET L'EXPLOITATION DES DONNÉES AU FIL DU TEMPS

Cette partie a pour objectif de montrer de façon chronologique les actions mises en place par les acteurs afin de collecter et d'utiliser les données pour le pilotage de la gestion de crise.

A. L'IRRUPATION DE LA CRISE A ENTRAÎNÉ UN BESOIN MAJEUR ET SOUDAIN D'UN NOMBRE EXCEPTIONNEL DE DONNÉES

Au démarrage de la crise, le Covid-19 correspondait à un phénomène nouveau : ses caractéristiques et ses effets étaient peu connus. Plus tard, des variants évoluant rapidement ont maintenu un contexte incertain.

La **situation s'est dégradée très rapidement**, notamment du fait de la vitesse de propagation rapide du Covid-19 (un agent volatile et contagieux, un taux d'incubation court...). En moins d'un mois, le système de soins était saturé (par exemple, une évolution de 400 personnes hospitalisées avec le Covid-19 à 32 000 a été enregistrée entre mi-mars et mi-avril 2020¹).

Cette situation a demandé une **réponse sanitaire d'une ampleur sans précédent**. Des moyens exceptionnels ont été mis en place par les acteurs pour répondre à la crise (par exemple, de nombreuses évacuations sanitaires par train ou par avion ont été effectuées, plus de 113 millions de tests ont été réalisés entre mars 2020 et août 2021, une campagne de vaccination massive a été engagée).

¹ Point épidémiologique Covid-19 Santé publique France.

Tout cela a mis en évidence un **besoin majeur de données** pour comprendre la situation, suivre la progression de l'épidémie et construire des réponses d'ampleur tout au long de la crise.

B. UNE VOLONTÉ POLITIQUE FORTE DE METTRE EN PLACE UNE GESTION DE CRISE S'APPUYANT SUR LES DONNÉES

• Un pilotage et une communication de crise s'appuyant sur les données

Dès le démarrage de la crise, **le président de la République et le gouvernement** ont donné **une impulsion forte : s'appuyer sur la science et les données pour assurer le pilotage de la crise**, ce qui s'est traduit notamment par plusieurs engagements emblématiques :

- **un pilotage fondé sur l'expertise scientifique, reposant notamment sur les données**, via la constitution de cellules d'experts en support des cellules de crise, comme le Conseil scientifique mis en place le 11 mars 2020 (organisation consultative indépendante composée d'une quinzaine de membres de différentes disciplines scientifiques et médicales complétée par un Conseil d'orientation de la stratégie vaccinale placé auprès du ministre chargé de la Santé mis en place en décembre 2020) ;
 - **une communication forte et fréquente auprès du grand public s'appuyant sur les chiffres de l'épidémie**, avec la tenue régulière de conférences de presse nationales, étayées par des analyses graphiques, des chiffres détaillés et l'expertise d'intervenants qualifiés ;
 - **une politique d'Open Data très marquée**, avec notamment la publication de nombreuses données dès le début de l'épidémie.
-
- **Les acteurs interrogés soulignent l'importance d'utiliser ces données pour quatre fonctions du pilotage : l'analyse de la situation, l'anticipation, la prise de décision et la communication**

L'importance d'utiliser les données pour quatre fonctions du pilotage a été soulignée par les acteurs interrogés. Il leur apparaît utile d'avoir des données pour : **analyser la situation** (détecter les signaux, recenser les moyens disponibles...) ; **anticiper** (réaliser des projections pour mettre en place des actions visant à limiter la propagation de l'épidémie, préparer les acteurs du système de soins...) ; **prendre des décisions** (affiner les décisions et les légitimer) ; et **communiquer** (informer les acteurs impliqués dans la gestion de crise et la population, pour rendre compte de l'action des pouvoirs publics et changer les comportements).

• De nombreuses données ont été identifiées pour répondre à ces quatre fonctions, avec des niveaux de disponibilité et de qualité variables au démarrage de la crise

Dès le début de la crise, **des données sanitaires essentielles au pilotage ont été identifiées par les acteurs interrogés dans le champ sanitaire :**

- données de suivi de la pression épidémique ;
- données de suivi de la capacité du système de soins ;
- données de suivi de la vaccination (à partir du démarrage de la campagne de vaccination) ;
- données de connaissance de la maladie et de ses impacts.

Suivi de la pression épidémique	Identification et suivi des contaminations et chaînes de transmission (Cas Covid, Cas Contact...)
	Suivi des consultations médicales pour suspicion Covid (médecine de ville, urgences...)
	Suivi des hospitalisations (Cas Covid hospitalisés...)
	Suivi de la situation au sein des EHPAD (Cas Covid, Décès...)
	Suivi des décès (Décès pour cause covid...)
Suivi de la capacité du système de soin	Suivi du nombre de lits disponibles (lits totaux, lits de réanimations...)
	Suivi des stocks de divers matériels médicaux (respirateurs, EPI, médicaments...)
	Suivi des tests (matériels, lieux...)
	Suivi du planning des établissements de santé (opérations programmées, reprogrammables...)
	Suivi du personnel soignant disponible (soignants disponibles par spécialité...)
	Suivi des capacités hospitalières disponibles (hôpitaux publics, privés...)
Suivi de la vaccination	Suivi de l'avancée de la campagne de vaccination (couverture vaccinale, répartition des doses...)
	Recensement des lieux de vaccination (pharmacies, centres de vaccination...)
	Mesure de l'efficacité de la vaccination et de ses effets (signalement des effets secondaires...)
Connaissance de la maladie et de ses impacts	Connaissance du contexte socio-démographique sur le territoire (décès annuels, comorbidités...)
	Suivi des effets du Covid (symptômes, létalité...)
	Efficacité / Fiabilité des protocoles de prise en charge et des traitements (protocoles, traitements...)
	Connaissance des travaux de recherche en France et à l'international (protocoles, traitements...)

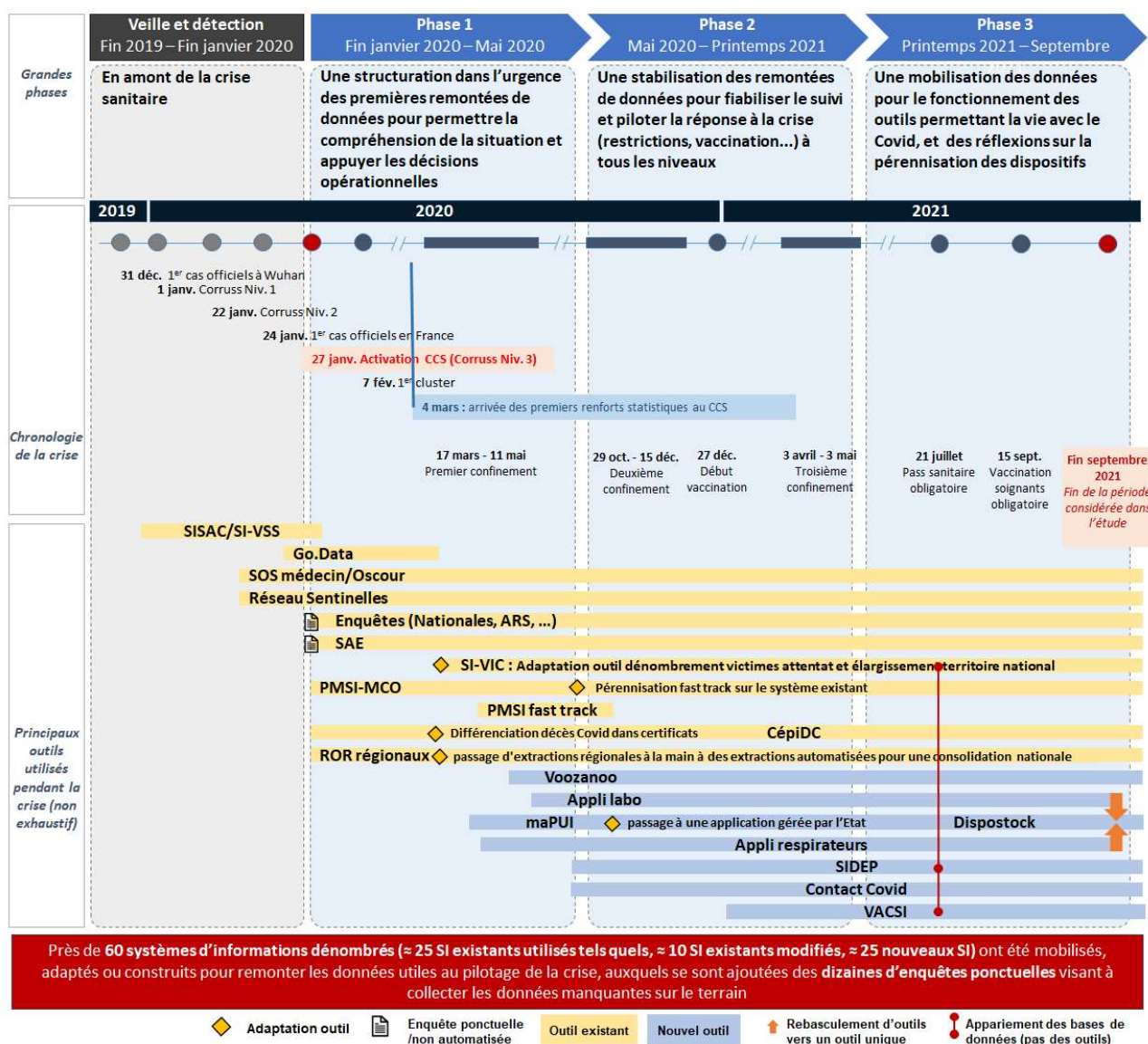
EPI = équipement de protection individuel.

Afin d'assurer le pilotage de la crise, ces données sont apparues nécessaires à **des fréquences différentes**. Par exemple, les données de suivi de la pression épidémique demandaient une fréquence de remontée au plus près du temps réel, en particulier dans les phases de pics épidémiques, alors que les données décrivant la maladie (tableau clinique, par exemple) nécessitaient notamment une connaissance globale du séjour des patients hospitalisés, voire de l'ensemble de leur parcours de santé, pour éclairer la pratique clinique des professionnels de santé et alimenter la recherche. Par ailleurs, **les niveaux de maturité de collecte de ces données différaient avant la crise, avec, dans l'ensemble, un niveau de maturité ou de réactivité faible** (secteur médico-social et libéral, causes médicales de décès, etc.) témoignant d'un manque de préparation du système en matière de collecte et d'utilisation des données dans le cadre d'une crise sanitaire.

C. LA STRUCTURATION DES REMONTÉES DE DONNÉES EN LIEN AVEC LA CRISE S'EST FAITE PROGRESSIVEMENT À PARTIR DE L'ACTIVATION DU CENTRE DE CRISE SANITAIRE FIN JANVIER 2020, MARQUANT LE DÉBUT DU PILOTAGE DE LA CRISE SANITAIRE

Depuis l'activation du Centre de crise sanitaire fin janvier 2020 et jusqu'à fin septembre 2021, la structuration des remontées de données s'est progressivement mise en place. On peut distinguer trois phases principales : **Phase 1**, de janvier à mai 2020, **Phase 2**, de mai 2020 au printemps 2021, et **Phase 3**, du printemps 2021 à septembre 2021 (fin de la période d'étude).

Au cours de ces phases, près de **60 systèmes d'information** ont été mobilisés, adaptés ou construits pour remonter les données utiles au pilotage de la crise, auxquels se sont ajoutées des **dizaines d'enquêtes ponctuelles** visant à collecter des données manquantes sur le terrain. Sur le graphique ci-dessous, les principaux outils sont représentés.



