



Santé Post Incendie 76

Une étude à l'écoute de votre santé

Etude de santé et de qualité de vie auprès des habitants exposés à l'accident industriel de
Rouen du 26 septembre 2019

Protocole d'étude

Version du 28 décembre 2020

Eléments de mises à jour du protocole à la date du 28 décembre 2020

Le protocole de l'étude « Santé Post Incendie 76 - Une étude à l'écoute de votre santé », présenté dans ce document, est l'aboutissement d'un travail mené en différentes étapes.

Une première version, élaborée par l'équipe projet, en lien avec le comité d'appui thématique et le Groupe Santé, a : 1) été présentée à la commission Démographie et questions sociales du Conseil national de l'information statistique (Cnis) pour avis d'opportunité (28 novembre 2019) ; 2) fait l'objet d'un rapport d'expertise du pré-label (15 mai 2020) ; 3) été présentée au comité du Label du Cnis (11 juin 2020).

Depuis cette dernière date, plusieurs ajustements ont été apportés, tenant compte 1) des recommandations formulées par le comité du Label ; 2) des conclusions de l'enquête test réalisée en juin/juillet 2020, 3) des échanges approfondis avec la Division Sondage ainsi qu'avec le département des méthodes statistiques de l'Insee ; 4) des éléments relatifs à la protection et à la sécurité des données ; 5) des ajustements au protocole de collecte discutés avec le prestataire de collecte Ipsos Observer ; 6) de l'intégration d'un volet complémentaire sur la santé mentale (mis en œuvre en novembre-décembre 2020) et 7) de l'inclusion dans la population d'étude des personnes résidant sur l'aire d'accueil des Gens du voyage du Petit-Quevilly (décembre 2020).

Les modifications notables (par rapport au protocole soumis au comité du Label en avril 2020) figurant dans la présente version du 28 décembre 2020 concernent : 1) le calendrier mis à jour de l'enquête et de la collecte ; 2) l'ajustement du rayon de la zone de proximité immédiate (700 m au lieu de 500 m), opéré suite au comptage réalisé par la Divisions Sondages de l'Insee, permettant d'obtenir un effectif de taille suffisante de la population la plus exposée ; 3) les simplifications apportées au questionnaire suite aux conclusions de l'enquête test et ayant permis d'atteindre l'objectif d'une durée de passation moyenne inférieure à 30 minutes ; 4) l'investigation de méthodes d'évaluation du biais d'auto-sélection.

Par ailleurs, en complément à ce protocole, une demande d'accès aux données du Système National des Données de Santé (SNDS) pour appariement des données des personnes tirées au sort pour la présente enquête a été déposée en août 2020. L'exploitation de ces données auxiliaires a un triple objectif : 1) documenter l'état de santé de la population d'étude au moment de l'enquête et de manière prospective ; 2) utiliser ces informations auxiliaires, disponibles à la fois chez les répondants et les non-répondants, afin de quantifier le biais d'auto-sélection et, si nécessaire, corriger les estimations en prenant en compte le biais dans la modélisation de la non-réponse, en complément des informations disponibles dans la base de sondage ; 3) enrichir les réponses au questionnaire par des données de santé à des fins d'analyses épidémiologiques et biostatistiques croisant les deux sources d'information. Cet accès aux données du SNDS a fait l'objet d'un amendement à la lettre-avis envoyée aux participants, les informant spécifiquement de ce traitement et des modalités d'exercice de leurs droits.

Contributeurs Santé publique France:

- Myriam Blanchard (Dire, Normandie)
- Christine Chan Chee (DSET)
- Anne-Claire Colleville (DATA)
- Sandrine Coquet (DiRe, Nouvelle-Aquitaine)
- Pascal Empereur-Bissonnet (DiRe, St Maurice)
- Franck Golliot (DiRe, St Maurice)
- Imane Khireddine (DMNTT)
- Emmanuelle Le Lay (DESin)
- Mélanie Martel (DiRe, Normandie)
- Pauline Morel (Dire, Normandie)
- Yvon Mottref (DMNTT)
- Gaëlle Pedrono (DATA)
- Anne-Laure Perrine (DATA)
- Jean-Baptiste Richard (DATA)
- Abdessatar Saoudi (DATA)
- Morgane Trouillet (DiRe, Grand-Est)
- Abdelkrim Zeghnoun (DATA)

Sommaire

1	CONTEXTE DE L'ENQUETE	5
2	AVIS DU CONSEIL NATIONAL DE L'INFORMATION STATISTIQUE	6
3	SERVICE PRODUCTEUR ET SERVICE REALISANT LA COLLECTE : MAITRISE D'OUVRAGE ET MAITRISE D'ŒUVRE	6
4	JUSTIFICATION DE L'ETUDE DE SANTE DECLAREE (OU SANTE PERÇUE)	6
4.1	PROBLEMES DE SANTE QUI POURRAIENT RESULTER DE L'EXPOSITION A L'ACCIDENT INDUSTRIEL	6
4.2	INTERET D'UNE ETUDE DE SANTE PERÇUE (SANTE DECLAREE) POUR EVALUER L'IMPACT DE L'INCENDIE SUR LA SANTE	8
5	OBJECTIFS DE L'ENQUETE DE SANTE DECLAREE (OU SANTE PERÇUE) EN POPULATION	9
6	MATERIEL ET METHODES	10
6.1	COMITOLOGIE ET DIALOGUE AVEC LA SOCIETE	10
6.2	CONSULTATION DES EQUIPES DE LA DREES ET DE L'INSEE	11
6.3	CALENDRIER DE L'ENQUETE	11
6.4	EXPLOITATION, DIFFUSION DES RESULTATS, ET MISE A DISPOSITION DES DONNEES	12
6.5	METHODOLOGIE STATISTIQUE	13
6.6	DONNEES COLLECTEES PAR QUESTIONNAIRE	35
6.7	ENQUETE PILOTE	37
6.8	TRAITEMENT DE DONNEES A CARACTERE PERSONNEL ET PROTECTION DU SECRET STATISTIQUE	38
6.9	LETTRES-AVIS, NOTICES OU PLAQUETTES D'INFORMATION	46
7	COUT DES ENQUETES	50
8	LISTE DES ANNEXES	51
8.1	AVIS DU COMITE DU SECRET STATISTIQUE	51
8.2	AVIS D'EXAMEN DU COMITE DU LABEL	52
8.3	PRECISIONS SUR LE CALCUL DU NOMBRE DE SUJETS NECESSAIRES	56
8.4	LETTRE-AVIS	59

1 Contexte de l'enquête

Un incendie est survenu le 26 septembre 2019, sur les sites des entreprises Lubrizol et NL Logistique situées dans la zone portuaire de Rouen. Cet accident industriel a potentiellement exposé des populations à :

- des stimuli visuels et olfactifs, un stress psychologique ou un sentiment d'insécurité et d'incertitude liés à l'événement accidentel, ses conséquences et sa gestion les jours qui ont suivi l'incendie ;
- des substances chimiques propagées dans l'environnement, selon trois circonstances : l'émission dans l'air de substances issues de la combustion, de la dégradation ou de la volatilisation de produits stockés sur les sites ; l'apport sur le site de substances utilisées pour l'extinction de l'incendie, et de celles visant à masquer l'odeur de substances émises pendant la dépollution du site ; la dispersion sur une longue distance et la déposition dans l'environnement des composés du panache de fumées noires qui a résulté de l'incendie. A cela s'ajoute la dispersion aux alentours du site incendié de nombreux débris de toiture en fibrociment susceptibles de contenir des fibres d'amiante.

Ces expositions, dont la nature et le niveau varient selon la distance à l'incendie et au panache, peuvent avoir un impact à court, moyen et long termes sur la santé des différents groupes de population concernés : riverains des sites incendiés, populations présentes sous le panache, personnels de secours qui sont intervenus pour maîtriser l'événement et travailleurs d'autres institutions publiques ou privées.

Le 8 octobre 2019, le directeur général de la santé a saisi Santé publique France afin de proposer une méthode d'évaluation de l'impact de cet incendie sur la santé des groupes de population exposés¹.

Dans sa réponse en date du 17 octobre 2019², Santé publique France propose de mettre en œuvre, au-delà de la surveillance sanitaire réalisée au moment de l'incendie à partir des données de la médecine d'urgence (passages aux urgences hospitalières et consultations de l'association SOS-Médecins), et destinée à identifier les effets sanitaires immédiats ayant motivé des soins d'urgence, un dispositif global de surveillance des effets sanitaires potentiels de l'accident industriel. Avec pour objectif d'évaluer les conséquences sanitaires à court, moyen et long termes, le dispositif Santé Post Incendie 76 s'appuie sur quatre approches différentes et complémentaires :

- la réalisation d'une enquête de santé déclarée (ou santé perçue), menée auprès d'un échantillon représentatif de la population de Seine-Maritime exposée aux conséquences de l'incendie, dont l'objectif est de décrire les états de santé à court et moyen termes de l'événement, et d'analyser leur relation avec les expositions, nuisances et stimuli pouvant avoir été subis par la population pendant et dans les suites de l'incendie.
- la mise en place pour la population exposée d'un suivi épidémiologique à moyen et long termes de l'état de santé, au travers d'une analyse à intervalles réguliers d'indicateurs de santé, destiné à décrire l'évolution dans le temps de la fréquence des événements de santé pouvant être en lien avec les nuisances générées par l'incendie.

¹ Suivi sanitaire suite à l'incendie de l'usine Lubrizol en Seine Maritime. Saisine de santé publique France par la direction générale de la santé, 8 octobre 2019.

² Réponse à votre saisine du 8 octobre 2019 concernant la surveillance sanitaire et le suivi des populations exposées aux conséquences de l'incendie de l'usine Lubrizol en Seine-Maritime. Santé publique France, 17 octobre 2019.

- de proposer en appui aux services de médecine du travail une méthode pour le suivi longitudinal de la santé des salariés des deux entreprises et des professionnels intervenus sur site lors de l'incendie et des jours qui ont suivi pour le nettoyage, les prélèvements.
- de mener une étude d'imprégnation biologique humaine si les résultats des mesures de polluants dans l'environnement montrent un risque d'exposition de la population à certaines substances toxiques dispersées par l'incendie.

2 Avis du Conseil national de l'information statistique

Le projet d'étude de santé déclarée (ou santé perçue) en population suite à l'incendie des entreprises Lubrizol et NL Logistique a été présenté au Conseil national de l'information statistique (Cnis) lors de la séance du 28 novembre 2019 de la commission « Démographie et questions sociales ». Un avis d'opportunité favorable lui a été délivré le 9 décembre 2019³. Le projet a également été présenté au Comité du Label de la Statistique publique le 22 juin 2020, et a obtenu un avis d'examen favorable (Annexe 2).

3 Service producteur et service réalisant la collecte : maîtrise d'ouvrage et maîtrise d'œuvre

Santé publique France assure la maîtrise d'ouvrage de cette enquête.

La maîtrise d'œuvre est assurée par Santé publique France, en s'appuyant sur une comitologie décrite infra.

La constitution de l'échantillon est réalisée par l'Insee.

La collecte auprès des personnes est réalisée par l'institut de sondage Ipsos Observer, titulaire du marché pluriannuel (2017-2021) concernant les enquêtes en méthodes aléatoires pour Santé publique France. L'institut Ipsos Observer travaille sous la supervision de la maîtrise d'ouvrage.

4 Justification de l'étude de santé déclarée (ou santé perçue)

4.1 Problèmes de santé qui pourraient résulter de l'exposition à l'accident industriel

4.1.1 A court terme (pendant et dans le mois suivant l'accident)

4.1.1.1 Une symptomatologie irritative et digestive pouvant être évocatrice d'une intoxication

Pendant un mois à partir du jour de l'incendie, Santé publique France a mis en place une surveillance des recours aux urgences hospitalières et à SOS médecins liés à l'incendie et a colligé les informations issues d'autres sources de données : signalements de symptômes associés à la perception d'odeurs à Atmo Normandie et Atmo Hauts-de-France ; appels reçus par les Centres antipoison et de toxicovigilance (CAP-TV) ; appels de particuliers reçus par le SAMU de Rouen et des retours sur l'activité en médecine de ville. Les troubles de santé et symptômes identifiés par ces différentes sources étaient essentiellement respiratoires (gêne respiratoire et crise d'asthme), ORL et ophtalmologiques (irritation de la langue, de la gorge, des yeux), digestifs (nausées, vomissement), des céphalées et des vertiges.⁴

Les résultats d'une pré-enquête qualitative menée au mois de février 2020 par Santé publique France, sous forme de focus groupes et d'interviews individuels dans quatre communes exposées à l'incendie,

³ https://www.cnis.fr/wp-content/uploads/2019/01/CR_2019_2e_reunion_COM_DQS.pdf

⁴ Santé publique France. Incendie de l'usine de Lubrizol. Surveillance du recours à la médecine d'urgence. Le Point épidémiologique, 28 octobre 2019. 10 p : <https://www.santepubliquefrance.fr/content/download/197168/2363806>

décrivent également des effets de type irritatif dans les sphères ORL (langue, gorge, yeux) et respiratoire, certaines personnes prédisposées évoquant une activation des crises d'asthme⁵. Des symptômes digestifs sont également décrits, comme des nausées, des vomissements et des maux de ventre, en particulier chez les enfants.

De tels symptômes de toxicité aiguë ont été mis en lien avec l'exposition à des substances émises lors d'accidents industriels ou lors de rejets de substances odorantes^{6,7}.

4.1.1.2 Des troubles de la santé mentale

L'anxiété, les troubles du sommeil, des cauchemars et un stress psychologique important ont été décrits lors de la pré-enquête, notamment par les riverains proches de l'incendie, mais également par les femmes enceintes. Ces effets étant encore présents pour certains au mois de février lors de la réalisation de la pré-enquête, et réactivés lors des épisodes olfactifs récurrents dans les suites de l'incendie. L'exposition à un accident industriel et la pollution qui en résulte sont associées, sans que l'on en connaisse tous les mécanismes, au risque de troubles psychologiques et de certaines pathologies psychiatriques comme l'anxiété, la dépression, le stress post traumatique et les troubles du sommeil⁸.

Ces problèmes de santé, en lien avec des expositions à des facteurs de risque chimiques et psychologiques, ont été également décrits par une enquête réalisée en octobre 2019 par l'Union régionale des médecins libéraux (URML) de Normandie auprès des praticiens de la zone géographique concernée par l'incendie.

4.1.2 A moyen et long termes (1 an et plus après l'accident)

La survenue tardive d'effets sanitaires, plusieurs années après l'évènement nourrit une inquiétude importante des habitants. Des études se sont intéressées aux conséquences de l'exposition à une source de pollution, notamment sur la survenue des cancers, des maladies cardiovasculaires et respiratoires (notamment l'asthme) et des malformations congénitales. La surveillance de ces maladies rencontre plusieurs limites méthodologiques. Leur temps de latence peut être long ; l'évaluation des expositions est le plus souvent rétrospective et entachée d'incertitude ; il est rarement possible de conclure à une association entre l'exposition à la pollution et la survenue de ces maladies, du fait d'un manque de puissance statistique lié au faible effectif de population exposée et à la relative rareté des événements de santé considérés. La réponse à ces inquiétudes peut être apportée dans certains cas, de manière prédictive par une démarche d'évaluation quantitative des risques sanitaires (EQRS), ou par l'observation épidémiologique avec la mise en place d'un suivi à long terme de l'état de santé de la population sous la forme d'une cohorte.