1. Phase 1, fin janvier 2020 – mai 2020 : une structuration dans l'urgence des premières remontées de données pour permettre la compréhension de la situation et appuyer les décisions opérationnelles

- Une mobilisation forte et rapide des acteurs sur la question des données, mais une dynamique peu coordonnée et des actions en ordre dispersé

Dès le démarrage de la crise, une gouvernance s'est mise en place aux niveaux national et local via la constitution de cellules de crise : Centre de crise sanitaire au niveau national (CCS) et cellule régionale d'appui et de pilotage sanitaire (CRAPS) au sein de chaque agence régionale de santé (ARS).

Afin de répondre au besoin majeur de données pour le pilotage, de nombreux acteurs avec des compétences en matière de données se sont mobilisés :

- **au sein de l'administration.** Au niveau national (par exemple, détachement de personnels de la DREES vers le CCS dès février 2020, mobilisation des agents de Santé publique France [SpF] et de la CNAM), et au niveau régional (par exemple, mobilisation des services statistiques des ARS pour appuyer les CRAPS, mobilisation identique dans les établissements de santé, etc.). Un support complémentaire **d'autres profils de l'administration** a été apporté (par exemple, un agent du Service de santé des armées est venu piloter un pôle dédié à la recherche et au partage des données et connaissances scientifiques sur le Covid-19)
- **en dehors de l'administration.** De nombreux acteurs externes se sont mobilisés pour apporter leur savoir-faire en matière de données. Par exemple, on peut citer la mise à disposition par des **collectifs civils de data scientists volontaires** d'outils de visualisation / représentations graphiques à partir des données mises en ligne par SpF, et la mobilisation de **nombreux chercheurs** dans des consortiums d'échanges et de collaboration autour de travaux liés au Covid-19 (par exemple, la création par le CNRS de la plateforme ModCov19 pour recenser les travaux de chercheurs).

Cependant, **ces mobilisations sont restées peu coordonnées**, avec un pilotage transverse insuffisant et une difficulté globale à cerner les rôles et compétences de chacun. Les initiatives sont restées souvent en silo, et les acteurs à différents échelons ont le sentiment d'agir « dans leur coin » faute d'une coordination nationale établie. Au-delà, de nombreuses initiatives volontaires sont peu entendues, et elles sont sous-exploitées selon certains acteurs (par exemple, des ARS qui ont eu l'impression d'être associées tardivement à la gestion de crise et aux enjeux de collectes).

- Une capitalisation immédiate sur les outils existants pour établir un suivi de la situation, mais des dispositifs rapidement dépassés

Au début de la crise, les outils de veille existants sont rapidement mobilisés pour signaler les premiers cas ou clusters sur le territoire national : **Sisac** (SI national pour signaler tout événement sanitaire des ARS à l'administration centrale) ; **SI-VSS** (SI régional pour signaler tout événement sanitaire au sein des ARS) ; **Go.Data** (SI développé par l'OMS pour assurer un suivi épidémiologique – le nombre de cas positifs – à petite échelle) ; **réseau Sentinelles** (outil de veille de SpF constitué d'un réseau de médecins généralistes signalant et remontant des données relatives aux infections respiratoires aiguës) ; **SurSaUD** (système de surveillance sanitaire des urgences et des décès, comprenant notamment des remontées de **SOS médecins** et des services d'urgences).

Cependant, ces outils sont très rapidement **dépassés face à l'ampleur de la crise** et à l'augmentation rapide des cas et des clusters sur l'ensemble du territoire national (ce sont des outils de signalisation et pas de dénombrement à grande échelle). Les équipes des cellules de crise sont souvent obligées d'effectuer un **travail manuel fastidieux** pour tenter de suivre les premières chaînes de contamination (sur Excel, sur papier...). Cela conduit rapidement à un **suivi partiel des cas et à une incapacité à suivre les chaînes de contamination** pendant quelques semaines, limitant le pilotage de la réponse à l'épidémie et la capacité à la maîtriser.

- Un besoin majeur et urgent de remontées de données structurées pour permettre un suivi des cas et de la situation hospitalière, dans un contexte d'évolution rapide de l'épidémie

Très rapidement, des fortes tensions hospitalières se font ressentir sur le territoire national, **avec près de 40 000 personnes hospitalisées** pour Covid-19 **sur le mois de mars 2020** dont près de 9 000 admises en soins critiques.

Dans ce contexte, **l'identification des données hospitalières et de la situation dans les Ehpad s'avère prioritaire** comme l'évoquent les acteurs interrogés (par exemple, les données d'occupation des lits, les données de matériels, les données de cas positifs ou de décès dans les Ehpad...).

À cette période, il aurait été également opportun, selon certains acteurs, de **prioriser la collecte de données scientifiques**, liées à la connaissance de la maladie, afin d'améliorer les dispositifs et les pratiques professionnelles de prise en charge.

- **Face à ces besoins, une difficulté à s'appuyer sur les outils existants et l'absence totale de données dans certains domaines**

Pour répondre à ces nouveaux besoins en données, les acteurs mobilisent les outils existants, par exemple la **Statistique annuelle des établissements de santé (SAE)**, pour connaître la capacité théorique des établissements de santé par service, ou le **Répertoire opérationnel des ressources (ROR)**, pour connaître le nombre de lits installés et le nombre de lits disponibles. Cependant, ces outils se révèlent rapidement inadaptés pour le pilotage d'une crise sanitaire, notamment en ce qui concerne la **maille géographique** (des outils permettant de récupérer des données à l'échelle locale ou régionale alors que les décideurs avaient besoin de données consolidées au niveau national en temps réel – par exemple, le ROR) et la **fréquence** (des outils récoltant des données de façon ponctuelle – collecte trimestrielle et/ou annuelle – alors que la crise impliquait un suivi quotidien, voire en temps réel – par exemple, la SAE ; ou des remontées décalées dans le temps, comme c'est le cas pour les causes médicales de décès ou pour le PMSI). Dans certains domaines, les données disponibles sont même anciennes, par exemple les données de situation au sein des Ehpad qui proviennent d'une enquête statistique de la DREES réalisée tous les quatre ans, les dernières données remontant à 2015 à ce moment-là. Les données statistiques structurelles apportent des premiers éléments de cadrage, mais elles ne répondent pas au besoin de pilotage de crise, car elles n'ont pas vocation à permettre un suivi conjoncturel, qui plus est dans le cadre d'une crise épidémique.

- **La mise en place dans l'urgence de remontées de données pour répondre aux besoins de pilotage – aux niveaux national et infranational**

Afin de répondre au manque d'outils utiles au pilotage de la crise, plusieurs chantiers sont conduits au niveau national :

- **des chantiers d'adaptation d'outils existants pour fiabiliser et compléter les remontées.** On retient notamment deux chantiers majeurs : **celui sur le ROR** conduit par la DGOS et la DREES en mars 2020 afin d'obtenir des données sur le nombre de lits disponibles au niveau national en temps réel ; et **celui sur SI-VIC**, outil conçu initialement pour suivre les victimes d'attentats (2015), conduit par l'ANS sous pilotage de la DGS en mars 2020, afin de suivre les hospitalisations liées au Covid-19, retours à domicile et décès, au plus près du temps réel ;
- **la construction de nombreux outils ad hoc pour répondre aux besoins non couverts.** On retient notamment **la construction de l'outil Voozanoo** par SpF en mars-avril 2020 sur une infrastructure informatique existante (portail de signalement des maladies du ministère des Solidarités et de la Santé) pour suivre les cas et les décès liés au Covid-19 au sein des Ehpad, et **divers chantiers de mise en place d'outils ad hoc** par les renforts de la DREES au **CCS**, en mars-avril 2020, pour répondre à des besoins urgents (application « Respirateurs » pour suivre les respirateurs disponibles ; application « Laboratoires » pour suivre les équipements pour les tests dans les laboratoires ; application « Épi-Stocks » pour suivre les équipements de protection individuels) ;

- **le lancement en parallèle de nombreuses enquêtes pour répondre à des demandes spécifiques.** De nombreuses enquêtes auprès des établissements sont également menées pour collecter des données ne faisant pas l'objet de remontées fiables via les outils mis en place (par exemple, les enquêtes de la DGOS auprès des établissements pour connaître les capacités en réanimation, en sus du ROR). Ces enquêtes sont souvent en doublon – notamment entre échelons national et régional – et sont réalisées à travers des canaux peu adaptés (Excel, WhatsApp, SMS, e-mails...).

Les acteurs régionaux et de terrain ont également structuré des remontées d'informations, pour leurs propres besoins de pilotage, et à un degré de coordination variable avec les acteurs de l'échelon national. Par exemple, l'ARS Île-de-France, disposant comme d'autres ARS de compétences mobilisables importantes (une cellule de données dédiée, de nombreux statisticiens mobilisés et un appui de leurs partenaires – GRADeS Sesan, observatoire régional de santé, cellule régionale de SpF...), a développé avant l'arrivée de Voozanoo un questionnaire auprès des Ehpad pour connaître le nombre de leurs cas et de leurs décès. Elle a également développé une plateforme, **Cas Covid Hébergement**, pour suivre le nombre de cas et de décès dans les hébergements sociaux.

- **Ces initiatives nationales et infranationales ont conduit à des processus complexes et à un foisonnement de données, et ont généré une pression forte sur les équipes de terrain**

Ces processus de collecte de données à tous les échelons s'avèrent complexes. **De nombreuses collectes de données restent non automatisées** et impliquent un fort investissement des équipes, notamment au niveau du terrain, pour effectuer le recueil et la saisie des données (par exemple, un remplissage à la main et en temps réel des données de SI-VIC, ROR...). **Les outils ont souvent été adaptés ou construits dans l'urgence, avec un manque de coordination**, conduisant notamment à des erreurs de collecte (par exemple, dans Voozanoo, le manque d'une doctrine de remplissage partagée a donné lieu à des incohérences dans les saisies de certains Ehpad, notamment sur le nombre de cas). **Ces difficultés dans les remontées ont donné par conséquent lieu à des retraitements et /ou une consolidation parfois complexes.**

Ces processus lourds ont également induit **une forte pression sur les équipes. Au sein des établissements**, elles ont été fortement mobilisées pour répondre aux nombreuses collectes de données, parfois redondantes. **Au niveau régional**, les ARS devaient assurer un rôle de pivot pour répondre aux nombreuses demandes de données au niveau national, aux demandes de partenaires régionaux (préfectures), tout en pourvoyant leurs propres besoins en données. Enfin, **au niveau national**, les équipes ont fourni des efforts conséquents pour retraiter, comparer et mettre en cohérence les données issues des différentes sources (par exemple entre SI-VIC et le ROR).

2. **Phase 2, mai 2020 – printemps 2021** : une stabilisation des remontées de données pour fiabiliser le suivi et piloter la réponse à la crise (restrictions, vaccination...)

- **L'aboutissement de chantiers nationaux majeurs permettant une automatisation des remontées de données et la construction d'indicateurs partagés en temps réel**

Au cours de la crise, le retour d'expériences et la possibilité de dégager un minimum de temps pour des investissements ont conduit à une **amélioration progressive** des circuits de collecte

et à la mise en place de remontées de données automatisées, facilitant la mise en place et le calcul des indicateurs clés pour le pilotage de la gestion de crise.

Trois chantiers majeurs illustrent cette amélioration progressive :

- **SI-DEP**, développé en avril 2020 par l'AP-HP et un consortium d'acteurs privés sous le pilotage de la Délégation ministérielle au Numérique en Santé (DNS), pour **remonter en temps réel les données sur l'ensemble des tests réalisés** sur la population, quel que soit le lieu de prélèvement ;
- **Contact Covid**, mis en place par la CNAM en mai 2020, pour **assurer le *tracing* des cas contacts et des chaînes de contamination** ;
- **Vaccin Covid**, développé en décembre 2020 par la CNAM dans le cadre d'une équipe projet pilotée par la DNS, pour **suivre en temps réel la campagne de vaccination**.