Par ailleurs, les conséquences psychologiques, sur plusieurs années, d'un accident industriel sont documentées. Elles comprennent l'anxiété, la dépression et le stress post traumatique.

⁵ Colleville A-C. Analyse des préoccupations, attentes et savoirs mobilisés par les parties prenantes autour du site incendié de Lubrizol et NL logistique. Groupe Santé, Rouen2020. p. 1-14.

⁶ Ricoux C. Incendie de l'usine SBM Formulation à Béziers. Évaluation de l'impact sanitaire immédiat de l'exposition par voie respiratoire. Institut de veille sanitaire - Cellule interrégionale d'épidémiologie languedoc Roussillon, 2006.

⁷ Eskenazi B, Warner M, Brambilla P, Signorini S, Ames J, Mocarelli P. The Seveso accident: A look at 40years of health research and beyond. Environ Int. 2018;121(Pt 1):71-84.

⁸ Shieh V, Huang J-J, Wu T-EG, Chiu J-Y, Chen Y-C, Lin G, et al. Rate of psychiatric disorders and associations with quality of life among community members following the Kaohsiung gas explosion: An 18-month cross-sectional follow-up study. Health and Quality of Life Outcomes. 2019;17.

De manière plus générale, l'exposition à une pollution environnementale ou la proximité d'un site industriel peuvent avoir un effet sur la santé globale perçue^{9,10}. Le concept de santé globale ou générale perçue recouvre différents aspects de la santé, la santé physique, psychologique et sociale en s'ouvrant à des dimensions positives de la santé comme le bien-être et la qualité de vie.

4.2 Intérêt d'une étude de santé perçue (santé déclarée) pour évaluer l'impact de l'incendie sur la santé

La littérature scientifique¹¹ et plusieurs travaux de l'Institut de veille sanitaire^{12,13} ont montré l'intérêt des études de santé perçue et de qualité de vie dans des problématiques locales de pollution environnementale. Fondées sur la déclaration des personnes, elles permettent de recueillir des événements de santé qui intègrent la représentation de la population interrogée de sa santé et la perception qu'elle a de ses problèmes sanitaires. Elles renseignent ainsi des états de santé et de bien-être qui ne peuvent pas être mesurés par une autre méthode, et qui ne sont consignés dans aucune source disponible d'informations notamment s'ils ne font pas l'objet d'un recours au système de soins : symptômes, gênes, incapacités, altération du sommeil, troubles psychologiques, dégradation de la qualité de vie, état de santé global. Ce faisant, les études de santé perçue répondent à la demande sociale qui souhaite que soient prises en compte les préoccupations et les plaintes de santé de la population exposée au risque, sans limiter l'observation à des événements tels que les maladies graves ou les décès classiquement étudiés en épidémiologie.

Un autre critère de pertinence et de validité de la mesure de la santé perçue est sa corrélation avec d'autres indicateurs sanitaires considérés comme objectifs, voire sa capacité à en prédire l'évolution, tels que la mortalité¹⁴ et la morbidité¹⁵, la prévalence et l'incidence des maladies cardiovasculaires¹⁶ ou encore la consommation de soins¹⁷.

Enfin, dans ces contextes d'exposition d'une population à des facteurs de risque générés par une source locale de pollution, l'étude de la santé perçue s'avère des plus utiles car elle permet aussi de réaliser une évaluation de la qualité de l'environnement fondée sur l'expérience sensorielle et

⁹ Chatenoud L, Mosconi P, Malvezzi M, Colombo P, La Vecchia C, Apolone G. Impact of a major thermoelectric plant on self-perceived health status. *Prev Med*. 2005;41(1):328-33.

¹⁰ Radon K, Peters A, Praml G, Ehrenstein V, Schulze A, Hehl O, et al. Livestock odours and quality of life of neighbouring residents. *Ann Agric Environ Med*. 2004;11(1):59-62.

¹¹ Daniau C., Dor F., Eislstein D., Lefranc A., Empereur-Bissonnet P., Dab W. Etude de la santé déclarée par les personnes riveraines de sources locales de pollution environnementale : une revue de la littérature. Première partie : les indicateurs de santé déclarée. *Rev Epidemiol Sante Publique* 2013 ; 61(4) : 375-87.

¹² Martin S., Guillam M.T., Cassagne E., Legout C., Daniau C., Ségala C. Étude des perceptions des habitants de la Vallée-de-Seine (Yvelines) relatives à leur environnement et leur santé : phase quantitative. *Rev Epidemiol Sante Publique* 2016 ; 64S : S249

¹³ Daniau C., Wagner V., Salvio C. *et al.* État de santé perçue de la population riveraine d'une plateforme industrielle chimique : Salindres. *Environ Risque Santé* 2018 ; 17(6) : 583-95.

¹⁴ Idler E.L., Benyamini Y. Self-rated health and mortality: A review of twenty-seven community studies. *J Health Soc Behav* 1997 ; 38(1) : 21-37.

¹⁵ Molines C., Sapin C., Simeoni M.C., Gentile S., Auquier P. Perceived health and migration: a new sanitary approach? *Rev Epidemiol Sante Publique* 2000 ; 48(2) : 145-55.

¹⁶ Van der Linde R.M., Mavaddat N., Luben R. *et al.* Self-rated health and cardiovascular disease incidence: results from a longitudinal population based cohort in Norfolk, UK. *PLoS One* 2013 ; 8(6) : e65290.

¹⁷ DeSalvo K.B., Fan V.S., McDonnell M.B., Fihn S.D. Predicting mortality and health care utilization with a single question. *Health Serv Res* 2005 ; 40(4) : 1234-46.

cognitive des personnes concernées, et ainsi d'étudier les relations entre la perception de la pollution environnementale et la santé perçue¹⁸.

La surveillance des effets sanitaires aigus mise en place au moment de l'incendie et dans le mois suivant par Santé Publique France, n'intègre pas l'impact potentiel de l'évènement sur le recours au soin en médecine de ville ni l'impact à moyen terme sur la qualité de vie. L'enquête « Une étude à l'écoute de votre santé » vient compléter les connaissances des effets aigus et à moyen terme.

5 Objectifs de l'enquête de santé déclarée (ou santé perçue) en population

L'objectif principal de l'enquête est d'évaluer les impacts sanitaires à court et moyen termes de l'incendie, et consécutifs à l'ensemble des expositions, nuisances et stimuli pouvant avoir été subis par la population pendant et dans les suites de l'incendie.

Il s'agit de décrire l'éventuel impact de l'évènement en termes de symptômes ressentis, de pollution perçue et de qualité de vie, en intégrant les dimensions psychologiques et sociales pouvant résulter d'une situation environnementale dégradée et d'un sentiment d'insécurité et d'incertitude quant aux risques.

Seront étudiés les effets immédiats ou à court terme, c'est-à-dire au moment de l'incendie ou dans les jours suivants, qu'ils aient ou non donné lieu à une prise en charge médicale ; ainsi que les effets sur la qualité de vie et la santé telles que perçues au moment de l'enquête, soit environ un an après l'accident.

Pour les troubles de santé rapportés par la population lors de l'incendie et les jours suivants, il s'agira de :

- décrire la fréquence des symptômes et la perception des expositions
- étudier la relation entre les symptômes et les expositions perçues;

Pour la qualité de vie et la santé telles que perçues au moment de l'enquête, il s'agira de :

- décrire les indicateurs de santé perçue : santé globale (dimension physique et psychique), satisfaction de vie, santé mentale ;
- comparer la fréquence et le score moyen des indicateurs de santé perçue dans la population exposée et dans une population témoin
- étudier l'association entre les perceptions des expositions à l'incendie et les indicateurs de santé perçue.

Un objectif secondaire est de décrire les comportements adoptés suite à l'incendie, la connaissance des informations et des recommandations des pouvoirs publics pour identifier les points d'amélioration dans la prise en charge, l'information ou la compréhension des messages.

¹⁸ Daniau C., Dor F., Eislstein D., Lefranc A., Empereur-Bissonnet P., Dab W. Etude de la santé déclarée par les personnes riveraines de sources locales de pollution environnementale : une revue de la littérature. Seconde partie : analyse des résultats et perspectives. Rev Epidemiol Sante Publique 2013 ; 61(4) : 388-98.

6 Matériel et méthodes

6.1 Comitologie et dialogue avec la société

6.1.1 Le Comité d'appui thématique, appui scientifique aux études menées

Sur le plan scientifique, le projet s'appuie sur un comité d'appui thématique indépendant intitulé : « **Evaluation de l'impact de l'incendie des entreprises Lubrizol et NL Logistique sur la santé de la population** », constitué de personnes qualifiées, externes à Santé publique France. Il s'inscrit dans la comitologie de Santé publique France.

Le comité d'appui thématique apporte aux équipes de Santé publique France un soutien scientifique dans des domaines d'expertise spécifiques au projet : toxicologie, métrologie, exposition professionnelle, sociologie des crises, épidémiologie, santé mentale, mesure psychométrique de la santé. Il a pour rôle d'aider aux choix scientifiques, à l'interprétation des résultats et à la formulation de propositions et actions de santé publique issues des travaux menés. Il vient en appui des équipes de Santé publique France.

6.1.2 Instances de dialogue avec la société

Les travaux sont menés dans le cadre d'une démarche participative, en lien avec un « Groupe Santé » issu du Comité de transparence et de dialogue (CTD) mis en place localement par le préfet de Seine-Maritime suite à l'incendie.

6.1.2.1 Le Groupe « Santé », une approche participative au niveau local

L'expérience que Santé publique France a acquise dans les études en population générale montre que l'approche participative devient de plus en plus nécessaire dans les questionnements de santé publique et notamment lorsque ceux-ci concernent l'environnement. Ainsi, des représentants, à divers titres, de la population exposée, ont été associés dès la phase de conception du dispositif d'études. Cette démarche permet de répondre aux attentes, d'améliorer la qualité des études réalisées, en faisant appel à l'expérience de la population locale et de faire comprendre ce qu'est une enquête scientifique, ses contraintes, mais surtout de pouvoir prendre en compte dans le dispositif d'enquête des aspects ou des questions qui pourraient être délaissés. Cette méthode de travail doit, dans cette logique, nous assurer de répondre au mieux à la question de santé publique et faciliter la compréhension et l'appropriation des résultats des travaux qui sont menés. Un groupe de représentants de la population issu des participants au CTD, nommé Groupe Santé (GS) est l'interlocuteur de Santé publique France pour les enquêtes en population générale (enquête de santé déclarée et suivi à moyen et long terme).

La coordination de cette approche participative est assurée par une personne extérieure à Santé publique France, **Marcel Calvez, professeur de sociologie à l'Université de Rennes**, qui a mené des recherches sur les relations scientifiques-citoyens engageant des questions de santé environnementale^{19,20,21}.

¹⁹ Calvez M. Qu'est-ce qu'une expertise citoyenne en santé environnementale ? Questions sur l'expérimentation rennaise *Ambassad'Air. Environnement, Risques & Santé*. 2018;17(5):498-504.

²⁰ Calvez M. Les risques de santé à l'épreuve des nuisances. Le cas des plaintes en santé environnementale. *Recherches sociologiques et anthropologiques*. 2016;47(1):187-203.

²¹ Calvez M. Les signalements profanes de clusters : épidémiologie populaire et expertise en santé environnementale. *Sciences sociales et santé*. 2009;27(2):79-106.

Le GS a été constitué par un appel à volontaires diffusé au mois de décembre 2019 au sein du CTD, aboutissant à constituer un groupe de 19 personnes dont les institutions sont listées :

- Associations : Rouen Respire, France asso santé – LCC 76 et Réseau environnement santé, UFC Que Choisir-Rouen ;
- Représentants de professionnels de santé : Ordre des médecins, URML, pathologies professionnelles – CHU Rouen, pneumologie – CHU Rouen, Ordre des infirmiers
- Elus et collectivités territoriales : maires de Canteleu, Petit Quevilly, La rue Saint Pierre, Rouen métropole Normandie, la Région Normandie- site de Caen, députée de la 2è circonscription (Est et Nord) et de la 10è circonscription (Nord) ;
- Université de Rouen (un anthropologue).

Le Groupe Santé se réunit à un rythme mensuel depuis janvier 2020. Une équipe projet de Santé publique France participe aux réunions du groupe, dont la coordination est assurée par Marcel Calvez (université Rennes 2).

Un second groupe a été constitué à partir des candidatures reçues représentant des acteurs pouvant apporter un éclairage au Groupe santé par leurs auditions, comme par exemple et sans que cela ne soit exhaustif : ARS Haute Normandie, Atmo Normandie, service de santé au travail des salariés et directions des entreprises concernées, etc.

6.2 Consultation des équipes de la DREES et de l'Insee

Des réunions d'échanges portant sur la méthode de l'enquête, concernant le périmètre de la zone d'étude, sur l'échantillonnage à mettre en œuvre ainsi que sur les méthodes statistiques pour prendre en compte l'effet du multi-mode ou le biais de sélection ont eu lieu avec le bureau Etat de santé de la population de la Drees et avec la Division sondages de l'Insee.

6.3 Calendrier de l'enquête

Les principales étapes sont décrites ci-dessous :

- **28 novembre 2019** : présentation du projet d'enquête au Conseil national de l'information statistique (CNIS) lors de la séance du 28 novembre 2019 de la commission « Démographie et questions sociales ».
- **9 décembre 2019** : réception d'un avis d'opportunité favorable du CNIS ²².
- **Février 2020** : réalisation d'une pré-enquête qualitative ayant pour objectif de recueillir, par entretiens, les préoccupations en termes de santé et les attentes de personnes habitant dans les communes impactées par l'incendie de l'usine Lubrizol et des entrepôts de NL Logistique, ainsi qu'un premier balayage des conséquences sur la santé et des symptômes qui s'y rattachent (bruit de l'explosion, incendie, nuage de fumée, différents types d'odeurs, suies, etc.) afin de préparer le questionnaire de l'étude. Quatre focus groupes, et 9 entretiens individuels ont été réalisés auprès de professionnels de santé et des maires des 4 communes : Petit Quevilly, Bois Guillaume, Préaux, Buchy ainsi que le centre social Relais accueil des Gens du voyage.
- **Février à mai 2020** : Les travaux menés en lien avec le Groupe santé ont eu pour objectif de finaliser le protocole de l'étude de santé déclarée en population : la finalisation du questionnaire et la rédaction du protocole, la définition de la zone d'étude. Santé publique France a par ailleurs conçu le plan de communication et les outils afférents (affiches, dépliants). Ces points ont été finalisés lors de la réunion du Groupe santé du 18 mai.
- **Mars 2020** : soumission en ligne du dossier au Comité du secret statistique.