Ces trois outils, construits dans des temps records (par exemple, trois semaines pour SI-DEP), sont identifiés par les acteurs comme une avancée majeure dans le pilotage de la crise. Pour y parvenir, des groupes de travail ont été mis en place avec l'ensemble des acteurs impliqués à tous les échelons (producteurs de données, responsables de la collecte et du traitement, utilisateurs des données) pour construire les outils et perfectionner dans le temps leurs usages.

En parallèle, des chantiers nationaux ont été conduits **pour affiner et fiabiliser les remontées de données mises en place et/ou utilisées lors de la phase précédente**. Un premier exemple est l'analyse des séjours SI-VIC d'une durée anormalement longue à l'été 2020, séjours qui ont été signalés aux ARS afin qu'elles puissent échanger avec les établissements de santé pour corriger si nécessaire les données. Des travaux de comparaison des données SI-VIC avec les données remontées « en routine » par les établissements de santé une fois les séjours clôturés et remontés dans le PMSI ont aussi été réalisés, pour identifier les écarts, et s'assurer que l'outil SI-VIC apportait une information fiable pour le suivi en temps réel de l'épidémie. Autre exemple, la reprise de l'outil privé MaPUI (suivi du stock des médicaments prioritaires) par l'ATIH et la DGOS sous forme d'un outil d'État en novembre 2020, **e-Dispostock**. L'objectif était d'avoir un outil réactivable en cas de besoin, limitant ainsi les coûts associés à l'externalisation d'un tel dispositif. Depuis, de nombreuses évolutions ont eu lieu notamment avec l'intégration de nouvelles fonctionnalités début 2021 (par exemple, le suivi des stocks de vaccins).

- **Une place croissante de la donnée dans le pilotage de la gestion de crise et un cadre d'utilisation des données plus sécurisant**

Dans l'ensemble, on observe au cours de cette phase une place croissante de la donnée **dans le pilotage de la gestion de crise. La mise en place de collectes de données fait partie intégrante des stratégies de réponse** à la crise, comme « tester, alerter, protéger ».

Si l'on reprend les quatre fonctions du pilotage :

- **le suivi de la situation est abouti** (existence d'indicateurs permettant un suivi de la situation en temps réel sur l'ensemble du territoire) ;
- **des processus de production et de diffusion de données sont mis en place pour alimenter la prise de décision** (prises de décision légitimées par l'utilisation d'indicateurs partagés) ;
- **la communication est renforcée** (utilisation forte des données et des indicateurs pour expliquer les mesures mises en œuvre) ;
- **une logique d'anticipation se met progressivement en place** (mise au point d'outils de prédiction par certains acteurs : Institut Pasteur, ARS Île-de-France...).

- **Cependant, des remontées et modalités d'utilisation des données encore cloisonnées pèsent sur le pilotage d'une crise sanitaire installée dans la durée**

Malgré ces avancées, on observe des difficultés persistantes dans les remontées de certaines données, notamment **celles peu structurées** (par exemple, l'outil Voozanoo de suivi des Ehpad, qui reste un outil de collecte peu normé) et **des processus non uniformisés** (par exemple, des ARS utilisant des outils de suivi des clusters différents).

Au-delà, il y a **une difficulté à avoir une vision unifiée de la situation et de l'effet des réponses**. En effet, les outils sont silotés et non interconnectés, ce qui implique toujours un travail fastidieux pour les acteurs impliqués dans la collecte et le traitement des données. Il existe également une profusion de tableaux de bord et d'indicateurs utilisés à tous les échelons, posant des questions de cohérence et de lisibilité de l'ensemble.

3. **Phase 3, printemps 2021 – septembre 2021** : une mobilisation des données dans un fonctionnement permettant la vie avec le Covid-19, et des réflexions sur la pérennisation des dispositifs

- **Des travaux autour des données visant à approfondir la compréhension de la situation, à affiner les actions de réponse et à mieux cerner leur impact ; mais les enjeux de collecte de données perdurent**

Dans un contexte de crise persistante, et de maîtrise progressive des rouages de la gestion de crise du Covid-19, **des travaux d'approfondissement autour des outils de crise sont conduits** :

- pour éprouver la robustesse de ces outils (par exemple, l'amélioration de la qualité du pseudonyme dans SI-DEP) ;
- pour affiner les prises de décision et soutenir la mise en place de nouvelles réponses de gestion de crise (par exemple, en juin 2021, des travaux d'appariement **entre les bases de données SI-VIC, SI-DEP et Vaccin Covid** ont été menés, ce qui a permis à la DREES et à SpF d'étudier l'efficacité de la stratégie vaccinale : lien entre statut vaccinal, infection et risque de développer une forme grave demandant une hospitalisation).

Cependant, de nouveaux défis en matière de données, liés à de nouvelles décisions de gestion de crise, **mettent également en évidence des enjeux restants en matière de collecte de données**. Ainsi en juillet 2021, dans le cadre de la décision de la vaccination obligatoire des personnels soignants, naît le besoin de connaître le statut vaccinal des soignants.

L'outil Vaccin Covid ne pouvant être utilisé, sauf dans certaines conditions et de façon subsidiaire, à des fins de contrôle de l'obligation vaccinale par les établissements pour des questions juridiques, les ARS et les établissements de santé ont dû utiliser des moyens alternatifs afin de vérifier le statut vaccinal du personnel des établissements de santé et des professionnels de santé libéraux.

- **Des premières réflexions sur les perspectives de pérennisation d'outils utilisés pour la gestion de la crise**

La crise a démontré l'intérêt de certains projets autour des données, voire a représenté un accélérateur fort. Les acteurs ont notamment évoqué le fait que la crise a confirmé **l'intérêt de mettre en place l'identifiant national de santé (INS) pour faciliter le partage de données**, afin d'apparier facilement les bases de données entre elles. L'INS est obligatoire dans les SI de santé depuis 2016. En revanche, son utilisation est toujours interdite en dehors du cadre du soin (toutefois une autorisation peut être donnée dans le cas d'une recherche). Mais son usage s'est beaucoup plus répandu depuis janvier 2021, ce qui va accélérer son

déploiement courant 2022. La crise a également permis d'accélérer la mise en place du ROR et l'uniformisation de son utilisation dans toutes les régions.

Des premiers travaux sont d'ores et déjà engagés dans une perspective de pérennisation de certains outils. Par exemple, concernant **le ROR, une trajectoire de centralisation progressive** est amorcée : l'objectif serait d'avoir un outil unique partagé par toutes les ARS, permettant de disposer d'une vision régionale et nationale quant au nombre de lits disponibles et à l'offre de soins. Toutefois, ce chantier de centralisation du ROR ne porte pas en soi d'amélioration de la qualité des données qui l'alimentent, en matière d'exhaustivité et de fraîcheur notamment. Des arbitrages doivent être rendus sur la priorisation des actions à mener, entre l'élargissement de son périmètre et la fiabilisation de l'offre de soins déjà intégrée, afin de répondre aux enjeux opérationnels de son utilisation par les professionnels et les établissements de santé.

D'autres outils pourraient également voir leur utilisation élargie à de nouveaux cas d'usage, comme c'est le cas de **SI-DEP qui devrait évoluer vers un nouvel outil concernant les résultats de biologie d'autres maladies** (par exemple, les bronchiolites infantiles).

II. LA MOBILISATION SANS PRÉCÉDENT DES ACTEURS AUTOUR DE LA DATA DANS LE CONTEXTE DE CRISE A MIS EN LUMIÈRE DES DÉFIS À TOUTES LES ÉTAPES DE LA COLLECTE ET DE L'UTILISATION DES DONNÉES

Dans cette troisième partie, une analyse détaillée des cinq grandes étapes de la collecte et de l'utilisation des données a été conduite afin d'étudier les modalités de fonctionnement mises en place pour chaque étape pendant la crise et de pouvoir en tirer des enseignements.

MODALITÉS D'IDENTIFICATION DU BESOIN, DE COLLECTE ET D'UTILISATION DES DONNÉES



Constats et enseignements tirés de la crise

Des difficultés de formulation et de priorisation des besoins en données, et un manque de coordination entre acteurs

La mise en place de la collecte et des remontées de données aux différentes phases de la crise s'est heurtée à des difficultés transverses

Le traitement, l'analyse et les travaux de visualisation des données ont été portés de manière très fragmentée par une multiplicité de parties prenantes

Un effort majeur de diffusion des données pendant la crise, mais des initiatives multiples et peu coordonnées, n'ayant pas répondu à l'ensemble des besoins

Les données ont été fortement mobilisées pour l'analyse de la situation, la prise de décision et la communication, mais insuffisamment utilisées pour l'anticipation

Gouvernance & Organisation

Une gouvernance et une organisation de crise multicentriques au niveau national, et présentant des difficultés de coordination entre les échelons (national / régional) qui ont nourri les difficultés rencontrées tout au long du processus : identification des besoins, collecte et utilisation des données

A. DES DIFFICULTÉS DE FORMULATION ET DE PRIORISATION DES BESOINS EN DONNÉES, ET UN MANQUE DE COORDINATION ENTRE ACTEURS

- En fonction des acteurs considérés et de leur rôle en matière de pilotage, la nature des besoins en données diffèrent

Plusieurs facteurs influent sur la nature des besoins en données des différents acteurs :

- **le périmètre et le domaine d'intervention** – les besoins sont ainsi/notamment différents si les acteurs ont en charge la gestion du système sanitaire (par exemple, le pilotage de l'offre de soins par le ministère de la Santé) ou plutôt la mise en œuvre de restrictions (par exemple, la mise en place du confinement et du déconfinement par le ministère de l'Intérieur) ;
- **la maille géographique** – périmètre national (gouvernement, HAS...), régional (ARS...) ou local (établissements de santé et médico-sociaux, professionnels libéraux, préfectures, mairies...) ;
- **la situation épidémique** – situation de pic épidémique (impliquant des besoins de données fréquents) ou situation de creux (impliquant une moindre fréquence de suivi, mais aussi une mise en alerte).

En conséquence, les besoins en données peuvent être, selon les acteurs, proches **mais avec des caractéristiques différentes** concernant la **nature des données** (le nombre de cas, le taux d'incidence, le nombre d'hospitalisations...), la **forme des données** (brutes ou agrégées, nominatives ou pseudonymisées), le **niveau de granularité** (national, régional, local), la **fréquence des remontées** (quotidienne, mensuelle...).

- Une difficulté dans la formulation des besoins en données pour la gestion de crise

Pendant la crise, de nombreux acteurs impliqués dans la production ou dans le traitement des données évoquent, dans l'ensemble, **des demandes insuffisamment contextualisées** : elles portent parfois sur « une donnée », sans précision de la question de pilotage sous-jacente et des éléments nécessaires à sa production (périmètre, fréquence de remontée attendue...). De plus, les demandes ont souvent lieu **dans une logique descendante**, avec des difficultés pour les acteurs de terrain à obtenir des précisions, ce qui complexifie leur capacité à fournir des réponses adaptées.

A été également constatée **une tendance à focaliser les besoins sur les données d'ores et déjà connues** et suivies hors crise, sans envisager l'ensemble du spectre. Ainsi, certains acteurs regrettent que l'effort de collecte de données concernant la maladie et ses effets, ou encore la médecine de ville, soit resté limité (effet « lampadaire »).

Aussi, **la coordination de la demande des acteurs du pilotage, en matière d'identification et de priorisation des besoins est restée limitée**. Les équipes en charge de la production et du traitement des données reçoivent de multiples demandes, proches mais non identiques, par des acteurs différents.

Par conséquent, **de nombreux besoins se sont avérés insuffisamment pris en compte dans la construction des premiers outils**, et ont engendré des difficultés opérationnelles dans l'utilisation des données puis des modifications *a posteriori*.