²² https://www.cnis.fr/wp-content/uploads/2019/01/CR_2019_2e_reunion_COM_DQS.pdf

- **29 avril 2020** : dépôt du dossier au secrétariat du Comité du label de la statistique publique.
- **18 juin au 8 juillet 2020** : réalisation d'une enquête pilote (test du questionnaire sur une centaine de personnes de communes limitrophes de la zone d'étude)
- **11 juin 2020** : présentation du protocole au comité du Label de la statistique publique
- **17 juillet 2020** : réception du comité du Label de l'avis d'examen définitif relatif à l'enquête
- **24 août 2020** : signature de la convention tripartite Santé publique France/Insee/DgFip
- **24 août 2020** : tirage de l'échantillon et livraison par l'Insee de l'échantillon
- **Fin août 2020** : communication annonçant l'enquête
- **Septembre-octobre 2020** : Phase de collecte des données (questionnaire principal)
- **Novembre-décembre 2020** : collecte des données du questionnaire santé mentale
- **2-3 décembre 2020** : passation du questionnaire aux Gens du voyage de l'aire d'accueil du Petit-Quevilly
- **Novembre-décembre 2020** : traitement aval (apurement, traitement de la non réponse et calcul des pondérations) et production des fichiers d'analyse et de données sur la participation
- **Début 2021** : diffusion des 1ers résultats
- **Mi 2021** : résultats complets

6.4 Exploitation, diffusion des résultats, et mise à disposition des données

La stratégie de valorisation des résultats de l'étude sera définie par le Comité éditorial (Comed) de Santé publique France. Compte tenu de l'organisation du projet, qui met en œuvre une approche participative menée localement, la séquence ci-dessous sera suivie :

- **En amont de la valorisation, partage de l'interprétation des résultats avec notamment les parties prenantes :**
 - Discussion des résultats en inter directions de Santé publique France (Direction scientifique, Direction des régions, Direction appui, traitement et analyse, Direction santé environnement, Direction des maladies non transmissibles) : première interprétation des résultats.
 - Présentation des résultats aux comités mis en place par Santé publique France pour cette étude :
 - Groupe Santé à Rouen : discuter des résultats, de leur signification et de leur apport aux questions posées localement.
 - Comité d'appui thématique mis en place par l'agence : consolider l'interprétation et dégager les enseignements et conclusions principales.
- **Présentation des résultats aux instances locales et nationales**
 - Présentation des résultats aux instances locales (ARS Normandie, Préfet de Seine Maritime, DREAL), à la Direction générale de la santé et à l'Anses.
 - Présentation des résultats lors d'une réunion à Rouen du Comité de transparence et de dialogue.
 - Information des maires et professionnels de santé de la zone d'étude et de la zone témoin
- **Diffusion des résultats auprès des personnes enquêtées et de la population**
 - La tenue de réunions publiques sera probablement difficile du fait du contexte COVID-19. Il faudra opter pour une information *via* les médias et réseaux sociaux, s'appuyant sur une vidéo pédagogique, qui accompagnera la mise en ligne du rapport et de sa synthèse.

- **Valorisation académique des résultats d'étude**
 - Articles scientifiques
 - Participation à des colloques
 - Partage des données avec des équipes de recherche universitaires (*open data*) et/ou projets de thèse de sciences

- **Conditions d'accès aux données**

Le partage des données sera réalisé selon les dispositions de mises à disposition des données actuellement utilisées à Santé publique France. L'ensemble des données pseudonymisées de la base consolidée sera accessible à des partenaires extérieurs. Une procédure d'accès à ces données a été mise en place. Tout tiers à l'agence qui souhaite accéder à des données scientifiques à des fins de recherche saisit l'agence par l'envoi d'un dossier de demande d'accès contenant : un formulaire de demande, le protocole de recherche, la justification de la réalisation d'une formalité adéquate pour les demandes portant sur des données à caractère personnel. La mise à disposition de données à caractère personnel est subordonnée à l'information préalable des personnes concernées par Santé publique France ou le demandeur, ou la justification d'une autorisation CNIL permettant la dérogation à l'information individuelle. Après instruction de la demande et en cas de validation, une licence d'utilisation est adressée au demandeur. Santé publique France s'engage à tenir un registre des données mises à disposition, et à veiller au respect des tiers des obligations souscrites notamment de destruction de données à l'issue de leur exploitation.

S'agissant d'une enquête de la statistique publique, un accès aux données collectées le plus large possible sera étudié, compatible avec les règles de protection des données :

- Information auprès des chercheurs de toutes disciplines, notamment ceux des sciences sociales, sur l'existence de l'enquête et les procédures d'accès aux données.
- Liens renvoyant sur le site de l'enquête sur des dispositifs généralistes comme le réseau Quetelet, ou le CASD, et mise à disposition des fichiers – éventuellement après une anonymisation complémentaire – sur ces centres de données.

6.5 Méthodologie statistique

6.5.1 Type d'étude

Cette étude est une enquête transversale de santé déclarée (ou santé perçue), réalisée sur un échantillon représentatif de la population exposée à l'incendie. L'appellation « santé déclarée ou santé perçue » est utilisée ici pour désigner l'ensemble des altérations de la santé recueillies à partir d'un questionnaire rempli directement par les individus ou par l'intermédiaire d'un enquêteur qui n'a pas pour mission de valider ni d'expertiser médicalement les déclarations des individus. Elle fait référence au mode de recueil des données de santé, en interrogeant la personne, qui intègre sa représentation personnelle de la santé²³. Ce type d'enquête permet de collecter des événements de santé dont l'existence ne peut être attestée que par la personne qui les ressent, et qui ne sont consignés dans aucune source d'information sanitaire a fortiori s'ils n'ont pas fait l'objet d'un recours au système de soin (Base médico administrative, registres, etc.). Dans la présente étude, le terme « santé déclarée ou santé perçue » comprend les événements de santé que les sujets déclarent, c'est-à-dire des plaintes, des symptômes, des maladies chroniques, des expositions environnementales, et la perception que les

²³ Bruchon-schweitzer M BE. Psychologie de la santé : concepts, méthodes et modèles. 2e éd ed2014 2014. 544p.

sujets ont de leur santé physique, psychologique et globale ainsi que leur satisfaction de vie. Ce type d'étude est donc particulièrement adaptée à une approche des situations d'exposition à des sources locales de pollution, car en couvrant à la fois le bien-être physique, psychologique et social, elle répond à un besoin d'approche globale de la santé que révèle la diversité des plaintes sanitaires exprimées par la population ^{11,18}.

6.5.2 Champ géographique : Définition de la zone exposée à l'incendie et de la zone témoin

6.5.2.1 Nuisances liées à l'incendie

L'incendie s'est déclaré le 26 septembre 2019 aux alentours de 2h40. Il a été contenu, puis maîtrisé le même jour à 15h00. Un feu couvant lié à des futs endommagés a perduré jusqu'au 7 octobre sur le site de l'entreprise NL Logistique.

L'incendie a provoqué un important panache de fumées noires, très dense et visible, qui s'est élevé à forte hauteur au droit du site et a ensuite poursuivi une trajectoire Nord-Est impactant les régions Normandie et Hauts-de-France. La colonne de fumées étant montée verticalement et assez haut (100-150 m), elle est passée au-dessus de la zone la plus proche (Rouen centre) au cours de la journée du 26 septembre. En revanche, le plateau nord-est de la ville a été plus directement touché par le panache, rabattu au sol par des pluies intermittentes. De nombreuses retombées de suie, visibles à l'œil nu, ont été constatées sous l'aire de dispersion du panache, dans une vaste zone au nord-est de Rouen. Le panache s'est dissipé le lendemain de l'incendie, le 27 septembre. Au total, il est estimé à 215 le nombre de communes impactées par les fumées noires ou les dépôts dont 112 en Seine-Maritime (111 après fusion de 2 communes).

Des odeurs incommodantes et persistantes (de soufre, de brûlé et d'hydrocarbures au début puis uniquement hydrocarbures) ont été perçues depuis le jour de l'incendie et l'ont été de façon intermittente jusqu'en septembre 2020. Elles ont été très fortes le mois qui a suivi l'incendie et ont été associées à des effets sanitaires tels que nausée, maux de tête et vomissement. Ce panache non visible mais constitué de molécules très odorantes avait pour origine les futs et résidus d'hydrocarbures présents sur le site incendié. La zone d'impact des émissions de substances odorantes couvre une grande partie de la Métropole de Rouen. Ces communes, qui ne sont pas forcément les communes touchées par les dépôts de suies, sont identifiées à partir des signalements sur l'application ODO mis en place par l'association agréée pour la surveillance de la qualité de l'air (Atmo Normandie)²⁴.

La population résidant à proximité de l'incendie a pu être exposée à d'autres types de nuisances spécifiques : auditive (bruits de l'explosion), visuelle (témoins des flammes de l'incendie), auxquelles s'est ajoutée une exposition possible d'ordre psychologique, caractérisée par un sentiment direct de menaces de propagation de l'incendie, voire d'explosion de l'usine, une peur de mourir. Au sein de cette population de proximité, les personnes résidant à moins de 500 mètres de l'usine ont en outre été placées en zone de mise à l'abri et invitées à rester chez elles. Cette population, située dans la commune de Petit-Quevilly, a été exposée encore plus intensivement à l'ensemble des nuisances, et a pu notamment ressentir une menace pour ses biens, son intégrité physique, voire une angoisse de mort (cf. résultats de la pré-enquête).

Par ailleurs, l'incendie et les explosions qu'il a provoqué ont détruit et dispersé la toiture en fibrociment d'un des bâtiments industriels, et de nombreuses retombées de débris ont été constatées dans les espaces publics et privés jusqu'à plusieurs kilomètres du site.

²⁴ <http://www.atmonormandie.fr/Lubrizon-NL-Logistique/Introduction>

6.5.2.2 Zone d'impact définie pour la surveillance environnementale

Pour décider du plan de surveillance environnementale à mettre en place, une zone d'impact a été rapidement délimitée par la cellule post-accident de la préfecture de Seine-Maritime, à partir de deux sources d'informations^{25,26} :

- Une simulation de la dispersion atmosphérique du panache et des dépôts de suies émises par l'incendie qui a été réalisée par l'Institut national de l'environnement industriel et des risques (Ineris) à partir du modèle mathématique à grande échelle CHIMERE, afin d'estimer l'étendue possible des retombées. Elle a conduit à définir l'étendue maximale des dépôts permettant ainsi de guider le choix de la zone dans laquelle effectuer des prélèvements. Cette modélisation a été réalisée rapidement après l'évènement et n'a pas été affinée depuis.
- En complément de cette modélisation, un constat visuel d'étendue du panache de fumées noires et des retombées de suies a été effectuée les 26 et 27 septembre par le Service départemental d'incendie et de secours (SDIS), la Direction départementale des territoires et de la mer (DDTM) et les gendarmes. Les données de ces constats ont permis de définir les communes dans lesquelles effectuer des prélèvements environnementaux. Ces observations ont montré que les dépôts détectables à l'œil nu n'étaient pas homogènes et ont pu être certaines fois très localisés (en « tâche de léopard »), ne permettant pas de graduer l'exposition à partir de la seule distance au foyer de l'incendie.

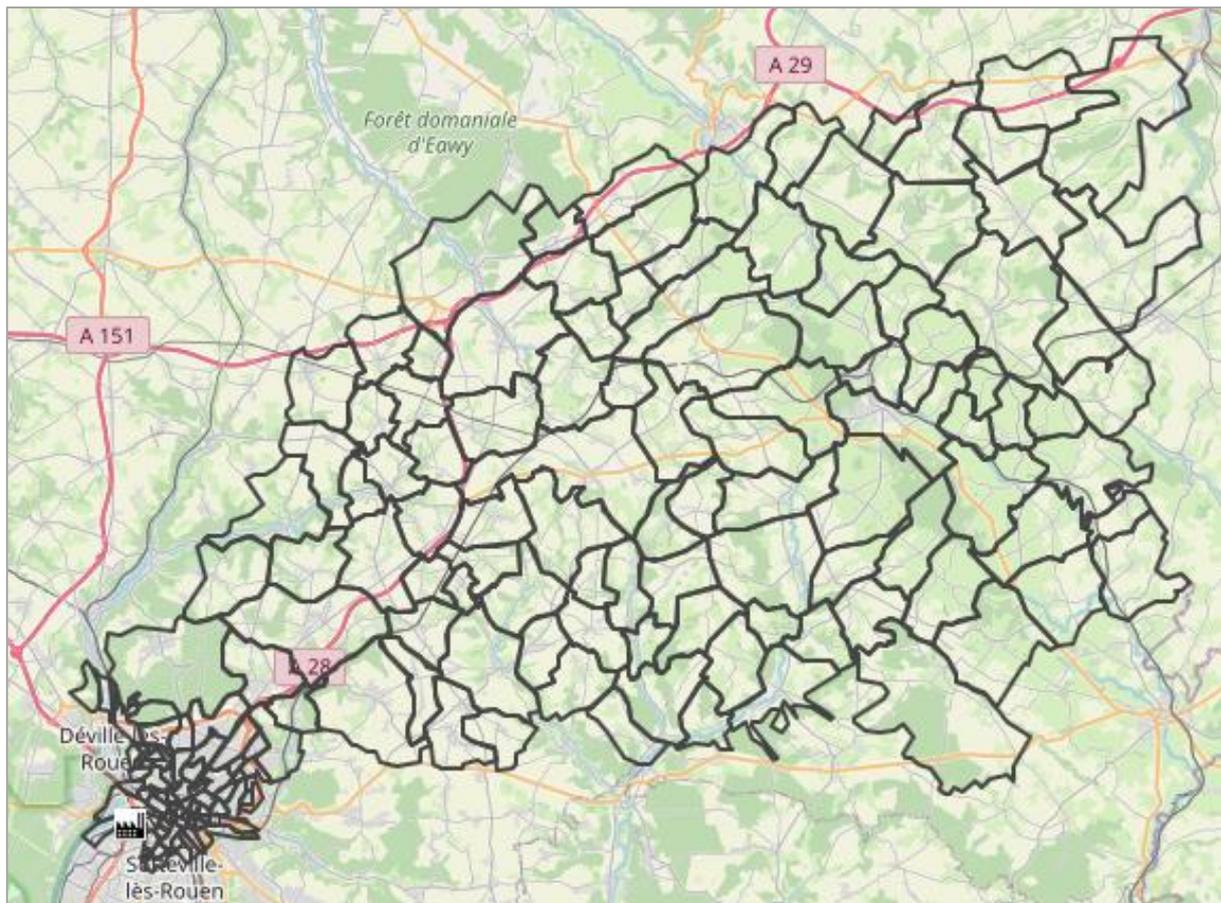
Le croisement de la modélisation de l'Ineris et des constats visuels a permis au préfet de définir les communes pour lesquelles une surveillance de l'impact environnemental devait être réalisée, soit 112 communes en Seine-Maritime (Figure 1).

²⁵ Arrêté du 14 octobre 2019, imposant à la société LUBRIZOL France (SIREN 542 070 958) des prescriptions de mesures d'urgence pour son site situé dans les communes de Rouen et de Petit-Quevilly. Direction régionale de l'environnement, de l'aménagement et du logement de Normandie.

²⁶ Arrêté du 14 octobre 2019, imposant à la société NL Logistique (SIREN 570 501 791) des prescriptions de mesures d'urgence pour son site situé dans les communes de Rouen. Direction régionale de l'environnement, de l'aménagement et du logement de Normandie.

La zone d'impact ayant servi à identifier les communes à inclure dans le plan de surveillance environnementale ne considère que l'exposition au panache de fumées noires et à ses retombées. Cette première zone a donc pu servir de base à la définition de la zone d'étude, elle a dû toutefois être complétée des communes qui n'ont pas été situées sous le panache, mais qui ont été exposées aux autres nuisances.

Figure 1 : 112 communes de Seine-Maritime concernées par l'arrêté préfectoral, exposées au panache de fumées et à ses retombées

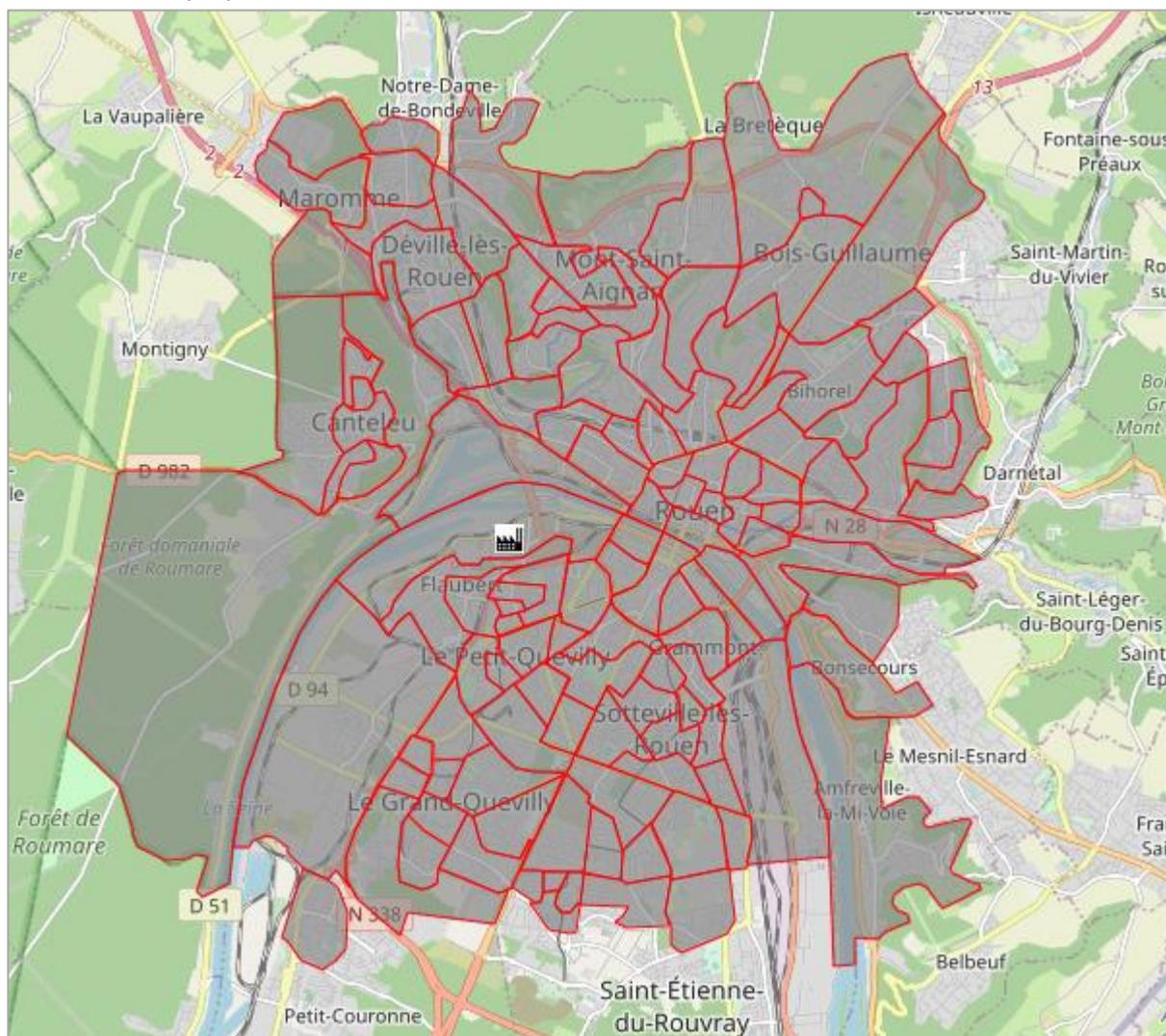


6.5.2.3 Zone concernée par les nuisances odorantes

Les données collectées par l'application ODO d'Atmo Normandie, issues de l'enregistrement du nombre de symptômes rapportés dans le cadre des signalements d'odeurs désagréables, montrent que des communes non exposées au panache de fumées noires ou à ses retombées, ont été exposées à des substances chimiques générant la perception d'odeurs très gênantes, associées à des maux de tête, vertiges, divers symptômes d'irritations de la sphère ORL et oculaire, vomissements, nausées et gênes respiratoires. Ce constat montre l'intérêt d'étendre la zone d'étude à ces communes.

Afin de cerner au mieux le périmètre de la zone exposée à des molécules odorantes, une extraction puis une cartographie de l'ensemble des signalements d'odeurs relevés par ATMO Normandie pendant une durée de 2 mois après l'incendie ont été réalisées. Ce travail a permis de définir la liste des IRIS compris dans la zone concernée par ces signalements d'odeurs (Figure 2).

Figure 2 : IRIS concernés par des signalements d'odeurs²⁷, pendant la période de 2 mois ayant suivi l'incendie du 26/09/2019.



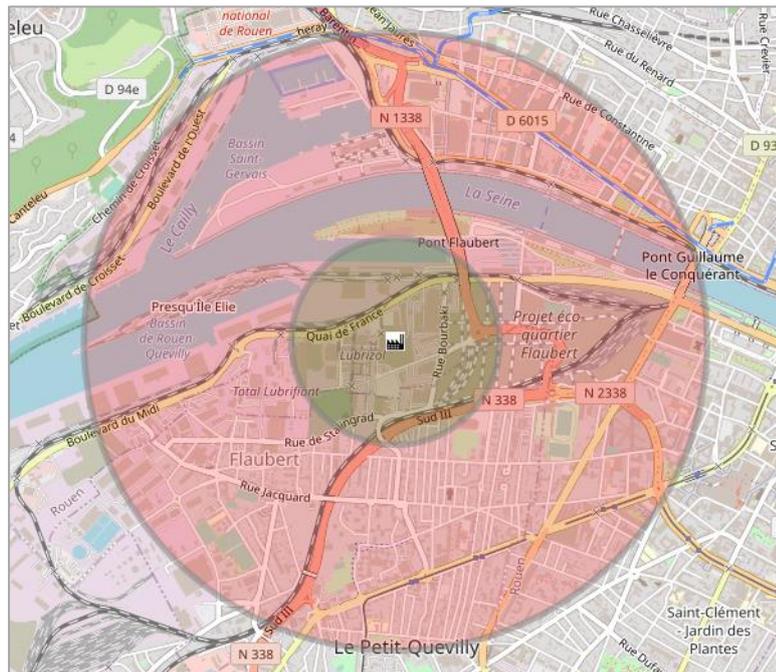
6.5.2.4 Zone de proximité

La définition de la distance reflétant au mieux la proximité au lieu de l'incendie a été discutée avec le Groupe santé et lors de la pré-enquête. Un consensus a été trouvé sur les zones suivantes :

- Une zone de proximité immédiate du lieu de l'incendie, située dans un rayon de 700 mètres autour de l'incendie, c'est-à-dire dans une partie de la commune du Petit-Quevilly (Figure 2.3). Cette zone permet de pouvoir inclure dans l'échantillon de l'enquête suffisamment de personnes pour analyser leurs expositions spécifiques et les éventuelles conséquences de l'accident sur leur santé ;
- Une deuxième zone de proximité située entre un rayon de 700 mètres et 1500 mètres autour de l'incendie (Figure 3), y compris de l'autre côté de la Seine ;

²⁷ Certains IRIS, dans lesquels seuls 1 ou 2 signalements avaient été recueillis, n'ont pas été pris en compte dans cette cartographie.

Figure 3 : Zones de proximité, rayon de 700 mètres (en vert) et de 1500m (en rouge) autour du lieu de l'incendie.



6.5.2.5 Choix d'une zone témoin

Compte tenu des objectifs de l'étude, visant à évaluer les expositions aux différentes nuisances de l'incendie et leurs conséquences sur la santé de la population, l'ajout à l'étude d'une zone témoin s'est avéré nécessaire afin de disposer de résultats dans une population de référence non exposée à l'accident industriel. Au-delà de son intérêt pour permettre d'évaluer l'impact de l'incendie sur la santé, la prise en compte de cette zone témoin doit permettre de mesurer l'importance des conséquences sur plusieurs indicateurs de santé de la période de confinement lié à l'épidémie COVID-19, débutée le 17 mars, commune aux populations de la zone d'étude et de la zone témoin. Cette zone témoin a été sélectionnée de façon à fournir le contrefactuel le plus satisfaisant pour l'évaluation des résultats concernant la zone d'étude.

Le choix de cette zone témoin a dû tenir compte de l'existence d'une exposition historique de la population de l'agglomération rouennaise à une pollution d'origine industrielle. Elle est définie :

- par une localisation géographique comprenant une exposition d'origine industrielle, de nature équivalente à celle de la zone d'étude ;
- par une localisation géographique suffisamment éloignée de Rouen pour limiter l'inclusion de personnes concernées par une exposition aux nuisances de l'incendie liée à leur activité professionnelle, et assez proche de Rouen pour en partager le climat, la culture et les modes de vie ;
- par des caractéristiques socio-économiques et démographiques comparables à celle de la zone d'étude, notamment du point de vue de la taille et de la densité de population (rural/urbaine) ;

Parmi les zones candidates évoquées lors de différents échanges avec le Groupe Santé et le Comité d'appui thématique, les agglomérations de Caen et du Havre se sont révélées pertinentes. Cette dernière présente l'avantage de la présence d'une activité industrielle historique comparable à celle

de l'agglomération rouennaise. Par ailleurs, l'analyse des caractéristiques socioéconomiques des deux zones candidates souligne une similitude plus importante entre la zone entourant Le Havre et la zone d'étude : indicateurs économiques, distribution de population selon la taille d'unité urbaine, indice FDEP (Tableau 1). La zone définie autour de Caen s'avère moins urbaine d'une part, d'un niveau socioéconomique plus élevé d'autre part.

Tableau 1. Comparaison d'indicateurs socioéconomiques entre la population d'étude et deux zones témoins candidates de taille de population identique : Le Havre et ses environs ; Caen et ses environs.

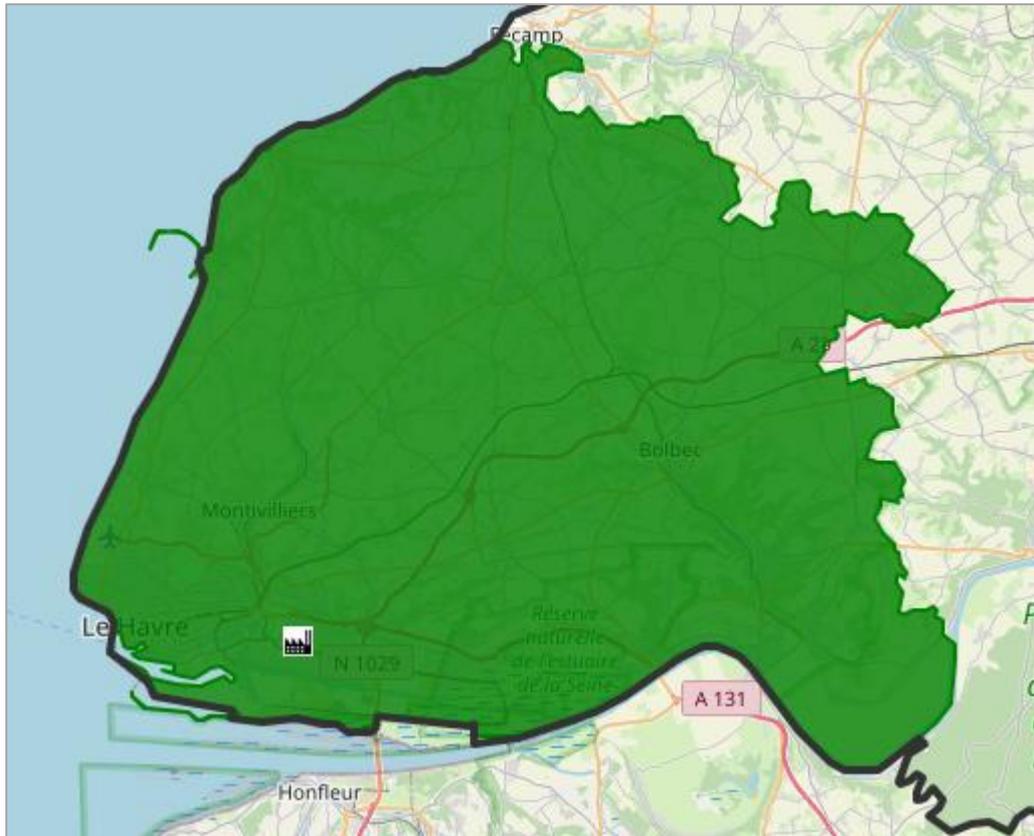
	Témoin - Le		
	Etude	Havre	Témoin - Caen
Effectifs	360 000	360 000	375 000
% chômeurs	11,8%	12,0%	9,6%
Revenus médian	20521	20133	21310
Part des ménages fiscaux imposés	50	48	52
Taux de pauvreté (M)	19,0	18,4	15,9
TUU			
Commune rurale	12,3%	21,4%	26,3%
UU < 5 000 hab.	5,3%	7,1%	22,3%
UU < 20 000 hab.	27,9%	25,7%	17,0%
UU >= 20 000 hab.	55,3%	46,0%	34,3%
FDEP - IRIS*			
Q1 (++)	31,2%	18,8%	32,9%
Q2 (+)	12,9%	16,6%	24,0%
Q3	14,2%	14,0%	24,6%
Q4 (-)	15,3%	14,8%	7,2%
Q5 (--)	26,3%	35,7%	11,3%

* construit à partir du taux de bacheliers, du taux de chômage et du taux d'ouvriers

Source : statistiques locales Insee ; indice FDEP 2015 disponible au niveau de l'IRIS.

Aussi, la zone témoin proposée inclut les communes de Seine-Maritime situées dans un périmètre de 29 km autour de la zone industrielle du Havre. Cette population sera comparable à celle de la zone d'étude sur plusieurs critères : taille (~360 000 habitants), diversité de répartition entre zone rurale et zone urbaine, niveau socioéconomique, historique et nature d'exposition à des activités industrielles (Figure 4).

Figure 4 : zone témoin, communes situées dans un rayon de 29 km autour de la zone industrielle du Havre.



6.5.2.6 Caractérisation socio-économique de strates

Le tableau ci-dessous présente quelques indicateurs socio-économiques selon les différentes strates de la zone exposée à l'incendie, et entre la zone exposée et la zone témoin.

Une des strates de la zone exposée se démarque nettement. En effet, la strate « odeurs et panache », qui est située au nord-est de l'incendie et au-delà de 1 500m, correspond aux quartiers les plus favorisés de Rouen et ses alentours. La population y est plus diplômée, et comporte notamment nettement moins d'ouvriers que dans les autres zones. La population de la zone témoin comporte plus d'ouvriers, est moins diplômée, et plus défavorisé selon l'indicateur FDEP que la zone exposée dans son ensemble. Le taux de chômeurs dans les deux populations est très comparable.