- Une amélioration progressive de la prise en compte des besoins des différentes parties prenantes dans la construction des nouveaux outils

On observe tout de même **une amélioration progressive dans la prise en compte des besoins** dans le cadre de la construction de nouveaux outils (ou de nouveaux usages) : exemples de SI-VIC, SI-DEP et Vaccin Covid.

En mars 2020, SI-VIC a été adapté pour suivre les hospitalisations. Cela s'est effectué dans l'urgence et sans coordination entre les acteurs, conduisant à des erreurs et des manques dans son usage (par exemple, une différenciation des types de lits non intégrée à l'origine entre réanimation et soins intensifs). Suite au besoin de données plus fines (distinction de la réanimation parmi les soins critiques par exemple), l'outil a été amélioré, malgré des limites persistantes (il n'a pas été conçu pour être utilisé aussi longtemps et pour une telle volumétrie).

En mai 2020, SI-DEP a été construit pour suivre les résultats des tests. Grâce à la constitution d'un groupe de travail multi-acteurs pour construire l'outil (DNS, DGS, DREES, SpF, industriels, laboratoires...), l'outil a pu couvrir un certain nombre de besoins exprimés, par les fournisseurs de données comme par les utilisateurs, même si le contexte d'urgence a limité la prise en compte de tous les besoins.

En décembre 2020, Vaccin Covid a été élaboré pour suivre la campagne de vaccination. On observe une **amélioration notable dans la prise en compte des** besoins des différentes parties prenantes, soulignée par les acteurs interrogés : mise en place d'un SI unique centralisé s'appuyant sur des groupes de travail multi-acteurs, anticipation du besoin de données de vaccination individualisées (ayant permis une génération rapide du pass sanitaire ensuite), construction associée d'un tableau de bord national utile à tous les acteurs, y compris en local (Datavac). Cependant, certains besoins qui au départ n'ont pu faire l'objet d'arbitrages se sont révélés indispensables, et ont causé des difficultés dans l'utilisation des données (par exemple, l'absence de données sur le lieu de résidence des personnes vaccinées).

B. LA MISE EN PLACE DE LA COLLECTE ET DES REMONTÉES DE DONNÉES AUX DIFFÉRENTES PHASES DE LA CRISE S'EST HEURTÉE À DES DIFFICULTÉS TRANSVERSES

- **Des difficultés d'expression et de priorisation des besoins qui se sont traduites dans les modalités de collecte devenues complexes, foisonnantes et redondantes au cours de la crise**

Les difficultés dans la formulation et la priorisation des besoins se sont directement traduites dans les collectes de données. Comme il l'a été montré dans la deuxième partie, de nombreux circuits de collecte se sont mis en place avec souvent des doublons à tous les échelons, demandant un travail fastidieux aux acteurs (ceux en charge de construire les outils ou les responsables des collectes et des traitements...). Cela a un impact direct sur la **charge** de collecte pour les fournisseurs de données, mais aussi sur la **qualité** des données collectées, ainsi que sur leur **sécurité**.

- **Des référentiels insuffisamment partagés entre les acteurs, ce qui a complexifié la mise en place des outils de collecte**

Pendant la crise, les acteurs sont confrontés à **un manque majeur d'usage des référentiels communs** (Finess², RPPS³...). Par ailleurs, certaines données au sein des référentiels existants ne sont **pas à jour** (par exemple, les contacts des Ehpad dans le Finess) ou ne sont **pas couvertes** (pas de référencement des Samu privés, vétérinaires...). Au-delà, les référentiels existants se sont parfois révélés **insuffisants pour un usage de crise** (par

² Tous les établissements sanitaires et sociaux sont enregistrés dans le Fichier national des établissements sanitaires et sociaux (Finess), répertoire national géré par le ministère des Solidarités et de la Santé.

³ Le Répertoire partagé des professionnels intervenant dans le système de santé (RPPS référence l'ensemble des données d'identification, de diplômes, d'activité, de mode et de structure d'exercice de tout professionnel de santé.

exemple, Finess, un référentiel pas toujours à jour ou ne comprenant pas forcément les contacts nécessaires pour la communication avec certaines structures en temps de crise).

Cela a conduit à la construction d'outils de collecte dans des conditions fastidieuses : les acteurs ont dû **consolider les référentiels dans l'urgence, parfois à la main**, pour pouvoir les utiliser (par exemple, en passant des appels téléphoniques pour mettre à jour la liste des laboratoires pouvant faire des tests).

Mais la crise a également amené à concevoir de nouveaux référentiels pour répondre à certains besoins (par exemple, un référentiel des centres de vaccination), en tirant des enseignements des situations passées.

- **Un manque de langage commun pour désigner les données, qui a conduit à des confusions dans le processus de collecte**

Les acteurs évoquent également **un manque de langage commun pour désigner les données**. Est notamment constaté au démarrage un manque de compréhension et d'appropriation du langage technique médical par certains acteurs (par exemple, les décideurs politiques). Certaines notions frontalières font également l'objet d'interprétations différentes sur le terrain en fonction des établissements (par exemple, sens différent des termes réanimation et soins critiques). Il y a également certaines divergences sur les périmètres géographiques (comme la distinction entre Guadeloupe, Saint-Martin et Saint-Barthélemy par l'ARS Guadeloupe *versus* le regroupement sous le « chapeau » Guadeloupe par SpF). Tout cela s'explique en particulier par le fait **qu'il n'y a pas eu d'étape de définition**, de partage et de communication auprès de tous les producteurs et utilisateurs de données sur les notions clés utilisées pour la collecte et la remontée de ces dernières, dans un contexte de « course contre la montre » face à l'épidémie.

Ce manque de langage commun entraîne des incompréhensions entre acteurs, conduisant à **des besoins de retraitements, du travail en doublon et des communications**, y compris externes, **paraissant comme incohérentes** (par exemple, la difficulté pour les administrés de Guadeloupe à comprendre les écarts de chiffres entre la source SpF intégrant Saint-Barthélemy et Saint-Martin, et la source ARS, focalisant sur les données Guadeloupe).

- **Un cadre juridique qui s'est adapté au contexte de crise pour garantir la protection des nombreuses données collectées, mais qui a entraîné une complexité forte dans la construction des outils et processus dans ce contexte d'urgence**

La collecte de données est régie par **plusieurs textes juridiques majeurs** (notamment le règlement général sur la protection des données et la loi Informatique et Libertés), et fait l'objet **d'un contrôle** de la Commission nationale de l'informatique et des libertés (CNIL). L'adaptation d'outils existants et la mise en place de nouveaux outils **suivent des étapes juridiques précises** : présentation de dossier CNIL, publication de décrets et d'arrêtés...

Le contexte de crise a impliqué **la mise en place d'un régime juridique temporaire et exceptionnel pour accélérer le traitement des dossiers**, tout en respectant la protection des données. Dans l'ensemble, les acteurs interrogés expriment un consensus sur **l'intérêt de ce système** et une **reconnaissance des efforts** de la CNIL pour accélérer la prise en charge des dossiers.

Ils soulignent toutefois la complexité de ces démarches, et **questionnent certaines limitations imposées dans l'utilisation des données**. En particulier, il était difficile pour les acteurs de **gérer en parallèle** la construction d'un outil tout en assurant le cadrage juridique (de nombreux allers-retours avec la CNIL). Par ailleurs, le nombre important de dossiers devant être traités par la CNIL tout en conservant un niveau d'exigence fort a impliqué **des**

temps de traitement de dossiers longs, notamment sur des dossiers utiles à la gestion de la crise. Certains dossiers d'adoption de nouveaux outils ont impliqué de longs débats parlementaires amenant parfois à imposer **des contraintes perçues comme trop fortes** par certains acteurs compte tenu de l'urgence sanitaire (par exemple, sur l'application TousAntiCovid permettant d'alerter les cas contacts via le système Bluetooth, l'accès à la date de contact a d'abord été refusé juridiquement, rendant le traçage inutile puisque l'utilisateur ne pouvait pas connaître son temps d'incubation et combien de personnes il avait pu contaminer. Cela a été rectifié seulement six mois plus tard). Certains acteurs ont été contraints, devant certaines situations urgentes, de devoir réaliser des modifications et de prévenir la CNIL en parallèle pour ne pas ralentir la collecte de données. D'autres jugent que **l'accès à certaines données a été trop restrictif** dans ce contexte de crise, freinant certains projets (notamment les travaux des chercheurs).

- **Un constat partagé par les acteurs sur le manque de certaines données majeures pour le pilotage de la crise**

Les acteurs et plusieurs rapports publics ont mis en évidence l'absence **de collecte de certaines données pouvant jouer un rôle majeur** pendant la crise, mais aussi en période hors crise, pour le pilotage du système sanitaire :

- des données pour lesquelles des solutions temporaires ont été mises en place dans l'urgence pendant la crise **demandant à être consolidées**, telles que les données des Ehpad. Avant la crise, ceux-ci faisaient l'objet d'une enquête seulement tous les quatre ans, réalisée par la DREES, la dernière datant de 2015. En mars 2020, devant le manque criant d'informations sur la situation en Ehpad, SpF met en place dans l'urgence un questionnaire en ligne « Voozanoo » sur la plateforme de signalement des maladies respiratoires aiguës. Cependant, ce questionnaire reste simple, basé sur du déclaratif et des traitements manuels, et demande à être consolidé et pérennisé ;
- des données disponibles mais qui sont lacunaires et pour lesquelles **la mise en place d'un outil partagé à l'échelle nationale paraît également indispensable** (par exemple, les données de la médecine de ville).

C. LE TRAITEMENT, L'ANALYSE ET LES TRAVAUX DE VISUALISATION DES DONNÉES ONT ÉTÉ PORTÉS DE MANIÈRE TRÈS FRAGMENTÉE PAR UNE MULTIPLICITÉ DE PARTIES PRENANTES

- **Les difficultés rencontrées dans le processus de collecte des données se sont traduites par des difficultés dans la compréhension, le traitement et la mise en qualité des données**

Les collectes de données complexes et parfois redondantes ont conduit à la remontée de données de qualité inégale. Cela a conduit à des travaux de **traitements de données fastidieux, chronophages et redondants entre les acteurs, par des équipes qui n'ont parfois pas d'agents aux compétences poussées en data** (pour comparer les données, comprendre et neutraliser les biais, etc.). À cela s'est ajouté un **manque de partage des règles** de traitement entre acteurs, notamment avec les ARS, conduisant à des **enjeux de cohérence et d'interprétation** des données (par exemple, un désaccord au démarrage sur la date retenue pour les tests – date de prélèvement *versus* date de résultat – donnant lieu à des présentations de données hétérogènes selon les acteurs, posant des questions de cohérence vis-à-vis du public).

Or, l'enjeu de **travailler de manière coordonnée, notamment pour identifier les erreurs**, est apparu comme majeur (par exemple, lors du repérage par les ARS de doublons dans les données SI-DEP liés à des erreurs de saisie, grâce au fait que les ARS ont accès aux données

nominatives – ce qui n'était pas le cas des acteurs nationaux). Dans ce contexte, **des initiatives coordonnées ont été saluées par les acteurs interrogés**, notamment le partage des programmes de traitement par **le service à compétence nationale des SI mutualisés des ARS** du secrétariat général du ministère des Solidarités et de la Santé. Celui-ci a pris l'initiative de partager avec les ARS, peu associées dans les échanges sur les collectes de données, des programmes standardisés permettant d'exploiter et de produire des tableaux de bord adaptés aux besoins, à partir des données produites par les grands SI nationaux (SI-VIC, SI-DEP...). Un autre exemple est la restitution par la DREES des données du ROR et de SI-VIC via des visualisations cartographiques à destination première des ARS (par exemple, la cartographie des lits installés et occupés) ; puis, dans un deuxième temps, d'un tableau de bord rassemblant toutes les données de suivi de l'épidémie, ouvert à toutes les parties prenantes du pilotage national et régional, et aux chercheurs.