Tableau 2. Comparaison d'indicateurs socio-économique selon les différentes strates

	Population	% population dans le quintile le plus défavorisé (FDEP)	Part de population (15+) avec un diplôme < Bac	% Chômeurs (15-64 ans)	% Ouvriers parmi les Actifs occupés (15-64 ans)
Strates de la zone d'étude					
Odeurs + panache	141415	13,4%	40,4%	11,3%	10,7%
Odeurs seules	124164	41,3%	63,5%	13,6%	24,2%
Panache seul (suiés)	72809	23,1%	58,5%	8,1%	22,4%
Proximité*	22116	35,6%	58,6%	15,6%	23,8%
Zone exposée vs zone Témoin					
Zone exposée (ensemble)	360505	26,3%	53,6%	11,7%	18,5%
Zone Témoin	359252	35,7%	61,4%	11,9%	25,2%

*Statistiques calculées parmi les IRIS contenant le périmètre de 1200 m autour de l'incendie ; la taille de population est similaire à celle résidant dans la zone de proximité de 1500m ;

6.5.3 Unités enquêtées

Le champ de l'enquête est défini par l'ensemble des adultes âgés de 18 ans et plus et des enfants âgés de moins de 18 ans le jour de l'incendie des usines Lubrizol et NL Logistique, le 26 septembre 2019, vivant dans un logement ordinaire situé dans la zone d'étude au moment de l'incendie, quel que soit le lieu de résidence actuel.

La population effectivement interrogée est celle des adultes âgés de 18 ans et plus au moment de l'incendie. Les enfants de moins de 18 ans seront interrogés indirectement à travers le questionnaire de leurs parents.

L'échantillon de l'enquête sera issu d'un sondage aléatoire simple stratifié réalisé par l'Insee dans la table d'individus du millésime 2019 du Fichier démographique d'origine fiscale sur les logements et les individus (Fidéli) conçu à partir des données de la Direction générale des finances publiques (DGFiP).

Critère d'inclusion pour l'échantillon exposé

- Etre âgé de 18 ans ou plus le jour de l'incendie pour les adultes
- Etre âgé de moins de 18 ans le jour de l'incendie pour les enfants
- Résider dans un ménage ordinaire dans la zone exposée le jour de l'incendie (les personnes ayant déménagé après l'incendie en dehors de la zone d'étude sont incluses dans l'étude (contact par numéro de téléphone / adresse mail ou suivi de courrier)
- Etre en mesure de comprendre, par téléphone, les implications de la participation à l'étude (audition, compréhension de la langue française).

Critère d'exclusion pour l'échantillon exposé

- Avoir emménagé dans la zone d'étude après l'incendie
- Avoir résidé dans la zone d'étude à titre occasionnel le jour de l'incendie (pour les vacances, de façon intermittente, pour des raisons professionnelles, etc.)
- Avoir un problème de santé grave rendant impossible l'entretien

Critère d'inclusion pour l'échantillon témoin

- Etre âgé de 18 ans ou plus le jour de l'incendie pour les adultes

- Résider dans un ménage ordinaire dans la zone témoin le jour de l'incendie
- Etre en mesure de comprendre, par téléphone, les implications de la participation à l'étude (audition, compréhension de la langue française).

Critère d'exclusion pour l'échantillon témoin

- Avoir travaillé dans la zone exposée le jour de l'incendie
- Avoir résidé dans la zone exposée à titre occasionnel le jour de l'incendie (pour les vacances, de façon intermittente, pour des raisons professionnelles, etc.)
- Avoir un problème de santé grave rendant impossible l'entretien

6.5.4 Paramètres d'intérêt de l'enquête

L'objet de cette enquête est de décrire précisément l'impact de l'événement :

- Exposition aux facteurs de risque : stimuli visuels, bruits, odeurs, retombées de débris de fibrociment, exposition directe au panache ou aux retombées de suie, attitudes adoptées au moment de l'incendie ;
- Symptômes (irritations, maux de tête, nausée...) ressentis dans les suites de l'incendie ;
- Santé globale actuelle, mesurées par différentes échelles (cf. partie 5)
- Comportements et perceptions de la population sur l'information disponible sur l'incendie et concernant les recommandations émises par les pouvoirs publics.

Plus précisément, les données recueillies viseront à comparer la fréquence des symptômes et de la pollution perçue entre les différentes zones d'exposition ; à étudier la relation entre la fréquence des symptômes et la pollution perçue et si possible l'intensité de l'exposition déclarée.

Il s'agira également de 1) décrire la santé et comparer cet état de santé entre les différentes zones d'exposition, avec la zone témoin et dans la mesure du possible à des références nationales ; 2) d'étudier la relation entre l'état de santé perçue et l'exposition aux facteurs de risques rapportée.

Dans la mesure où l'évaluation de la santé perçue est fortement liée aux caractéristiques socioéconomiques des populations, les résultats seront analysés en tenant compte de nombreuses variables disponibles dans la base de sondage ou recueillies dans l'enquête : sexe, âge, situation financière, type de ménage, situation matrimoniale, nombre d'enfants, commune de résidence, nationalité, situation par rapport à l'emploi, profession, profession des parents pour les élèves et étudiants, revenus, diplôme le plus élevé obtenu.

Dans la zone témoin seront recueillies, à titre de données de référence, les données concernant la survenue de symptômes sur une période des quatre dernières semaines, ainsi que l'ensemble des questions relatives à la santé perçue au moment de l'enquête. L'ensemble des variables socioéconomiques sera également recueilli.

Cette enquête doit permettre de répondre aux questionnements sur l'impact sanitaire de l'incendie du 26 septembre 2019, à court et moyen termes, en donnant des résultats sur l'exposition ressentie par la population et sur sa santé perçue, puis en cherchant une relation entre elles. Ces résultats pourront notamment permettre d'identifier des pistes d'analyses plus poussées, en mobilisant notamment dans un second temps le Système national des Données de santé (SNDS) et en orientant les indicateurs de santé les plus pertinents à suivre.

6.5.5 Description du sondage

6.5.5.1 Base de sondage

L'échantillonnage et la constitution de la base de sondage sont réalisés par la Division Sondages de l'Insee, dans le millésime 2019 du Fichier démographique d'origine fiscale sur les logements et les individus (Fideli). Cette base de données est construite à partir des données de la Direction générale des finances publiques. Il s'agit d'un recoupement de données issues des déclarations de revenus, des fichiers de la taxe d'habitation et des propriétés bâties. Ces données sont complétées par des données issues d'autres sources : les répertoires des communautés, des résidences hôtelières ainsi que des logements locatifs des bailleurs sociaux. Relativement à une enquête par ménage, le choix d'une enquête individuelle est justifié par le fait que nous supposons que les expositions aux nuisances liées à l'incendie, et les conséquences de santé potentiellement associées, présentent plus de variabilité d'un ménage à l'autre qu'au sein d'un même ménage.

Pour pouvoir utiliser les données de Fideli, une convention tripartite a été signée entre la DGFIP, l'Insee et Santé publique France. Ce projet a obtenu un avis favorable du Comité du secret statistique pour autoriser l'accès aux sources statistiques et administratives, la communication de données fiscales, et pour permettre au sous-traitant d'accéder aux données nominatives demandées.

Le recours à cette base pour tirer l'échantillon aléatoire a plusieurs avantages :

- Bénéficier d'un défaut de couverture faible. Il manquera cependant dans le champ de l'enquête les individus qui résidaient le 26 septembre 2019 dans la zone exposée mais qui n'y vivaient pas au 1er janvier 2019 et ceux qui ne sont pas identifiés dans cette base ;
- Pouvoir tirer au sort des individus résidant dans la zone exposée aux nuisances de l'incendie grâce à la géolocalisation des ménages disponible dans Fideli ;
- Pouvoir disposer des coordonnées complètes de ces individus ou de la personne de référence du ménage : nom, prénom, âge, adresse, numéros de téléphone et mails pour une partie d'entre eux. Il sera donc possible de les contacter par téléphone filaire ou mobile disponible pour environ 70% de la population adulte avant passage à l'annuaire téléphonique ;
- Pouvoir utiliser les variables socio-économiques disponibles dans Fideli pour la correction de la non réponse et le redressement de l'échantillon final, car la base de sondage permet de disposer de variables auxiliaires pour toutes les unités de l'échantillon, répondantes comme non-répondantes ;

Certains points d'attention doivent cependant être soulevés et pris en compte :

- Certaines informations de contact (mail et numéro de téléphone) ne seront pas nécessairement celles de la personne échantillonnée mais peuvent être celles d'un autre membre du ménage ;
- Le prénom n'est que partiellement disponible pour les mineurs ;
- Aucune information n'est disponible concernant la profession ou le diplôme ;

6.5.5.2 Echantillonnage

Un échantillonnage stratifié est mis en œuvre. Dans chacune des strates, un tirage aléatoire des individus est réalisé. Un adulte par logement est tiré au sort. Lorsque l'adulte tiré au sort est parent d'enfants de moins de 18 ans, un de ses enfants est tiré au sort et fera l'objet du module « enfant » du questionnaire auquel le parent tiré au sort sera invité à répondre.

6.5.5.2.1 Strates

Les strates retenues pour l'enquête sont construites en fonction de la zone d'étude définie précédemment. Elles prennent en compte les diverses nuisances subies par la population à savoir, l'exposition au panache de fumées et ses retombées, la perception des odeurs, le bruit, les stimuli visuels et les conséquences spécifiques de l'incendie concernant les personnes résidant à proximité.

Six strates sont ainsi proposées et définies de la manière suivante :

- **Strate de proximité immédiate avec mise à l'abri** : personnes résidant dans un rayon de 700m autour de l'incendie ;
- **Strate de proximité sans mise à l'abri** : personnes résidant entre 700m et 1500m autour de l'incendie ;
- **Strate « odeurs seules »** : personnes résidant dans des IRIS concernés par les signalements d'odeurs mais non exposées au panache de fumées et à ses retombées ;
- **Strate « odeurs et panache »** : personnes résidant dans des IRIS concernés par les signalements d'odeurs et exposées au panache de fumées et à ses retombées ;
- **Strate « panache seul »** : personnes résidant dans des IRIS exposés au panache de fumées noires et à ses retombées mais non concernés par les signalements d'odeurs ;
- **Strate témoin** : personnes résidant dans les IRIS de Seine-Maritime situés à moins de 29 km de la zone industrielle du Havre ;

Un ajustement de ces zones a été réalisé par un travail mené localement au sein du Groupe Santé. Chaque membre du groupe a été consulté individuellement pour donner un avis sur le périmètre géographique initial, pour aboutir au périmètre définitif après obtention d'un consensus.

Un seul adulte par logement sera tiré au sort, avec le souhait d'interroger le plus de logements différents et l'hypothèse de réponses plus homogènes au sein d'un même logement. Après un échange avec la Division Sondages de l'Insee, nous proposons :

- Dans la strate exhaustive 0-700 m : de tirer au sort un adulte dans chaque logement en utilisant un sondage à deux degré (équivalent à un sondage à un degré avec le tirage d'un seul adulte par ménage) ;
- Dans les autres strates, dans lesquelles le taux de sondage est inférieur à 10% :
 - o De trier la base des logements par type de ménage et par revenus fiscal de référence, puis par identifiant de logement et identifiant individu ;
 - o De réaliser un tirage systématique, avec un pas de sondage constant dans chacune des strates, permettant d'éviter de tirer plusieurs individus dans un même logement, le pas de sondage étant supérieur au nombre maximum de 10 personnes adultes par foyer (au-delà, les logements sont considérés comme des communautés et sortis du champ de l'enquête).

6.5.5.2.2 Taille d'échantillon

Le calcul du nombre de sujets nécessaire repose sur des hypothèses faites a priori en fonction des principaux objectifs de l'étude. Pour les effets ressentis par la population lors de l'incendie, il s'agira de calculer un nombre de sujets nécessaire permettant 1) d'estimer la prévalence des symptômes ressentis avec une précision donnée ; 2) de comparer la prévalence des symptômes ressentis selon différentes zones d'expositions ; 3) et d'étudier la relation entre l'intensité de l'exposition et la fréquence des symptômes ressentis. Ces calculs ont été réalisés pour un risque d'erreur $\alpha=5\%$, une puissance de 80% et un effet plan égal à 1,5. Ils prennent en compte le facteur de correction de population finie lorsque le taux de sondage est élevé (>15%).

Au total, 5 000 individus adultes sont souhaités sur l'ensemble de la zone d'étude (zone exposée et zone témoin). L'interrogation de 4 000 individus permettra d'estimer une prévalence de 10% avec une précision relative proche de 10% pour la zone exposée de l'étude, de moins de 25% pour chacune des strates y compris la strate témoin, et de mettre en évidence un écart de 10% entre les prévalences de deux strates en ajustant sur d'autres facteurs. Les calculs détaillés sont présentés par ailleurs (Annexe 3).

Pour les enfants, sous les hypothèses précédemment formulées, et avec l'estimation supplémentaire que 28% des individus de 18 ans et plus interrogés ont au moins un enfant mineur, 1 400 questionnaires concernant un enfant mineur seront ainsi administrés sur l'ensemble de la zone d'étude (1 148 sur la zone exposée). Cet échantillon est suffisant pour les objectifs de publication des résultats sur la zone exposée de l'étude, et aucune surreprésentation des parents d'enfants ne semble nécessaire. Une précision relative autour de 20% est obtenue pour une prévalence de 10% sur l'ensemble de la zone exposée.

6.5.5.2.3 Allocation par strate

Compte tenu de la faible étendue des zones de proximité et que celles-ci sont majoritairement des zones industrielles, le nombre de personnes résidant à une distance de moins de 1 500 mètres autour de l'incendie est faible : 1 200 habitants à moins de 700 mètres, 23 300 à moins de 1 500 mètres. Une allocation proportionnelle dans les six strates ne nous permet pas d'atteindre les objectifs en termes d'inclusion afin de répondre aux différents objectifs de l'étude. Ainsi, compte tenu de la taille d'échantillon souhaitée sur l'ensemble et par strates, il est proposé (Tableau 3) :

- D'interroger un adulte dans chacun des logements résidant dans la strate de proximité avec mise à l'abri ;
- D'atteindre un effectif de répondants minimal de 1 000 répondants par strate, à l'exception des strates de proximité où 1 000 répondants seront obtenus sur l'ensemble de la zone 0-1 500m ;

Tableau 3 : Populations par strates et allocations correspondantes

Libellé des strates	Population adulte	%	Nb 18 + tirés au sort	Taux de mobilité estimé	Taux de réponse attendu	Nb de répondants 18+ attendus	Taux de sondage	Nb de 0-17 sélectionnés
Proximité avec mise à l'abri (<=700 m)	1 061	0,4%	613	10%	60%	331	57,8%	134
Proximité sans mise à l'abri (700-1500m)	14 913	5,5%	1 326	10%	60%	716	8,9%	510
Nuisances olfactives seulement	93 436	34,7%	2 020	10%	55%	1 000	2,2%	722
Nuisances olfactives et panache	104 419	38,7%	2 020	10%	55%	1 000	1,9%	592
Panache seulement	55 693	20,7%	2 020	10%	55%	1 000	3,6%	841
Ensemble Zone d'étude	269 522	100%	7 999			4 047	1,9%	2 799
Zone témoin	269 727		2 778	10%	40%	1 000	1,0%	1 063

10 777 individus adultes sont ainsi tirés au sort pour atteindre nos objectifs. Les hypothèses ayant conduit à cette estimation sont les suivantes :

- Un taux de mobilité, tenant compte des déménagements, décès et départs en institution ayant pu survenir entre la date de référence de la base de sondage (1er janvier 2019) et la période de collecte (septembre 2020, soit 21 mois plus tard), est estimé à 10% ;
- Le taux de réponse est estimé à 55%. Une hypothèse d'une meilleure participation est formulée dans les strates de proximité (60%), de moins bonne participation dans la zone témoin (40%) ;

Cette allocation a pour conséquence de créer de forts écarts de poids. Néanmoins, ces écarts de poids seront principalement portés par la strate de proximité 0-700 mètres. Dans les autres strates, la distribution des poids demeurera relativement équilibrée, d'un facteur 1 à 4 dans la zone d'étude, et il nous semble ainsi que le risque de points aberrants très influents est limité. Si tel était le cas, après correction de la non-réponse et après redressement, nous pourrions mobiliser différentes méthodes comme la troncature des poids de sondage ou, au moment des estimations des variables d'intérêt, la méthode dite de « winsorisation ».

Au sein de chaque strate, les variables de tri utilisées sont les suivantes, dans l'ordre: le décile de niveau de vie (les valeurs manquantes étant distinguées des déciles) ; le type de ménage (où il a été distingué en plus des modalités habituelles les couples avec enfants mineurs / majeurs et les familles monoparentales avec enfants mineurs / majeurs) ; l'identifiant de logement ; l'identifiant de l'individu.

Ces variables sont classiquement associées aux indicateurs de santé perçue, et il est important de s'assurer d'une bonne représentativité de chaque strate de l'échantillon selon ces variables, diminuant ainsi la variance des estimations sur la santé perçue. De tels effectifs permettraient ainsi, en tenant compte la correction en population finie, d'obtenir un intervalle de confiance de plus ou moins 2,0 % pour une prévalence de 10% dans chacune de strates de l'étude, et de +/- 1,1% pour l'ensemble²⁸ (Tableau 4).

Tableau 4 : IC à 95% attendus dans les différentes strates, pour une prévalence de 10%

Libellé des strates	Nb de répondants 18+ attendus	f	IC 95% autour de 10 % +/-
Proximité avec mise à l'abri (<=700 m)	331	0,42	2,0%
Proximité sans mise à l'abri (700-1500m)	716	0,04	1,9%
Odeurs seulement	1 000	0,01	2,0%
Odeurs et panache	1 000	0,01	2,0%
Panache seulement	1 000	0,01	1,9%
Ensemble Zone d'étude	4047	0,01	1,1%
Zone témoin	1000	0,00	2,0%

²⁸ Formule de variance pour chaque strate : $\hat{V}(p) = (1 - f) \frac{p(1-p)}{n-1}$; et pour l'ensemble de la zone d'étude : $\hat{V}(p) = \sum_{i=1}^5 \left(\frac{N_i}{N}\right)^2 * \hat{V}(p_i)$

6.5.5.2.4 Echantillon de réserve

Étant donné que le tirage pour chaque échantillon est réalisé de manière indépendante au sein de chacune des strates définies au moment de l'échantillonnage, un échantillon de réserve de 20% de la taille de chacune des strates de l'échantillon principal est constitué, soit 2100 contacts supplémentaires (Tableau 5).

Ces échantillons seront tirés au sort en même temps que l'échantillon principal. Ces réserves seront ouvertes avec des règles de déclenchement par strate, selon le taux de participation observé, et exploitées en intégralité.

Tableau 5 : Taille de l'échantillon principal et de l'échantillon de réserve, par strates.

Libellé des strates	Echantillon principal	Echantillon de réserve
Proximité avec mise à l'abri (<=700 m)	613	0
Proximité sans mise à l'abri (700-1500m)	1326	265
Odeurs seulement	2020	404
Odeurs et panache	2020	404
Panache seulement	2020	404
Ensemble Zone d'étude	7999	1477
Zone témoin	2778	556
TOTAL	10 777	2 033

6.5.6 Traitements statistiques

6.5.6.1 Identification et traitement des hors-champ

Les individus attendus hors-champ sont les suivants : ceux ayant déménagé à l'étranger entre le 1er janvier 2019 (date d'observation des données issues de Fidéli) et la date de collecte, ceux décédés entre le 1er janvier 2019 et la date de collecte, ceux ayant rejoint une institution entre le 1er janvier 2019 et la date de collecte, de même que les individus qui ne résidaient pas dans la zone d'étude au moment de l'incendie. Pour une partie de ces individus qui répondront à l'enquête, les informations permettant de les caractériser seront recueillies en début de questionnaire, lors la phase de vérification de l'éligibilité. Le traitement des hors champs sera réalisé conjointement avec celui de la non-réponse totale. Avec une estimation de 10 % de déménagements par an, le taux de hors-champ devrait donc être limité, sans doute inférieur à 5%, du fait 1) de la disponibilité d'une base récente de Fidéli (9 mois d'écart entre le 1^{er} janvier et le 26 septembre 2019) et 2) du fait que les individus ayant déménagé au sein de la zone d'étude feront partie du champ de l'étude.

6.5.6.2 Correction de la non-réponse partielle

La non-réponse (NR) partielle est attendue lorsqu'un individu refuse de répondre à une question particulière. La nécessité d'un traitement de la non-réponse partielle sera évaluée après analyses des données recueillies et du volume concerné (la non-réponse partielle est souvent limitée dans le cadre

d'enquêtes hétéro-administrées). Le cas échéant, des techniques d'imputations pourront être mobilisées, consistant à remplacer les valeurs manquantes par des données "plausibles", en général estimées à partir de celles des répondants. Les variables utilisées pour l'imputation seront donc des variables issues du questionnaire de l'enquête. Différentes méthodes pourront alors être utilisées: imputation par régression, hot-deck ou imputation par équations chaînées.

Le protocole de collecte prévoit deux modes de collecte : par téléphone (CATI) et par Internet (CAWI). Le taux de NR partielle attendu par CATI est très faible, du fait de l'intervention de l'enquêteur qui permet de la limiter. La correction de la NR partielle sera donc principalement mise en œuvre au sein des répondants au CAWI, avec donc un mode de réponse unique entre donneur et receveur. Cette correction sera réalisée préalablement au traitement de l'effet mode.

Nous nous attacherons en priorité à évaluer le biais de sélection, puis le biais de mesure. Concernant la méthode utilisée pour corriger un éventuel effet de mode, son choix fera l'objet d'une étude spécifique, et nous projetons de discuter ces éléments en collaboration avec le programme multi-mode de l'Insee et notamment le GT-Aval. Nous avons pris connaissance de différentes solutions rapportées dans la littérature^{29,30,31}, telles que les méthodes d'imputation, déterministes ou aléatoires, ou le recours à des méthodes de calibration. Au-delà de l'application directe au traitement des données de cette étude, une collaboration scientifique est envisagée avec ce groupe dans le but d'enrichir les connaissances et d'écrire un article scientifique.

6.5.6.3 Traitement de la non-réponse totale

La présence de non réponse dans une enquête est un problème inévitable malgré toutes les précautions que l'on peut apporter à la préparation de celle-ci. Elle recouvre différentes situations : impossibilité à joindre l'individu tiré au sort, refus de répondre (explicite ou implicite, certaines personnes n'étant plus joignables après un premier contact établi), abandon du questionnaire après avoir commencé à y répondre.

Le traitement de la non-réponse totale sera réalisé en utilisant la méthode des scores³². Cette méthode est basée sur le principe de réponse homogène des unités échantillonnées à l'intérieur de groupes de réponse homogènes (GRH). Ces GRH sont constitués en modélisant la probabilité de réponse en fonction de variables connues pour les répondants et les non-répondants. Le principe général est le suivant :

- Identification des facteurs liés à la probabilité individuelle de répondre par une régression logistique ;
- Constitution des GRH : par classement les unités échantillonnées selon leur probabilité de réponse estimée et regroupement de l'échantillon (répondants et non-répondants) en groupes de tailles approximativement égales. Une fois les GRH constitués, la probabilité de réponse est supposée indépendante d'un groupe à un autre et constante au sein d'un même groupe ;

²⁹ Klausch L.T. (2014), *Informed Design of Mixed-Mode Surveys. Evaluating mode effects on measurement and selection error*, Ph.D. dissertation, Utrecht University: Utrecht, The Netherlands.

³⁰ Legleye et al. (2018). Agréger les échantillons d'une enquête multimode en limitant l'effet de mesure : une proposition d'imputation raisonnable et pragmatique. *JMS* 2018.

³¹ Kolenikov, S. and Kennedy, C. (2014). Evaluating three approaches to statistically adjust for mode effects. *Journal of survey statistics and methodology*, 2(2), 126–158 ; De Leeuw E.D. Mixed-Mode: Past, Present, and Future, *Survey Research Methods*, 2018, 12(2):75-89.

³² Haziza, D. et Beaumont, J.-F. (2007). : On the construction of imputation classes in surveys, *International Statistical Review*, 75, 25-43.

- Calcul de la pondération finale : la pondération finale pour une unité donnée est obtenue en multipliant sa pondération initiale par l'inverse de la probabilité de répondre au sein de son groupe homogène ;

L'utilisation de Fidéli comme base de sondage permet de disposer d'informations chez les répondants et les non-répondants. Les variables à intégrer sont celles qui sont à la fois liées aux variables d'intérêt de l'enquête et au fait de répondre ou non à l'enquête. Ainsi, les variables pouvant être utilisées pour corriger la non-réponse totale sont :

- La tranche d'âge de l'individu
- Le sexe de l'individu
- Le type de logement (maison individuelle, appartement...)
- Le statut d'occupation du logement (propriétaire, locataire...)
- La présence d'un enfant de moins de 18 ans dans le ménage pour l'échantillon des adultes
- La présence d'un enfant en bas âge dans le logement ;
- La situation matrimoniale
- Le type d'individu fiscal (référént, conjoint, pac)
- Le niveau de vie dans le logement approché par le revenu fiscal de référence par part
- Perception d'allocations chômage ou préretraite
- Indicatrice de ménage pauvre à 60 %
- La taille d'unité urbaine
- Le type de ménage
- Le nombre d'habitants dans le logement,
- Variables de contact : indicatrice de présence de mail et de téléphone (fixe/mobile), indicatrice d'absence de mail et de présence de téléphone, indicatrice de présence d'informations de contact directes de l'individu...)

Dans le cas où une non-réponse totale différente serait observée au sujet des questions concernant les enfants (si certains individus par exemple refusent de répondre au sujet de leur enfant sélectionné), la méthode des scores pourra être appliquée séparément pour cet échantillon.

D'autres variables, issues de l'appariement des données avec le SNDS, seront mobilisées pour établir le modèle de non-réponse. Il apparaît en effet important de pouvoir contrôler un éventuel biais d'auto-sélection en lien avec la santé, variable d'intérêt de l'enquête, ces mécanismes pouvant être différents entre zone d'étude et zone témoin.

6.5.6.4 Calage sur marge

Le principe général du calage sur marge est d'ajuster les poids de sondage après traitement de la non-réponse des unités enquêtées de façon à ce qu'elles reproduisent des distributions connues par ailleurs sur la population de référence et concernant des variables en lien avec la variable d'intérêt. Ceci permet, sous certaines conditions, d'améliorer la précision des estimateurs. Les poids de sondage corrigés de la non-réponse sont repondérés en utilisant un calage sur marges.

Les marges de calage seront issues de la base de sondage. Les variables de calage peuvent être sélectionnées parmi les variables suivantes : la tranche d'âge et le sexe des individus, le type de ménage, la TUU, le niveau de diplôme, la PCS, la présence d'un enfant de moins de 18 ans dans le foyer (pour les adultes), le niveau de revenu.

6.5.6.5 Calcul des poids finaux

Concernant les adultes, le plan de sondage étant stratifié avec un tirage aléatoire simple au sein de chaque strate, la pondération initiale de chaque répondant est égale à l'inverse du taux de sondage

dans sa strate de sondage. La probabilité d'inclusion est calculée en prenant comme dénominateur l'ensemble des adultes de 18 ans et plus. En notant h la strate de tirage, n_h^A le nombre d'adultes tirés au sort dans la strate h et N_h^A le nombre d'adultes de la strate h dans la base de sondage, le poids de sondage initial pour les adultes de 18 ans et plus tirés dans la strate h s'écrit :

$$W_h^A = \frac{N_h^A}{n_h^A}$$

Après correction de la non-réponse, le poids de chaque individu correspondra **au poids de sondage divisé par la probabilité de réponse dans le GHR** estimée par le modèle. Ces poids corrigés de la non-réponse sont ensuite utilisés en entrée de la fonction de calage, et les poids en sortie sont les poids calés utilisés pour la diffusion et les analyses.

Les enfants sont tirés au sort conditionnellement au tirage de l'adulte (c). Une fois l'adulte tiré au sort, il s'agira de tirer exhaustivement leur enfant s'il en a un seul, et de tirer aléatoirement un de ses enfants s'il en a plusieurs.

En notant :

- h la strate de tirage ;
- W_h^A le poids de sondage de l'adulte tiré au sort ;
- f_E le nombre d'enfants du ménage ;

Le poids de l'enfant tiré au sort est donné par : $W_h^E = f_E * W_h^A$

Lorsque les enfants vivent en garde alternée chez deux parents vivant dans la zone d'étude, leur poids sera ajusté en utilisant la méthode de partage des poids^{33, 34}. Des questions spécifiques figurent dans le questionnaire afin de déterminer l'éligibilité des parents en cas de garde alternée.

Les enfants de famille nombreuse seront sous-représentés, mais se verront attribués un poids d'autant plus important que la taille du foyer est importante. Ce choix se justifie par l'hypothèse d'une plus grande variabilité inter-ménages qu'au sein d'un même logement concernant les conséquences sur la santé de l'incendie.

6.5.6.6 Evaluation de l'effet lié au mode de collecte

Un arbitrage a dû être effectué avec le recours à un mode unique de collecte. Ce choix aurait résolu la question de l'effet mode, mais aurait présenté d'autres limites, et notamment un taux de collecte plus faible en monomode (CATI ou CAWI) qu'en multimode, avec une introduction potentielle d'un biais de sélection si les non-répondants au mode de collecte retenu (CATI ou CAWI) présente des caractéristiques particulières, et d'un biais de couverture, du fait que pour environ 20% des individus sélectionnés on ne dispose pas de numéro de téléphone, et que d'autre part une part encore importante de la population française ne dispose pas ou n'est pas en capacité d'utiliser facilement internet³⁵. Si le mode de recueil est associé à la nature des réponses collectées, notamment sur des

³³ Deville, J.-C., Lavallée, P. (2006). Sondage indirect : les fondements de la méthode généralisée du partage des poids. Techniques d'enquête, Vol. 32, No 2, p. 185.

³⁴ La méthode du partage des poids. Cyril Favre-Martinoz et Emmanuel Gros. Département des méthodes statistiques, Version no 1, diffusée le 10 octobre 2017 (<https://www.insee.fr/fr/information/2838097>).

³⁵ Selon le baromètre numérique 2019 du Crédoc, 88% des Français sont connectés à internet en 2019. Cependant, la part des français qui n'utilisent jamais les outils informatiques et numériques ou qui abandonnent à la moindre difficulté est de 16%. <https://www.credoc.fr/publications/barometre-du-numerique-2019>

échelles de santé perçue pouvant être sensible à l'administration ou non par un enquêteur, il semble important de pouvoir mesurer l'effet du mode de collecte.

L'enquête proposant deux modes de collecte, en CATI et en CAWI, il est nécessaire de mesurer l'impact potentiel du mode de collecte sur les principales variables d'intérêt. L'évaluation de l'effet mode sera réalisée en distinguant le biais de sélection des répondants ayant répondu sur un mode de collecte spécifique d'une part, l'effet lié à une différence dans les estimations obtenues en contrôlant l'effet de sélection d'autre part.