- **Des travaux peu axés sur l'anticipation et la construction de modèles prédictifs**

Les travaux de traitement décrits par les acteurs interrogés ont été fortement centrés sur l'analyse de la situation : **une minorité évoque des travaux conduits dans une logique d'anticipation**. Cela s'explique au démarrage de la crise par le **contexte d'urgence** et le **manque de données disponibles**, conduisant les acteurs à prioriser la mise au point d'un suivi en temps réel. Cela s'explique également par **un manque de compétences** requises pour réaliser des modèles prédictifs au sein de l'administration – des compétences data et statistiques, mais également des connaissances métiers (épidémiologie...) – et par l'absence de partenariats institutionnalisés avec des équipes de recherche. **L'Institut Pasteur et des figures clés de la recherche** (Vittoria Colizza, Simon Cauchemez...), aux fortes compétences en épidémiologie, **ont joué un rôle majeur** en proposant des premiers modèles de prédiction à l'échelle nationale. Mais ils l'ont fait dans un cadre peu formalisé, et les résultats de ces modèles n'étaient pas toujours mobilisables « tels quels » pour les besoins opérationnels (par exemple, la modélisation de flux et non de stocks). Au fil de la crise, d'autres acteurs, à différents échelons, ont **engagé des travaux autour de modèles prédictifs**, qui ont démontré leur intérêt pour le pilotage et l'anticipation (par exemple, des modèles construits par des infectiologues au sein des établissements ; l'outil de prédiction des occupations des lits STEP de l'ARS Île-de-France).

- **Des outils de visualisation qui se sont multipliés à tous les échelons, invitant à une réflexion sur les leviers de simplification et de mutualisation, tout en préservant une flexibilité pour répondre aux besoins spécifiques de chaque acteur**

Pendant la crise, **les visualisations et tableaux de bord ont joué un rôle majeur pour présenter les données et ont démontré le caractère essentiel de ces instruments** pour permettre aux décideurs d'accéder rapidement à une vision synthétique et parlante des données et pour communiquer auprès du grand public.

Cependant, pendant la crise, le manque de coordination a conduit à **une multiplication des outils de visualisation** à tous les échelons du pilotage impliquant à chaque fois un travail important d'identification des données clés, de conception des modes de représentation adaptés, etc. (par exemple, au niveau local – production d'un tableau de bord par chacune des ARS à partir des mêmes types de bases de données, comme SI-DEP ou Vaccin Covid ; au niveau national – production d'un tableau de bord individuel par chaque acteur, SpF, DREES, cellule interministérielle de crise, CCS, CNAM...).

Cela invite à réfléchir à des leviers de simplification et de mutualisation. Il existe en effet **un intérêt fort à construire un outil de visualisation commun** pour assurer la cohérence des

données partagées ; limiter le travail en doublon (lorsque chaque acteur produit ses visualisations) et mutualiser les moyens ; permettre un accès élargi aux données, de façon sécurisée.

Il existe tout de même la nécessité **d'une certaine flexibilité dans l'outil pour répondre aux besoins propres** à chaque acteur (par exemple, ceux d'une ARS et d'une cellule de crise nationale – qui relèvent de différences d'échelle, locale ou nationale, de types de données, de fréquence, de public cible...). **Une amélioration** progressive a été constatée avec la mise en place pour le suivi de la campagne de vaccination de **l'outil de visualisation unique Datavac**, qui, dès sa conception, se voulait adapté aux besoins de tous les acteurs (acteurs nationaux, ARS...).

D. UN EFFORT MAJEUR DE DIFFUSION DES DONNÉES PENDANT LA CRISE, MAIS DES INITIATIVES MULTIPLES ET PEU COORDONNÉES, N'AYANT PAS RÉPONDU À L'ENSEMBLE DES BESOINS

- **Une diffusion des données parfois peu aisée au sein de l'État**

Les acteurs ont mis en place certaines actions pour assurer une diffusion centralisée des données aux responsables du pilotage des différents échelons. On peut notamment noter, au niveau central, **la mise en place fin avril 2020 par la DREES, en lien avec la DNUM et le pôle « data plateforme recherche innovation » dédié au suivi de la recherche et au partage des données du CCS, d'une plateforme de partage** rassemblant l'ensemble des sources de données provenant de différents acteurs, et destinées aux pilotes nationaux, aux ARS, mais aussi accessibles aux chercheurs. Au niveau local, **de nombreuses diffusions** de données, sous forme de tableaux de bord, **par les ARS à leurs partenaires**, notamment les préfetures, à une fréquence très rapprochée, voire quotidienne.

Cependant de nombreux acteurs évoquent **une difficulté persistante à obtenir des données en interne** tout au long de la crise, notamment au niveau local : des données non communiquées, avec du retard, ou par un mauvais canal de diffusion. Par exemple, ce manque de partage a conduit certaines ARS à devoir **attendre la publication des données en Open Data pour le grand public** et à les recopier à la main (comme les données de dépistage autour de juin 2020). A également été constaté **un sentiment de frustration dans les établissements de santé**, lié au fait de passer du temps à remonter des données à travers de multiples outils, avec peu de visibilité en retour. Cela s'est traduit dans certains cas par un manque de motivation pour répondre aux enquêtes et par des besoins non assouvis, comme par exemple celui de disposer de données très localisées, ainsi que des données des établissements voisins pour anticiper les besoins.

Les chercheurs expriment également un manque d'accès à des données utiles à des travaux pouvant faire avancer la compréhension de la situation et améliorer les prises de décision.

Cela est dû, selon les acteurs, à plusieurs enjeux :

- **une culture du partage de la donnée peu répandue au sein de l'État** (des organisations silotées et une peur soulignée par les acteurs de diffuser leurs données pour des raisons de sécurité ou de mésusage, notamment entre le niveau national et le niveau régional) ;
- **un fort enjeu de visibilité des administrations et une forme de « concurrence » entre acteurs renforçant ce manque de partage** (un défi à être le premier à publier en Open Data ou à exploiter un jeu de données, pouvant s'expliquer par un désir de valoriser tout le travail réalisé en amont) ;
- **des enjeux de sécurité et des contraintes juridiques dans l'utilisation des données de santé, afin de protéger les personnes concernées** (des contraintes

notamment ressenties par certains chercheurs, qui regrettent de ne pas avoir pu accéder plus facilement à des données pour leurs travaux, même s'il faut souligner la mise en place de travaux de recherche beaucoup plus rapide qu'en temps habituel) ;

- **des modalités opérationnelles de partage non adaptées au sein de l'État, en matière de processus et d'infrastructures** (un manque d'outil commun de partage adapté – avec un niveau de sécurité garanti, une capacité à faire transiter des flux importants de données, etc.) ;
- **un coût important de préparation et de mise en forme des données, d'automatisation des extractions et de la production des statistiques, de la documentation des données.**

Les coûts techniques, juridiques et en ressources humaines pour la diffusion et le partage des données sont à prendre en compte dans la gestion de crise.

- **Un effort exceptionnel de l'État pour la mise à disposition de données en Open Data se révélant un atout clé pour mobiliser l'expertise d'acteurs externes et diffuser l'information auprès du grand public**

Les efforts de l'État en matière de partage de données en Open Data pendant la crise ont été soulignés par les acteurs. Ils se sont traduits par la **mise à disposition de nombreuses bases de données sous forme de fichiers « plats »**, avec un accompagnement d'Etalab, sur Data.gouv.fr ; **le partage de nombreuses visualisations très parlantes** lors de conférences de presse et sur divers tableaux de bord en ligne (site du gouvernement, du ministère de la Santé...) ; **de nombreux temps d'échanges entre les acteurs publics et la presse** pour expliquer les données (par exemple, la tenue hebdomadaire de conférences avec la presse par SpF).

Ce développement de l'Open Data a pu s'appuyer sur **un écosystème d'acteurs privés** (collectifs citoyens, *data scientists* indépendants) qui s'est largement mobilisé pour partager du contenu qualitatif en Open Data, comme **des visualisations originales** – ce fut notamment le cas de sites de visualisations largement consultés par le grand public pour leurs vertus pédagogiques comme CovidTracker (lancé par Guillaume Rozier), Germain Forestier Info (conçu par Germain Forestier) et Météo Covid (créé par Guillaume Saint-Quentin).

De nombreux avantages sont reconnus à l'Open Data pour améliorer l'usage des données :

- un contrôle qualité renforcé avant la publication des données par les producteurs ;
- **un contrôle externe de la cohérence des données** (par exemple, des *data scientists* se saisissent des données disponibles, les analysent et identifient parfois des incohérences, qu'ils font remonter auprès des acteurs publics) ; en temps de crise et malgré le contrôle qualité interne, les données sont diffusées très rapidement après leur production, et des incohérences sont possibles du fait des délais de production très courts ;
- **une stimulation des initiatives privées via la mise en place de visualisations intéressantes** (par exemple, des cartes thermiques pour montrer le taux de positivité en fonction de l'âge par Germain Forestier) **et via la conception de nouveaux outils opérationnels rendant la gestion de crise efficace** (par exemple, la création de l'outil ViteMaDose par Guillaume Rozier pour trouver et réserver rapidement un créneau de vaccination) ; les visualisations ont rarement été produites par l'État, non pas par manque de compétence, mais du fait que les ressources humaines en *data scientists* au sein de la cellule de gestion de crise ont été priorisées sur la production des données, leur analyse et leur restitution pour le pilotage ; de plus les circuits de validation n'étaient pas adaptés à une communication rapide d'outils de visualisation dédiés au public ;

- **un outil de communication pédagogique auprès du grand public** (par exemple, les visualisations graphiques utilisées lors des conférences de presse, ou encore le compteur des vaccinations sur la façade du ministère).
- Cette politique d'ouverture des données a également mis en lumière certains enjeux pour améliorer encore l'impact de l'Open Data

Tout d'abord, les acteurs mettent en avant un besoin de rationalisation des nombreux canaux de diffusion de données disponibles en Open Data comme **condition clé de la cohérence des données diffusées et de la crédibilité de l'État**. Parmi les canaux de diffusion utilisés au cours de la crise, on peut citer Data.gouv.fr (Etalab), les géodes et sites officiels de SpF, le site du gouvernement, le site du ministère des Solidarités et de la Santé, Santé.fr, le site de la CNAM (pour les données de vaccination), le site de la DREES (appariement des données SI-VIC, SI-DEP et Vaccin Covid).

On retrouve parfois sur ces sites des données différentes (par exemple, le lieu de résidence utilisé par la CNAM *versus* le lieu de vaccination disponible pour SpF, problème ensuite résolu), ce qui peut rendre difficile l'identification de l'information cherchée en l'absence d'explication claire et visible des différences entre les sources. Dans d'autres cas, les données sont proches mais avec certaines subtilités qui posent des difficultés d'interprétation aux acteurs (par exemple, la diffusion sur le site Data.gouv.fr des données de vaccination *versus* la publication sur le site de la CNAM des données de vaccination associées à des causes de comorbidités).

Les acteurs mettent également en évidence **un questionnement plus large sur le rôle de l'État en matière de diffusion auprès du grand public**, et la limite pertinente de son action : l'État doit-il uniquement produire et mettre à disposition des données (en laissant des citoyens se charger des visualisations) ou bien doit-il également mettre à disposition du public des visualisations abouties ?

Les acteurs évoquent notamment un malaise **au sein de l'administration au sujet des initiatives privées devenues très visibles et mises en avant dans la communication notamment des responsables politiques**. S'est exprimé le sentiment que les efforts en arrière-plan de l'État sur la production et la mise à disposition de données n'étaient pas reconnus et valorisés, alors qu'ils représentent un investissement très important en ressources techniques et humaines, et que l'État dispose également d'une capacité propre à produire des outils de visualisation grand public. Mais il doit tenir compte de nombreuses exigences, que ne connaissent pas les acteurs privés (charte d'accessibilité de la communication de l'État, par exemple).