La méthode proposée repose sur l'utilisation successive :

- D'un modèle logistique visant à obtenir un score de propension à avoir répondu par un mode de collecte (CATI vs CAWI) en fonction de différentes caractéristiques sociodémographiques et de variables disponibles dans la base de sondage ;
- De la définition d'une zone commune, définie par un score propension compris entre le minimum du score de propension observé pour le mode CATI et le maximum du score de propension observé pour le mode CAWI ;
- Au sein de cette zone commune, il sera possible :
 - o De comparer les individus selon les 2 groupes (CATI / CAWI) sur plusieurs variables à l'aide de tests statistiques, de différences standardisées ;
 - o D'analyser les différences entre groupes avant et après correction de l'effet de sélection par appariement sur score de propension ou par repondération par l'inverse du score de propension ;
 - o D'envisager si nécessaire une méthode de correction de l'effet de mesure par imputation pour les individus aux caractéristiques sociodémographiques semblables dans les deux modes mais présentant des réponses différentes pour la variables d'intérêt.

Parmi les variables à prendre en compte dans cette analyse, les informations de contacts disponibles dans Fideli (adresse mail, téléphone fixe et mobile) seront mobilisées. En outre, la prise en compte des usages d'Internet et du téléphone, nécessaires comme variable de contrôle, seront estimées par l'intégration des questions suivantes dans l'enquête, portant sur la fréquence d'usage d'internet et sur le fait de décrocher son téléphone (mobile ou fixe) lorsqu'il sonne.

- Décrochez-vous, ou une personne de votre ménage décroche-t-elle lorsque votre téléphone **fixe** sonne ? : Oui, mais seulement lorsque vous connaissez le numéro qui essaie de vous joindre/ Oui, sans vous soucier du numéro qui essaie de vous joindre/ Non, jamais / Non, il ne sonne jamais / vous n'avez pas de téléphone fixe ;
- Décrochez-vous, ou une personne de votre ménage décroche-t-elle lorsque votre téléphone **mobile** sonne ? : Oui, mais seulement lorsque vous connaissez le numéro qui essaie de vous joindre/ Oui, sans vous soucier du numéro qui essaie de vous joindre/ Non, jamais / Non, il ne sonne jamais / vous n'avez pas de téléphone mobile ;
- Au cours des trois derniers mois, en moyenne, vous avez utilisé internet ? Tous les jours ou presque ? / Pas tous les jours, mais au moins une fois par semaine ? Moins d'une fois par semaine ? Jamais ? Vous n'avez pas accès à internet (*Consigne : Quel que soit le lieu ou le type d'accès. Nous nous intéressons aux usages dans tout type de lieu (à votre domicile, au travail ou à un autre endroit) et sur tout type de support (ordinateurs, tablettes, smartphones, liseuses, PDA, consoles de jeu...)*)

6.5.7 Description du mode de collecte

6.5.7.1 Protocole de collecte

La collecte des questionnaires repose sur un multimode séquentiel en deux phases, internet (CAWI) puis téléphone (CATI). La collecte en face-à-face n'est pas privilégiée, sauf exception, pour des raisons de coûts et de durée de collecte, ainsi que du contexte sanitaire actuel. La possibilité de mobiliser un mode de collecte supplémentaire par questionnaire papier a été discutée, permettant notamment d'offrir une alternative aux personnes peu familières ou peu à l'aise avec Internet. Toutefois, la charge supplémentaire, tant pour le prestataire de collecte (maquettage du questionnaire, gestion manuelle d'un nouveau mode de recueil, saisie optique) que pour les équipes en charge du traitement aval (phase d'apurement, traitement de l'effet mode avec un mode supplémentaire), ceci dans un calendrier de collecte et de restitution des résultats relativement contraints, nous a amené à écarter ce choix.

6.5.7.1.1 *Précisions concernant la collecte*

- Seul l'individu sélectionné sera interrogé pour répondre au questionnaire. Si cette personne a des enfants mineurs, elle sera invitée à répondre également à un court module de questions qui concernera l'un de ses enfants, tiré au sort en amont de la prise du contact ;
- Le questionnaire n'est rempli qu'après validation / vérification de l'identité du répondant ;
- Une enquête test a été mise en place, afin de valider la bonne compréhension des questions et la durée moyenne du questionnaire ;
- La collecte durera de 6 à 8 semaines. Elle sera réalisée par téléphone et en ligne, à l'aide d'un site mis à disposition pour répondre à l'enquête et pour renseigner sur l'enquête ;

6.5.7.1.2 *Données disponibles pour le contact*

Ce dispositif bénéficie de la diversité des informations de contact disponibles dans Fidéli du millésime 2019 :

- Nom, Prénom pour la quasi-totalité des individus, à l'exclusion d'une partie importante des mineurs qui seront contactés en utilisant leur année de naissance ;
- L'adresse postale pour 100 % des individus ;
- Au moins un numéro de téléphone (individu sélectionné ou une personne dans le logement) pour 70 % à 75 % des individus ;
- Au moins un numéro de téléphone portable (individu sélectionné ou une personne dans le logement) pour environ la moitié des individus ;
- Une adresse mail (individu sélectionné ou une personne dans le logement) pour environ deux tiers des individus ;
- A contrario, pour environ 15% des individus, le seul mode de contact possible est le courrier postal ;

6.5.7.1.3 *Enrichissement (numéros de téléphone) des données contact*

Une fois le fichier livré au prestataire, ce dernier effectuera un contrôle de qualité des données présentes (présence des champs nécessaires, contrôle des variables de contacts), afin de contrôler plus efficacement le routage des courriers postaux. Une fois les fichiers contrôlés, un travail d'enrichissement en numéros de téléphone sera effectué, par une recherche automatique réalisée par un seul prestataire référencé, en indiquant les numéros de téléphone trouvés à partir du Nom et de l'adresse postale. Cet enrichissement sera effectué uniquement sur les adresses n'ayant pas de numéros de téléphone renseignés (environ 30%) dans la base Fideli.

En complément de cette étape d'enrichissement, les coordonnées (numéros de téléphone, adresse mail) des individus pourront également être actualisées directement par la personne elle-même sur le site Internet dédié à l'enquête, via le module de mise à jour des coordonnées, ou par un contact à la hotline.

6.5.7.1.4 Phase de contact

Sont mis en place un numéro vert et un site internet dédié permettant à la personne sélectionnée, après connexion (avec mot de passe) de prendre rendez-vous pour un entretien téléphonique et de préciser le numéro de téléphone sur lequel elle peut être jointe. La procédure de contact des individus tirés au sort suivra les étapes suivantes :

- **Une lettre-avis** est envoyée à tous les foyers comportant une personne échantillonnée, accompagnée d'un dépliant d'information. Cette lettre est accompagnée d'un **mail-avis** (si une adresse mail est disponible). Le message du mail précise le prénom et le nom de la personne échantillonnée à qui il est demandé de transférer le mail. Lorsque l'on dispose d'un numéro de téléphone portable pour un foyer, **un SMS** est également envoyé, invitant à participer à l'enquête et précisant le prénom et nom de la personne échantillonnée. La lettre-avis permet de notamment de proposer l'enquête aux personnes pour lesquelles on ne dispose ni d'un mail, ni d'un numéro de téléphone portable dans Fidéli dont on sait qu'ils ont des profils spécifiques sans doute lié au sujet de l'enquête (plus âgés par exemple).
- La lettre-avis est adaptée pour les personnes dont on ne disposerait que de l'année de naissance dans Fidéli (pas de prénom/nom). Dans ce cas, un mail ou SMS si possible et la lettre-avis seront envoyés à la personne de référence, précisant l'année de naissance de la personne sélectionnée et demandant de transmettre l'information. Par ailleurs, l'enquêteur demandera à parler à la personne née l'année correspondante dans le foyer.
- L'adresse du site internet auquel se connecter, ainsi qu'un QR code, seront mentionnés dans tous les modes de contacts (courrier postal, courriel et SMS).

6.5.7.1.5 Phase de collecte et relances

La collecte est réalisée initialement par internet, puis par internet et par téléphone, afin de favoriser la participation des personnes répondant peu sur Internet (personnes plus âgées ou plus défavorisées par exemple). Un site internet est mis à disposition pour répondre à l'enquête et pour renseigner sur l'enquête. La durée de collecte est estimée de 6 à 8 semaines (Tableau 6). La durée du questionnaire est estimée à 30 minutes.

10 jours après l'envoi de la lettre-avis, les foyers qui ont un numéro de téléphone mais qui n'ont pas déjà rempli le questionnaire sur internet sont alors sollicités par téléphone. Les appels seront réalisés en priorité auprès des personnes pour lesquelles aucun mail n'est disponible, afin d'augmenter le temps de collecte par internet auprès des personnes pouvant être relancées par mail, et réduire ainsi le nombre de numéros de téléphone à solliciter dans la suite.

Les appels téléphoniques sont réalisés avec un coefficient d'insistance de 15 tentatives maximum à différentes heures de la journée et différents jours de la semaine, y compris le samedi. Dans le cas où une personne sélectionnée mentionnera ne pas avoir été informée de l'enquête, il lui sera proposé de transmettre la lettre-avis par courriel, ou par courrier postal si besoin.

Les personnes sélectionnées seront relancées pour participer à la collecte :

- 4 fois par mail et SMS (j+5 et j+10), lorsque ces informations sont disponibles ;
- 1 fois par lettre-avis, pour les personnes dont le numéro de téléphone n'est pas renseigné et pour lesquelles le seul mode de collecte possible est internet. Pour ces personnes, les modes

de contacts sont de fait réduits, puisque limités aux seuls envois de courriers postaux dans lesquels plusieurs possibilités de réponses sont proposées (prise de rdv par téléphone, questionnaire en ligne).

Tableau 6 : Calendrier de la collecte et des relances

Enquête principale : CAWI exclusif (10 j) puis mode mixte CAWI CATI (23 j)	21 août - 27 octobre.
<i>Livraison du fichier contact par l'Insee</i>	<i>21-août</i>
<i>Fichier nettoyé/enrichi</i>	<i>31-août</i>
<i>Envoi des lettres-avis</i>	<i>04-sept</i>
<i>Envoi des mails-avis et SMS</i>	<i>07-sept</i>
<i>Relances par mails et SMS</i>	<i>11 sept ; 18 sept; 9 oct ; 16 oct</i>
<i>Relance par lettre-avis (personnes sans numéro de téléphone)</i>	<i>18 sept ; 8 oct</i>
<i>Formation des enquêteurs</i>	<i>14-sept</i>
<i>Collecte Téléphonique (personnes sans mail en priorité)</i>	<i>15-sept</i>
<i>Fin de collecte</i>	<i>27-oct</i>

6.5.7.1.6 *Echantillon principal et échantillon de réserve*

Étant donné que le tirage pour chaque échantillon (principal et réserve) est réalisé de manière indépendante au sein de chacune des strates définies au moment de l'échantillonnage, si les taux de réponse quatre semaines après le lancement de la collecte ne sont pas jugés satisfaisants (inférieurs de 10 points aux taux de réponses attendus), un échantillon de réserve de 20 % de la taille de l'échantillon principal pourra être activé, de façon indépendante pour chacune des strates. Ce ou ces nouveaux échantillons seront alors contactés et relancés selon le même principe que l'échantillon principal.

6.5.7.2 Collecte des données pour les personnes résidant sur l'aire des gens du voyage

Les Gens du voyage résidant sur l'aire de Petit-Quevilly font partie de la population de la zone exposée : ils habitent dans la strate de proximité (0-700 m) dans laquelle tous les ménages sont inclus.

Pour cette population, le sondage ne pouvait pas être réalisé à partir de la base Fideli. Par conséquent l'enquête auprès de cette population a été réalisée selon des moyens spécifiques dans le but de pouvoir intégrer ce groupe de personnes éligibles dans l'étude et lui administrer le questionnaire d'enquête selon des modalités adaptées.

Un adulte est tiré au sort parmi les adultes (+18 ans) de chaque foyer (ou emplacement). Le tirage est réalisé à partir d'informations sur la constitution des foyers. Lorsque l'adulte tiré au sort a un ou des enfants, il est interrogé également au sujet de cet enfant, ou de l'un d'entre eux tiré au sort.

Les critères d'inclusion suivants ont été définis : les personnes enquêtées étaient présentes sur l'aire de Petit-Quevilly au moment de l'incendie. Les familles ayant déménagé entre-temps de l'aire de Petit-Quevilly seront également incluses si elles peuvent être jointes par le médiateur santé ou aux autres familles et si leur lieu de vie actuel se situe dans les environs de Rouen.

Un binôme d'enquêteurs issu de la cellule régionale Normandie de Santé publique France se rendra sur place pour administrer le questionnaire, accompagné du médiateur.

6.5.8 *Enquête complémentaire sur la santé mentale*

Afin d'être en mesure de réaliser une post-enquête auprès de la population interrogée, une question spécifique est posée en toute fin de questionnaire pour permettre aux enquêtés de communiquer leur

accord ou refus quant à la conservation de leurs coordonnées. Il est demandé au répondant de préciser ses informations de contact pour d'éventuelles interrogations ultérieures. A l'issue de la collecte, ne seront conservées que les données nominatives des personnes qui auront donné leur accord. Elles seront archivées de manière sécurisée sur les serveurs internes de Santé publique France.

La réalisation d'une post-enquête est justifiée par la contrainte de temps de passation du questionnaire, souhaité d'une durée moyenne inférieure ou égale à 30 minutes, qui a nécessité de réaliser des arbitrages dans les thématiques abordées. En particulier, la passation de questionnaires standardisés (PHQ9, GAD7, PCL5) portant sur la santé mentale (respectivement la dépression, l'anxiété et le stress post-traumatique) a été écartée, au profit d'un questionnaire de santé globale. Afin de permettre de recueillir ces informations importantes, il est proposé de poser ce questionnaire complémentaire dans une post-enquête dédiée.

6.5.9 Appariements

Les données collectées seront enrichies de certaines informations disponibles dans la base de sondage (revenu, information géographique, caractéristiques des logements, composition des ménages, niveau de vie), ceci permettant d'alléger la charge d'enquête.

Sous réserve d'autorisation de la CNIL, Les données collectées seront également enrichies avec les informations concernant les consommations de soins des personnes interrogées (soins en ville et à l'hôpital, causes médicales de décès) issues du Système national des données de santé (SNDS).

Cet appariement avec les données du SNDS fait l'objet d'une demande d'autorisation auprès du CESREES et de la CNIL.

6.6 Données collectées par questionnaire

Le questionnaire d'enquête est destiné à être administré au téléphone, par un professionnel de la société Ipsos, pour une durée d'environ 30 minutes, ou en mode auto-administré en ligne pour les personnes ne pouvant pas répondre au téléphone (absence d'équipement, handicap auditif) ou étant plus à l'aise avec ce mode de collecte.