Au contraire, certains acteurs au sein et hors de l'administration sont à l'aise avec l'idée d'un rôle de l'État s'arrêtant à la mise à disposition des données, charge ensuite aux citoyens de s'en saisir, de les « challenger », dans une logique constructive, et d'imaginer des modes de visualisation (par exemple, les travaux de Germain Forestier) et des usages (par exemple, Vite ma dose) innovants.

E. LES DONNÉES ONT ÉTÉ FORTEMENT MOBILISÉES POUR L'ANALYSE DE LA SITUATION, LA PRISE DE DÉCISION ET LA COMMUNICATION, MAIS INSUFFISAMMENT UTILISÉES POUR L'ANTICIPATION

Dans l'ensemble, les données ont immédiatement **représenté un pilier du pilotage de la gestion de crise**, en particulier pour décrire la situation et communiquer, puis progressivement pour prendre des décisions, et, dans une moindre mesure, pour anticiper la situation.

1/ Analyser la situation. Dès les premiers mois de la crise, les acteurs ont su mettre en place **un suivi de la situation** appuyé sur les données, via la réalisation de nombreux tableaux de bord à tous les échelons fournissant un point de situation en temps réel aux décideurs. Des analyses de données ont été proposées sur de nombreux sujets pour suivre la situation (par exemple, la pression épidémique) et pour suivre l'impact de décisions (par exemple, la vaccination). **Les enjeux** identifiés sur cet axe sont :

- de **lever les difficultés** juridiques et techniques **de partage de données entre les différents échelons du pilotage** (pas de serveur partagé entre les échelons régional et national où exploiter les mêmes bases à partir des mêmes programmes, par exemple) ;
- de **renforcer les collaborations** entre les acteurs pour rationaliser les travaux d'analyse de situation (par exemple, la construction de tableaux de bord) ;
- d'**exploiter plus rapidement des données** permettant une meilleure connaissance de la maladie.

2/ Communiquer. Les données ont également été **largement utilisées depuis le début de la crise pour communiquer**, notamment auprès du grand public (par exemple en tenant des conférences de presse avec des présentations de graphiques, ou la mise en place de tableaux de bord en ligne). Cette démarche s'est révélée efficace pour informer sur la progression de l'épidémie et faire preuve de pédagogie en invitant la population à changer ses comportements. **Les enjeux** identifiés sur cet axe sont :

- de **veiller à la cohérence des données** diffusées ;
- de **définir le rôle de l'État** en matière de diffusion de données et de visualisations.

3/ Prendre des décisions. Pendant la crise, les données ont été **progressivement intégrées dans le pilotage** de la gestion de crise (ainsi, au cours des premières vagues épidémiques, des décisions de confinement et de déconfinement ont-elles été prises en fonction de l'évaluation de plusieurs indicateurs clés – taux d'occupation des lits de réanimation, nombre de cas...). Cependant, plusieurs **difficultés** ont été évoquées par les acteurs interrogés :

- **le manque de certaines données indispensables pour le pilotage**, total ou partiel (par exemple, le taux d'occupation des lits via le ROR) ;
- **un manque de culture scientifique au sein des instances de pilotage**, rendant parfois difficiles la compréhension des données et la prise en compte de leurs biais ;
- **un questionnement par certains acteurs sur le niveau de prise en compte des données au-delà du champ hospitalier et sanitaire pour prendre les décisions**, posant la question de la capacité à recroiser les données sanitaires et non sanitaires dans la préparation des prises de décision.

Les enjeux identifiés sur cet axe sont :

- **de couvrir** les « angles morts » dans les données ;
- d'**améliorer la prise en compte des données et leur bonne compréhension** dans le champ décisionnel.

4/ Anticiper. L'utilisation des données à des fins d'anticipation s'est effectuée dans un **second temps de crise**, et a concerné seulement certains acteurs ayant les compétences nécessaires. En effet, au démarrage de la crise, face à l'impréparation du système sanitaire et de la collecte de données, les efforts se sont concentrés sur la mise en place de collectes de données pour alimenter en priorité les trois autres axes du pilotage (analyse de la situation, communication, prise de décision). Pourtant, il existe un intérêt majeur à mettre en place, dès que possible, des outils prédictifs pour anticiper au mieux et préparer une réponse de crise efficace. Une fois atteinte une certaine maîtrise de ces trois axes, certains acteurs ont mis en

place des outils de modélisation prédictifs (comme l'Institut Pasteur, l'ARS Île-de-France, des CHU...), mais cela implique des compétences spécifiques pas forcément disponibles dans chaque organisation. Les **enjeux** identifiés sur cet axe sont :

- de **favoriser le développement des compétences utiles à la modélisation prédictive** parmi les acteurs de la gestion de crise (via des compétences internes ou des partenariats), de renforcer la coordination entre acteurs sur les modèles prédictifs ;
- de **mettre en place des leviers pour diffuser les modèles** élaborés et les adapter aux besoins de chaque acteur (modélisation et formats de restitution).
-

F. UNE GOUVERNANCE ET UNE ORGANISATION DE LA GESTION DE CRISE SUR LA PARTIE DATA FRAGMENTÉES, QUI ONT NOURRI LES DIFFICULTÉS RENCONTRÉES TOUT AU LONG DU PROCESSUS D'IDENTIFICATION DES BESOINS, DE COLLECTE ET D'UTILISATION DES DONNÉES

- **Une absence de volet « données » dans la gouvernance de crise qui a conduit chacun des acteurs à agir sans coordination**

Une grande partie des acteurs interrogés identifie au sein du ministère des Solidarités et de la Santé un manque de gouvernance des systèmes d'information et de la donnée, particulièrement ressenti dans le contexte de crise. Il a été nécessaire de sécuriser dans des délais courts de nombreux flux de données et de coordonner l'avancement des chantiers liés à ces dernières. Par exemple, **un pôle « indicateurs »** a été mis en place au niveau du centre de crise (CCS) du ministère de la Santé, mais uniquement à partir d'avril 2020. Ce pôle indicateurs a été chargé de consolider et d'analyser les données collectées pour appuyer le pilotage et la prise de décision. Cependant, il a eu surtout **un rôle technique**, mais **pas de pouvoir d'arbitrage dans le champ des données** (notamment vis-à-vis des producteurs – SpF, CNAM...), **ni de rôle de pilotage transverse** des chantiers liés aux données. Ce pôle ne couvrait par ailleurs pas tous les sujets (par exemple, les données de vaccination remontant par un circuit *ad hoc*, avec notamment une *task force* dédiée).

De plus, l'administration dispose de compétences clés autour des données, mais elles sont dispersées et inégalement représentées au sein des acteurs du pilotage de la crise (par exemple, on trouve des compétences pointues à la DREES : *data scientists*, développeurs... ; des compétences plus limitées au sein de certaines directions métiers ; et des compétences fortes dans les ARS, notamment des statisticiens, mais inégales selon les régions). Le **manque de vision transverse de ces compétences** en matière de données au début de la crise a rendu plus difficile leur mobilisation rapide et peut expliquer pourquoi certains acteurs ont le sentiment de n'avoir pas été suffisamment associés (par exemple, les ARS).

Ce manque de gouvernance a entraîné **un manque de coordination global** au démarrage de la crise qui a ensuite perduré tout au long de la crise, à mesure que les acteurs construisaient leurs propres processus de pilotage. Il a été d'autant plus difficile d'établir cette coordination transverse que **les parties prenantes** (directions ministérielles, opérateurs...) **n'avaient pas l'habitude de travailler régulièrement ensemble sur des sujets liés aux données**, avec une répartition des rôles et responsabilités en temps de crise **non clairement définie** (par exemple, une répartition des périmètres d'intervention floue entre les principaux responsables de traitement des données au niveau national – SpF, CNAM et DREES – donnant lieu à des situations de flottement, mais aussi à des redondances).

La coordination a également été limitée avec le pilotage de crise hors du ministère des Solidarités et de la Santé, de nombreux acteurs mentionnant par exemple le **manque de liens entre le CCS et le centre interministériel de crise**. Les actions ont semblé plus

coordonnées au niveau régional et au sein des ARS, avec un appui sur les équipes statistiques existantes au sein des ARS et chez leurs partenaires (comme GRADeS), mais aussi sur des **liens bien définis** entre le directeur territorial de l'ARS et les préfets de chaque département.

- **Un pilotage en silo des chantiers « données » pendant la crise du Covid-19 et une vue d'ensemble difficile à obtenir**

De fait, la répartition des rôles et responsabilités de chacun des acteurs autour des données **s'est faite de façon spontanée, en fonction de leurs compétences respectives**, de leur « bande passante » au moment du besoin et des infrastructures dont ils disposent (par exemple, SI-DEP a été confié à l'AP-HP et un consortium d'acteurs privés notamment du fait de leur capacité à faire, des fortes compétences SI de leurs équipes, de la possibilité de s'appuyer sur des serveurs de résultats de biologie et une infrastructure préexistants et à même de remonter des résultats de laboratoires de biologie médicale).

L'organisation a de plus été très évolutive : des **structures** de pilotage adaptées ont été créées au fil de la crise (par exemple, avec la mise en place du CabCovid, qui préparait les synthèses à destination des instances décisionnaires comme le Conseil de défense et de sécurité nationale). Cette structuration avec un cabinet consacré au Covid-19 est considérée comme peu lisible en matière de rôle et de composition par les acteurs interrogés, y compris ceux impliqués dans le pilotage national de la crise.

En parallèle, en fonction des chantiers des **modalités d'organisation spécifiques ont été instaurées**. Par exemple, différentes *task forces* ont été mises en place pour piloter des grands chantiers (*task force* test, vaccin...). Cette organisation a été **plutôt bien comprise** au niveau national et local. Cependant, **cette organisation est restée très compartimentée** (par exemple, pour la vaccination, la *task force* vaccination assurait la remontée de données auprès du cabinet ministériel, sans passer par le CCS et le pôle indicateurs – si cela a un intérêt quant à la répartition de la charge de travail, le manque de coordination a rendu difficile la mise en cohérence des données).

- **Cette absence de pilotage transverse des sujets « données » a causé des difficultés opérationnelles**

Des **difficultés opérationnelles ont été particulièrement relevées sur les chantiers SI en raison de l'absence de pilotage transverse**. L'adaptation des systèmes d'information était **parfois tâtonnante**, car faisant face à des injonctions diverses de la part des acteurs du fait de l'absence d'un « *board* décisionnel data » rassemblant l'ensemble des parties prenantes⁴. Par exemple, l'Agence nationale du numérique en santé (ANS) a adapté SI-VIC au départ sans appui extérieur et a dû mettre à jour plus d'une dizaine de fois la doctrine de l'outil après les retours du terrain (par exemple, différencier les types de lits).

Par ailleurs, **des travaux ont été réalisés en doublon ou « pour rien »** du fait de l'absence de communication et de liens ou de ponts entre les différents chantiers (par exemple, en mars 2020, une ARS a lancé une application pour suivre les respirateurs et a dû tout changer au dernier moment pour utiliser à la place l'application nationale).

Au-delà, on observe que ces difficultés **s'inscrivent dans la durée** (par exemple, des régions continuent de ne pas utiliser le ROR, rendant difficile la visibilité sur les capacités de lits dans certaines régions au niveau national, et perpétuent l'emploi de canaux autonomes – comme des enquêtes).

⁴ Combes, S., Bacry, E. (2020, juin). Note de synthèse de la *task force* Data vs Covid-19.

Enfin, le manque de pilotage au niveau central a pu conduire à des phénomènes de « concurrence » entre administrations, pouvant notamment freiner la diffusion et le partage de données (par exemple, une volonté limitée de partager certaines données). Le faible niveau de formalisation des liens entre institutions a aussi conduit à les faire reposer en premier lieu sur les personnes, rendant ces liens vulnérables en cas de départ ou de mobilité des collaborateurs, même si, au fil de la crise, des habitudes de travail en commun se sont développées.

III. PISTES DE RÉFLEXION

Dans cette dernière partie, différentes pistes ont été envisagées pour accompagner la consolidation des systèmes de collecte et d'utilisation des données à des fins de pilotage du système sanitaire.

A. LA CRISE DU COVID-19 A ACTÉ L'IMPORTANCE STRATÉGIQUE DE COLLECTER ET D'UTILISER DES DONNÉES POUR LE PILOTAGE DU SYSTÈME SANITAIRE, ET A DÉMONTRÉ L'IMPORTANCE DE DÉFINIR « À FROID » LEURS MODALITÉS D'UTILISATION EN TEMPS DE CRISE

- La crise du Covid-19 a acté l'importance stratégique de collecter et d'utiliser les données pour le pilotage du système sanitaire, notamment en temps de crise

L'utilisation des données pour le pilotage de la crise du Covid-19 est devenue une évidence pour les acteurs du système sanitaire. Dès le démarrage de la crise, il y a une **volonté politique** affichée de **faire de la donnée un outil central de la gestion de crise** avec des enjeux de collecte et d'utilisation des données très visibles dans le pilotage de la crise sanitaire et plus largement du système de santé. Les acteurs interrogés ont également souligné **l'importance d'utiliser les données** pour les **quatre fonctions du pilotage** : l'analyse de la situation, l'anticipation, la prise de décision et la communication. Aussi, depuis deux ans, les acteurs ont **orchestré des remontées quotidiennes de données aux niveaux national, régional et local**, afin d'assurer un pilotage fin de la crise et la construction d'un panel d'indicateurs afin d'appuyer les décisions. Cela a notamment conduit à un **ancrage progressif des modes de fonctionnement entre organisations** sur la question des données (réunions de partage entre institutions...).

La crise du Covid-19 a également marqué un tournant pour les usagers du système de soins et la société civile en matière d'accès à l'information. **Les acteurs en charge du pilotage ont communiqué quotidiennement des données de situation** au grand public **pour rendre compte de l'action des pouvoirs publics et faire de la pédagogie** et ont **grandement poussé à la diffusion de la donnée en Open Data**. Les usagers se sont fortement approprié les données, notamment avec l'appui des médias et des initiatives de citoyens compétents en matière de données.

- Les échanges avec les acteurs interrogés ont montré que la construction du volet « données » de la gestion de la crise s'est faite parallèlement à la gestion opérationnelle

Au démarrage de la crise, **peu de données faisaient l'objet de remontées régulières et le schéma de crise n'était pas connu**, générant ainsi un **manque de visibilité** sur la situation et sur les remontées de données à mettre en place. **Les dispositifs** de gestion de crise existants **n'intégraient pas des modalités précises d'utilisation des données** (on peut ainsi mentionner l'existence des différents niveaux du CORRUSS, dont le CCS au sein de la DGS, mais sans compétence en données, appelant à la **mise en place d'un pôle indicateurs**

dans l'urgence). La culture de la donnée et son utilisation pour le pilotage étaient **encore peu développées** au sein du ministère des Solidarités et de la Santé.

Cette situation a impliqué de **construire dans l'urgence l'ensemble du volet data** de la gestion de crise. Pour cela, les acteurs ont conduit de nombreuses actions dans un laps de temps très court :

- **un travail sur la compréhension de la crise** (nature du virus, impacts...) ;
 - **l'identification des données utiles au pilotage** (pour le suivi de la situation, la prise de décision...) ;
 - **l'alignement des acteurs autour des mêmes référentiels et même notions** ;
 - **la mise en place des nouveaux systèmes d'information** (par exemple, la construction de SI-DEP, Vaccin Covid, etc.) ;
 - **la construction d'un cadre juridique** permettant de protéger les personnes, tout en accélérant les procédures.
- **Les forces et les faiblesses du pilotage du système sanitaire mises en évidence dans le contexte de la crise du Covid-19**

Pour construire ce volet données de la gestion de la crise, les acteurs ont pu s'appuyer sur plusieurs **forces** :

- **un réseau d'acteurs en charge du traitement et de l'analyse des données** (au niveau national, avec l'existence notamment de SpF, de la DREES, de plusieurs bureaux de la DGOS ; au niveau régional, avec les équipes de statisticiens des ARS et des cellules régionales de SpF ; au niveau local, au sein des établissements) ;
- **des compétences en collecte de données et en construction de systèmes d'information (informaticiens et statisticiens)** au sein de l'administration centrale et de l'écosystème ;
- **la mobilisation d'épidémiologistes spécialistes de la modélisation** (Institut Pasteur notamment) ;
- **une politique préexistante à la crise de développement de l'Open Data avec des procédures établies pour la diffusion des données** ;
- **un cadre juridique imposant une réflexion exigeante** sur l'utilisation et les usages des données collectées, même si ce cadre a pu paraître parfois bloquant.

Mais la construction de ce volet données a révélé aussi plusieurs faiblesses :

- **une organisation de gestion de crise sur la data qui n'avait pas été anticipée, et qui s'est reposée essentiellement sur la mobilisation personnelle** des acteurs ;
- **des périmètres non couverts dans les outils existants** de remontée de données, donnant lieu à des « angles morts » ;
- **une absence de références partagées entre l'ensemble des acteurs** de l'écosystème, notamment en matière de référentiels (méconnus ou insuffisants) et de langage commun (signification des données recueillies) ;
- **un manque d'articulation et d'interopérabilité entre les différents outils** de remontée de données et de gestion de crise ;
- **l'absence d'un responsable des données dans la gouvernance de crise pouvant arbitrer**, qui s'est traduite par un déficit de coordination et d'arbitrage dans la plupart des chantiers autour des données ;
- **un manque de culture de la donnée chez les décideurs** ;
- **un nécessaire accroissement de la littératie numérique du grand public** permettant la compréhension des dispositifs d'Open Data mis en place.

- **Des faiblesses qui ont généré une forte tension sur l'ensemble des acteurs du pilotage de la crise du Covid-19 et démontrent l'importance de définir « à froid » les modalités d'utilisation des données en temps de crise**

Il ressort des entretiens une **forte tension sur l'ensemble des acteurs** impliqués dans le pilotage et dans la gestion des données (décideurs, analystes et statisticiens, équipes SI, équipes métiers, etc.), à tous les échelons géographiques, notamment pendant la première phase de la crise avec une mobilisation exceptionnelle et une forte augmentation du temps de travail, mais simultanément des collaborations exceptionnelles entre les structures publiques (par exemple, l'intervention de collaborateurs DREES pour la mise en place du pôle indicateurs du CCS).

Le manque de préparation du système en matière de données a également eu **un impact fort sur la continuité des activités non liées au Covid-19**, avec une dépriorisation et un report d'une grande partie des activités pendant de nombreux mois consécutifs (ainsi, plus de la moitié de l'effectif de la sous-direction en charge de l'observation de la santé et de l'assurance maladie de la DREES a été mobilisée sur les aspects « données » pendant les premiers mois de la crise). Dans certaines situations, **des sujets sensibles ont dû être traités « à chaud », sans prise de recul et sous la pression des délais**, avec en particulier des sujets nécessitant l'implication des parlementaires et de la société civile, difficiles à gérer dans un contexte d'urgence (par exemple, le cadre juridique d'utilisation et de réutilisation des données personnelles en situation de crise).

Tout cela a montré l'importance **de mieux anticiper le volet données en cas de prochaine crise**. Plusieurs sujets, notamment, pourraient faire l'objet d'une meilleure anticipation :

- **le cadre juridique** d'utilisation des données ;
- **les outils de collecte, de traitement, de visualisation et de diffusion** des données (mobilisation des outils quotidiens en contexte de crise avec des fonctionnalités étendues spécifiques aux situations de crise, construction d'outils *ad hoc*...) ;
- **l'existence d'un socle de données permettant un suivi de premier niveau de la situation sanitaire et de l'offre de soins** ;
- **les infrastructures et moyens techniques** (puissance de calcul, serveur partagé entre l'échelon national et les ARS pour assurer une production de données cohérente et rationalisée) ;
- **les modalités de coordination** (qui fait quoi, quand ?) **et de gouvernance** autour des sujets de données en contexte de crise ;
- **le développement de la culture de la donnée chez les décideurs, via la formation initiale et continue.**

B. L'ÉTUDE A PERMIS DE METTRE EN LUMIÈRE TROIS AXES CLÉS POUR RENFORCER LE PILOTAGE DES CRISES SANITAIRES GRÂCE À L'UTILISATION DES DONNÉES, EN PRÉPARATION D'UNE CRISE FUTURE

Trois axes clés sont avancés afin d'améliorer la collecte et l'utilisation des données dans le cadre du pilotage de la gestion de crise sanitaire.

1 – La mise en place d'une vision prospective des besoins en données nécessaires au pilotage des futures crises sanitaires, via l'élaboration de scénarios de crise

Le premier axe clé à développer concerne la réalisation, pour l'ensemble des types de crises sanitaires envisageables, **d'un exercice approfondi de prospective sur la nature des données à collecter, les outils à construire et le cadre à donner**, en associant les élus, les

usagers et leurs représentants, afin de **valider un cadre partagé** et de permettre aux acteurs de se concentrer sur la gestion opérationnelle lors de l'arrivée d'une crise.

Les crises sanitaires peuvent être considérées comme *tout événement, de caractère inédit, affectant la santé d'un grand nombre d'individus, avec un impact possible sur le niveau de mortalité*. Elles peuvent être de **différentes natures**, par exemple : épidémie ou pandémie liée à un virus (grippe, Covid...) ; incident nucléaire ; événement climatique extrême ; propagation d'un agent chimique (à la suite d'un accident ou d'une attaque terroriste par exemple) ; événement engendrant une traumatologie lourde (à la suite d'un accident ou d'une attaque terroriste là aussi).

Un modèle unique de crise sanitaire n'existe pas. Chaque type de crise sanitaire a des enjeux spécifiques en matière de pilotage, selon plusieurs dimensions :

- **le périmètre géographique touché** – localisé, sur un territoire plus ou moins étendu (par exemple, suite à un événement climatique), ou évolutif jusqu'à l'ensemble du territoire (par exemple, suite à une épidémie ou pandémie) ;
- **la fréquence de remontée des données** (quotidienne, voire en temps réel, hebdomadaire ou mensuelle) ;
- **les principaux effets sur la santé et la nature de la prise en charge nécessaire** – un effet connu (par exemple, traumatologie à la suite d'une attaque) ; des effets non connus (par exemple, pandémie liée à un nouveau virus) ;
- **l'impact sur l'offre de soins** – services mobilisés dans les établissements (par exemple, réanimation) ; éventuelle destruction de capacités (par exemple, événement climatique extrême engendrant la disparition d'infrastructures de soins) ;
- **les besoins d'anticipation** et de modélisation ;
- **les besoins d'adaptation des prises en charge des patients** (pratiques professionnelles) ;
- **la communication aux citoyens – rôle de l'État** en matière de diffusion de données et de visualisations ;
- **d'autres caractéristiques** pouvant nécessiter un suivi de données spécifiques pour piloter la réponse sanitaire (par exemple, éventuel impact sur les modalités de communication et de déplacements lors d'un événement climatique extrême, etc.).

L'objectif de ce travail de prospective serait d'élaborer des **scénarios de crises** possibles, au niveau national, et d'anticiper les besoins de pilotage et de données associés, en assurant une bonne prise en compte du spectre de crises envisagées. Pour cela, il conviendrait **d'impliquer l'ensemble de l'écosystème** potentiellement en charge de la gestion de crise afin de garantir un continuum entre la théorie et la mise en œuvre et de **capitaliser sur les initiatives existantes** d'élaboration de scénarios prospectifs aux niveaux national (SpF) et régional (ARS en lien avec les préfectures).

Les besoins en données des acteurs peuvent être proches **mais avec des caractéristiques différentes** en matière de **nature des données** (nombre de cas, taux d'incidence, nombre d'hospitalisations...) ; de **forme des données** (brutes ou agrégées, nominatives ou pseudonymisées) ; de **niveau de granularité** (national, régional, local) ; de **fréquence de remontée** (quotidienne, mensuelle...). Par ailleurs, il faudra particulièrement cibler les champs dans lesquels les données ont manqué lors de la crise du Covid-19, notamment le médico-social et la ville. Il est également important dans ces travaux de prospective de travailler sur les données scientifiques à collecter, afin de nourrir les travaux de modélisation et d'améliorer les dispositifs et les pratiques professionnelles adaptées à la prise en charge des patients. Enfin, un tel travail devrait s'intéresser aussi à l'intégration des données sociales permettant de caractériser l'impact de la crise dans les différentes parties de la société.

Sur chaque scénario, un volet données devrait être associé, visant à déterminer :

- **les données et outils existants utiles** aux fonctions de pilotage sur lesquels capitaliser ;
- **les initiatives prioritaires à conduire en matière de données** (collectes additionnelles à lancer, éventuelles adaptations d'outils existants, appariement entre les bases, langage commun pour désigner les données dans les processus de collecte et d'analyse, serveur partagé entre les échelons régional et national où exploiter les mêmes bases à partir des mêmes programmes, outils de visualisation communs, acteurs et compétences à mobiliser), **et le cadre associé à ces actions** (cadre juridique et de sécurité des données notamment) ;
- **les acteurs devant accéder aux différents types de données et les usages associés** (afin d'anticiper notamment le cadre juridique à donner).

Pour cela, l'analyse des quatre fonctions du pilotage est un instrument utile pour définir le type de données utiles à collecter.

Fonctions de pilotage		Type de données à collecter	Exemple de données utilisées dans le contexte de la crise COVID
Suivi de la situation	Besoins nouveaux auxquels doit répondre le système de santé	Nature et niveau des impacts sanitaires de la crise	Populations à risque, taux d'hospitalisation
		Recensements des cas	Nombre de tests + sur un territoire donné
	Offre de soins disponible	Référentiels de prise en charge	Protocoles de soins mis en place dans le cadre d'une hospitalisation
		Capacité du système sanitaire au regard des référentiels de prise en charge	Nombre de lits de réanimation disponibles
Prise de décision	Attitude population	Niveau de perception des actions conduites	Niveau d'adhésion à la politique vaccinale
	Choix et allocation des moyens / dimensionnement de la réponse	Mesures de santé publique	Niveau de circulation du virus, pour décider d'un confinement
		Moyens humains	Suivi des ETP disponibles au sein des établissements
		Moyens matériels	Suivi des stocks EPI ou des respirateurs disponibles au sein des établissements
Communication		Moyens financiers	A préciser
	Suivi des actions de réponse à la crise sanitaire	Suivi du déploiement des actions	Suivi du déploiement de la campagne de vaccination
		Mesure de l'impact des actions	Impacts de la vaccination sur les taux de contamination ou d'hospitalisation
	Vers les acteurs de la réponse	Diffusion des données clés de situation et des mesures prises aux acteurs du pilotage	Données consolidées sur la dynamique épidémique
Anticipation	Vers la population générale	Information grand public sur la situation	Données consolidées sur la dynamique épidémique
	Anticipation de l'évolution des cas	Évolution des cas	R0, données sur la mobilité de la population
	Anticipation de l'effet des actions	Effet des actions de réponse	Données de recherche sur l'impact de la vaccination

Un tel exercice amènerait sans aucun doute à mettre en avant des données absentes pour le moment et utiles aussi en situation « hors crise », comme un suivi des motifs de consultation en médecine de ville ou des remontées plus rapides des causes médicales renseignées sur les certificats de décès.

2 – L'intégration de la question des données dans la conception et la préparation d'un modèle de gouvernance de crise

Le deuxième axe proposé est la **prise en compte et l'intégration complète de la question des données dans la définition du modèle de gouvernance** de crise, afin de **garantir une fluidité entre acteurs autour des données** (avoir anticipé le « qui fait quoi », les modèles de collaboration, etc.).

S'il n'existe pas de « gouvernance type » de crise en matière de données, la crise sanitaire due au Covid-19 a mis en lumière le besoin d'une stratégie globale des données, afin de

disposer d'une vue d'ensemble des outils existants et des projets en cours et d'arbitrer sur l'ensemble du périmètre pour mettre en œuvre la stratégie définie. L'objectif est d'assurer la meilleure qualité de données possible, d'éviter les redondances comme les manques et d'optimiser les ressources (humaines, techniques, financières...).

Pour cela, trois axes d'intervention ont été identifiés :

- **renforcer les dispositifs existants de gestion de crise** en matière de **compétences en données** – assurer au CCS la présence de personnes avec des compétences en données, en période de crise notamment (identifier en amont les ressources à mobiliser) ;
- **désigner un pilote unique des chantiers liés aux données** – créer un rôle de pilote unique des chantiers liés aux données, avec une fonction pivot entre décideurs, producteur et utilisateurs des données, et un positionnement « haut niveau » pour arbitrer les sujets données ;
- **mettre en place une cellule intégrée à la gouvernance de crise, dédiée au sujet des données** – sécuriser au sein de l'organisation de la gestion de crise une équipe en charge de l'ensemble des chantiers de données, en appui du pilote unique.

De plus, ce volet données de la gouvernance de crise devrait être mis en œuvre en s'appuyant sur plusieurs principes directeurs :

- **assurer une lisibilité forte du volet données de la gouvernance de crise** vis-à-vis de l'ensemble des parties prenantes ;
- **s'appuyer sur des liens forts et pérennes entre institutions**, renforcés par les collaborations hors crise – au niveau national et entre les niveaux national, régional et local ;
- **systématiser l'association de l'ensemble des parties prenantes** dans la conduite des chantiers sur les données, pour coordonner les actions et initiatives (nationales, régional, local) ;
- **soutenir la prise d'initiative des acteurs de terrain et repérer les outils pouvant profiter à un périmètre plus large d'acteurs** ;
- **organiser la diffusion et le partage des données** aux différents acteurs nationaux, régionaux, locaux, et à leurs partenaires hors champ de la santé ;
- **identifier les modalités de retour d'information aux acteurs locaux, notamment ceux qui remplissent les SI** ;
- **piloter la stratégie de diffusion en Open Data**, en tenant compte des **coûts associés** (techniques et humains notamment), **et garantir la cohérence des données publiées** par les différents producteurs.

3 – Dès maintenant, la réduction des obstacles à la collecte et à l'utilisation des données

Le troisième axe à approfondir est la **sécurisation de réponses robustes aux défis identifiés** par la crise du Covid-19 **sur l'ensemble de la chaîne de production et d'utilisation** des données.

Pour cela, il convient d'agir sur trois plans :

- **définir et appliquer un cadre commun à l'ensemble des parties prenantes en matière de données**. L'objectif sera de faciliter les échanges et de limiter les erreurs et le temps de coordination, dans le fonctionnement nominal et dans un contexte de crise. Pour cela, il s'agira de mettre à jour et de diffuser des référentiels communs (par exemple, dans le champ hospitalier du fait de l'autonomie des établissements dans leur choix de logiciels de gestion), avec des règles d'utilisation explicites (par exemple, Finess), de définir le langage commun définissant les variables enregistrées, de définir

des règles et des modalités de calcul des indicateurs clés partagées, et de mobiliser le cadre d'interopérabilité des systèmes d'Information de santé ;

- **identifier les « angles morts » de la collecte de données hors temps de crise et construire des modalités de collecte adaptées.** Il s'agira notamment de développer et d'améliorer les outils de collecte et de remontées rapides des données des établissements et services médico-sociaux (Ehpad notamment), de la médecine de ville, des ressources humaines en santé (professionnels de santé libéraux et salariés), des causes de décès, et de développer une vision « parcours de santé » des patients entre les différents compartiments du système de santé ;
- **engager une réflexion plus globale sur le schéma directeur des systèmes d'information, tenant compte du retour d'expérience de la crise.**

Pour y parvenir, plusieurs enjeux ont été identifiés :

- **évaluer le niveau de déploiement et d'appropriation des outils existants**, y compris les SI mis en place en urgence sur la pérennisation desquels il faut statuer, et le niveau de qualité des données produites et échangées ;
- **développer le numérique en santé dans les établissements et services de santé et médico-sociaux**, condition clé de la production de données avec un niveau de fiabilité et d'automatisation minimum, mais aussi fiabiliser et entretenir un annuaire d'adresses fonctionnelles pour les établissements et les services afin qu'ils puissent échanger efficacement avec eux ;
- **assurer l'existence d'un socle de données permettant un suivi de premier niveau de la situation sanitaire et de l'offre de soins** (par exemple, en développant le suivi des motifs de prises en charge dans le secteur libéral) ;
- **renforcer l'interopérabilité des systèmes d'information** (y compris au sein des établissements sanitaires, sociaux et médico-sociaux, et professionnels libéraux). Afin de : 1/ s'appuyer sur les SI existants pour collecter les données nécessaires à l'évaluation de chaque crise et son pilotage, et **ne pas créer de nouveaux SI qui nécessitent des saisies redondantes et peu optimisées** (par exemple, repenser le mode de remplissage de SI-VIC à partir des données hospitalières déjà existantes) ; 2/ faciliter l'échange de données entre les systèmes d'information natifs et ceux constitués pour la crise ; 3/ et ainsi faciliter les échanges entre les acteurs ;
- **définir les capacités et « niveaux de services » des outils et de leurs modes d'utilisation**, entre un fonctionnement « nominal » (par exemple, fréquence mensuelle, sur une partie du champ) et un fonctionnement « de crise » (remontées quotidiennes sur l'ensemble du champ), en lien avec les travaux de prospective sur les scénarios de crise ;
- **faire émerger un outil de collecte partagé** entre l'ensemble des acteurs de l'écosystème, permettant le lancement d'enquêtes et de collectes de données dans un cadre sécurisé, optimisé et connu de tous (au lieu d'enquêtes multiples via des canaux variés et peu sécurisés), intégrant si possible des interfaces avec les SI des établissements de santé et médico-sociaux, pour améliorer *in fine* la qualité des données collectées et le temps global dédié à la collecte et au traitement ;
- **concevoir une « infrastructure socle » partagée, sécurisée et performante pour l'exploitation de données** – partage de bases, de programmes, puissance de calcul.

Annexe. Rapports publics utilisés

- Assemblée nationale (2020, juin). *Impact, gestion et conséquences dans toutes ses dimensions de l'épidémie de Coronavirus-Covid-19* (rapport d'information).
- Combes, S., Bacry, E. (2020, juin). Note de synthèse de la *task force* Data vs Covid-19.
- Comité de contrôle et de liaison Covid-19 (2022, mars). *Données ouvertes en santé : une condition de l'agir ensemble, un levier pour la démocratie en santé* (avis).
- Cour des Comptes (2021, mars). *Rapport public annuel 2021. Tome I.*
- Ministère des Solidarités et de la Santé (2020, juin). *Projets numériques portés par la puissance publique dans le cadre de l'épidémie Covid-19.*
- Pittet, D. et al. (2021, mars). *Mission indépendante nationale sur l'évaluation de la gestion de la crise Covid-19 et sur l'anticipation des risques pandémiques* (rapport).
- Sénat (2020, décembre). *Évaluation des politiques publiques face aux grandes pandémies à la lumière de la crise sanitaire de la Covid-19 et de sa gestion* (rapport).