La conception du questionnaire est fondée sur un ensemble de sources :

- l'expérience de Santé publique France qui a déjà réalisé plusieurs études de santé déclarée, notamment autour de la plateforme industrielle de Salindres¹³ ;
- les apports d'une pré-enquête qualitative réalisée en février 2020 dans quatre communes autour de Rouen. Cette étude sociologique, conduite par focus groupes et entretiens individuels³⁶, a permis de recueillir les préoccupations en termes de santé et les attentes des habitants des communes impactées par l'incendie, ainsi qu'avoir un aperçu des nuisances subies (bruit d'explosion, incendie, fumées, types d'odeurs, suies, etc.) ;
- les échanges avec les membres du Groupe Santé - constitué d'élus et de professionnels de santé locaux, de représentants associatifs, d'un anthropologue - réalisés par courriels et lors de deux réunions organisées en mars et avril 2020 sur la pré-enquête et le questionnaire (sujets d'intérêt, structure, équilibre entre thèmes) ;
- les résultats d'une enquête menée en octobre 2020 par l'Union régionale des médecins libéraux (URML) de Seine-Maritime, auprès des praticiens exerçant dans les communes touchées, qui ont permis d'affiner la liste de troubles de santé ressentis par les patients exposés aux nuisances générées par l'incendie ;

³⁶ Salvaing L. RJ. Analyses des préoccupations, attentes et savoirs mobilisés par les parties prenantes autour du site incendié de Lubrizol et Normandie Logistique. Note de synthèse – Etude qualitative. Kantar, 2020 février 2020. Report No.

- l'échelle SF12 pour évaluer la santé actuelle des personnes interrogées : le questionnaire SF-12 est un instrument général d'évaluation de l'état de santé ; il combine une information synthétique qui cumule un score sur la dimension physique et un score sur la dimension mentale. Il comprend 12 items, répartis sur les mêmes 8 dimensions que le SF-36, la version longue dont il est extrait.³⁷

6.6.1 Questionnaire pour l'échantillon exposé à l'incendie

Mis à part les modules consacrés au contact et à la validation des coordonnées des personnes enquêtées, le questionnaire est formé de deux grands ensembles de modules qui se réfèrent à des temporalités différentes de l'état de santé qui vont être explorées par l'étude :

1) Une partie qui porte sur la période de l'accident et ses suites, dont les modules constitutifs renseignent sur :

- les expositions perçues durant l'accident : proximité de l'incendie, panache de fumées noires, odeurs, suies, débris de fibrociment ;
- les comportements adoptés suite à l'incendie ;
- le niveau d'information, la connaissance et l'utilité ressentie des recommandations ;
- les symptômes et problèmes de santé attribués à l'accident (Adulte et enfant) ;
- la grossesse et l'allaitement.

L'approche participative qui a été mise en place avec le Groupe Santé, complétée par une pré-enquête de nature sociologique, a permis une co-construction avec des membres et représentants de la population d'étude. Cette partie du questionnaire est donc particulièrement adaptée au contexte de l'accident industriel ainsi qu'aux attentes et sujets de préoccupation de la population qui en a été victime.

Le recueil de ces données est nécessaire pour mesurer précisément l'exposition des personnes et les éventuels problèmes de santé consécutifs à cette exposition.

2) Une partie qui porte sur la santé actuelle, composée des modules suivants (entre parenthèse sont indiquées les sources des modules non spécifiques à l'étude) :

- Santé globale : échelle internationale SF12 ;
- Indicateurs en lien avec la santé : taille, poids, recours aux soins, consommations de tabac et d'alcool ;
- Attitudes à l'égard de l'environnement ;
- Effet COVID-19 ;
- Conditions de vie / Profession ;
- Lien avec les entreprises Lubrizol et NL Logistique ;

Une question portant sur l'issue de grossesse a été posée, pour les femmes qui déclarent avoir été enceinte au moment, ou depuis l'incendie. Cette question, qui concernera vraisemblablement peu de personnes, vise à enrichir les connaissances sur d'éventuelles conséquences de l'incendie sur l'issue de grossesse (dont le recours ou non à une IVG), sujet d'inquiétude important depuis l'incendie et souligné par le Groupe Santé.

³⁷ Gandek B, Ware JE, Aaronson NK, Apolone G, Bjorner JB, Brazier JE, et al. Cross-validation of item selection and scoring for the SF-12 Health Survey in nine countries: results from the IQOLA Project. International Quality of Life Assessment. J Clin Epidemiol 1998;51(11):1171-8.

Le module relatif à la COVID-19 a été introduit afin de contrôler l'effet de l'épidémie et ses conséquences sur les réponses des personnes interrogées. Une question portant sur le fait d'avoir été infecté par la COVID-19 a été posée pour en tenir compte dans les analyses portant sur l'état de santé actuel.

De même, les consommations de tabac et d'alcool ont été intégrées afin de tenir compte de ces facteurs de risque important sur l'état de santé.

Un module relatif à l'usage du téléphone et d'internet a été introduit afin de permettre de contrôler l'effet mode de collecte induit par le multimode.

6.6.2 Questionnaire pour l'échantillon témoin

Outre la validation de leurs qualités métrologiques en langue française, un critère de choix des instruments de mesures psychométriques est l'existence de données de référence observées dans la population générale en France à une période récente. Il est motivé par l'objectif de comparer les résultats obtenus dans l'échantillon étudié à une référence externe pertinente. Or, l'épidémie COVID-19 produit des effets déjà documentés sur la santé mentale de la population française, ce qui compromet la comparabilité des données de l'étude à des références produites dans un contexte moins perturbé. En conséquence, il a été décidé de conduire l'étude sur une population témoin, elle aussi touchée par l'épidémie mais n'ayant pas été exposée à l'accident industriel survenu à Rouen en septembre 2019.

Un questionnaire a donc été élaboré pour cet échantillon témoin. Il ne comporte pas les modules portant sur les expositions perçues durant l'incendie, les comportements adoptés suite à l'incendie, le ressenti et la confiance des personnes par rapport aux informations et la gestion de l'événement. Pour le reste, il est identique au questionnaire administré à l'échantillon exposé à l'accident industriel.

Sa structure est la suivante :

- Santé globale : échelle internationale SF12 ;
- Indicateurs en lien avec la santé : taille, poids, recours aux soins, consommations de tabac et d'alcool ;
- Attitudes à l'égard de l'environnement ;
- Effet COVID-19 ;
- Conditions de vie / Profession ;
- Lien avec les entreprises Lubrizol et NL Logistique ;

6.6.3 Coordonnées des personnes interrogées

Le recueil des coordonnées des participants à l'enquête sera proposé, dans le but de leur transmettre les principaux résultats de l'étude et de les recontacter dans le cadre d'une post-enquête.

6.7 Enquête pilote

Les caractéristiques de l'enquête pilote sont les suivantes :

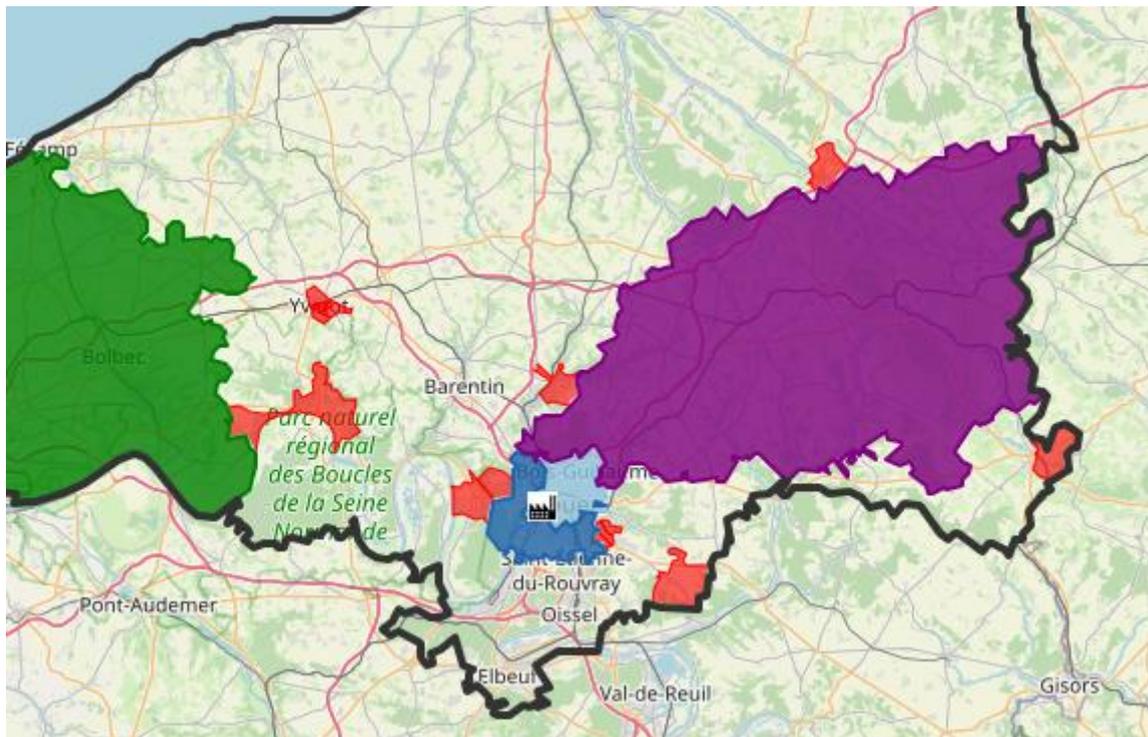
- 9 communes ont été sélectionnées (figure 5), en dehors de la zone d'étude de l'étude principale, dans 3 regroupements prédéfinis :
- regroupement pour le test de la zone témoin ;
- regroupement pour le test de la zone de proximité : les communes retenues sont celles dans lesquelles quelques signalements d'odeurs (moins de 6, les excluant de la zone d'étude) ont été recueillis dans les suites de l'incendie ;
- regroupement pour le test des communes éloignées ;

- Inclusion de communes de tailles différentes ;
- 900 adresses ont été commandées, 300 par zone ;
- Un objectif de 100 questionnaires est souhaité ;
- Les deux modes de collectes sont testés, 5 jours étant accordés au CAWI et 3 jours au CATI ;

Le protocole et le calendrier de collecte de l'étude pilote sont les suivants :

- Envoi des lettres-avis, comprenant un lien pour répondre en CAWI : 11 juin ;
- Formation des enquêteurs : 15 juin matin ;
- Terrain d'enquête CATI : 15 juin après-midi – 18 juin puis du 5 au 8 juillet ;

Figure 5 : Communes (en rouge) sélectionnées pour l'étude pilote



6.8 Traitement de données à caractère personnel et protection du secret statistique

6.8.1 Fondement légal

Le traitement des données à caractère personnel mis en œuvre pour cette étude est fondé sur l'intérêt public. Santé publique France est responsable de ce traitement. Les finalités et les modalités de mises en œuvre de l'enquête en santé déclarée, ainsi que la présente étude, ne correspondent pas à celles d'une recherche impliquant la personne humaine et n'a donc pas été soumise à l'avis d'un CPP. Cette étude s'inscrit dans la mission légale de Santé publique France relative à la surveillance de la santé de la population en France (L.1413-1 du code de la santé publique). Elle contribuera à apporter des informations très attendues sur l'exposition de la population aux nuisances de l'incendie du 26 septembre 2019, sur d'éventuelles conséquences sur la santé, sur les points d'améliorations dans l'information diffusée lors d'événements de ce type.

6.8.2 Responsabilité de traitement

Santé publique France est responsable du traitement mis œuvre. Il fait appel à Ipsos comme sous-traitant pour réaliser la collecte.

6.8.3 Secret statistique

Cette sous-traitance a été déclarée et a obtenu l'avis favorable du Comité du secret statistique. Pour la réalisation de cette enquête, l'Insee fournit à Santé publique France un échantillon d'individus à interroger. Les fichiers transmis par l'Insee correspondent à des échantillons construits en tenant compte des spécificités précisées par Santé publique France en vue de réaliser l'enquête. Le tirage des échantillons est réalisé dans la base Fidéli. L'utilisation par Santé publique France des données transmises par l'Insee sera conforme aux lois et réglementations sur l'obligation et le secret en matière de statistique et d'informatique actuellement en vigueur.

6.8.4 Personnes concernées

Seront contactés 10 777 adultes répondant aux critères définies au point 6.5.3. Au regard du recueil par questionnaire, on estime les réponses au questionnaire à :

- 5000 pour les adultes
- 1400 pour les enfants (les données sont recueillies auprès de l'adulte interrogé).

6.8.5 Respect du droit des personnes

Les personnes bénéficient d'un droit d'accès, d'opposition, de rectification, d'effacement ou de limitation de traitement pour les données les concernant ou concernant leur enfant pendant la période de conservation des données d'identification. Les personnes seront informées de leur liberté de participer et de leurs droits quant à l'accès et à la rectification des données recueillies.

Les informations utiles pour faire valoir ce droit seront mentionnées dans la lettre-avis, le courriel-avis, la plaquette d'information et le site internet de l'étude. Les demandes d'exercice des droits seront adressées par courrier électronique ou par courrier au DPO de Santé publique France.

Les refus exprimés lors de la passation du questionnaire par un enquêteur seront enregistrés par Ipsos et transmis à Santé publique France. Les demandes relatives à l'exercice des autres droits reçus par IPSOS seront transmises au DPO de Santé publique France.

6.8.6 Description des bases de données relatives à l'enquête (hors appariement).

Différentes bases de données seront créées pour la réalisation de l'enquête. Le Tableau 7 résumé les accès et durées de conservation de ces bases de données.

- **Bases issues de l'échantillonnage réalisé à partir du Fichier Fidéli (Insee)**

- **Fichier CONTACT :**

- Le Fichier CONTACT contient les informations personnelles issues du fichier Fideli permettant de contacter les individus : identifiant Insee (Id), identité des personnes ainsi que les coordonnées postales, téléphoniques et électroniques pour la personne tirée au sort ainsi que pour le référent fiscal principal et secondaire, présence d'enfant mineur dans le foyer et identité de l'enfant tiré au sort (incluant l'année de naissance pour pallier l'éventuelle absence du nom et du prénom).

- Conformément aux dispositions de la loi n°51-711 du 7 juin 1951 sur l'obligation, la coordination et le secret en matière de statistiques, le fichier « CONTACT » est transmis au prestataire de Santé publique France Ipsos suite à l'avis favorable du Comité du Secret Statistique en date du 22 juin 2020 et du 9 juillet 2020 et à l'autorisation de la DGFIP.
- Le Fichier CONTACT est :
 - Transmis par l'Insee à Ipsos ;
 - Utilisé par Ipsos uniquement pour contacter les répondants à l'enquête, puis supprimé dans un délai maximal de 3 mois à compter de la fin de la collecte.

➤ Fichier **REDRESSEMENT** :

Ce fichier pseudonymisé contient des informations relatives aux caractéristiques de tous les individus du logement, du ménage et du logement de la base de sondage, en vue de procéder aux redressements. Ce fichier contiendra également les informations concernant le poids statistique de l'enfant tiré au sort et ses composantes (poids statistique de l'adulte, nombre d'individus majeurs dans le logement, nombre d'individus mineurs dans le logement, probabilité de sélection de l'enfant au sein du logement). Il ne contient aucune information permettant une identification directe. Ce fichier sera constitué en parallèle et réservé à la correction de la non réponse.

Le fichier REDRESSEMENT est transmis à Santé publique France suite à l'avis favorable du Comité du Secret Statistique en date du 22 juin 2020 et du 9 juillet 2020 et à l'autorisation de la DGFIP.

Il est transmis par l'Insee à Santé publique France et supprimé une fois les poids de sondage finaux pour les différentes analyses statistiques envisagées réalisées, dans un délai maximal de six ans à compter de la réception du fichier ;

Santé publique France utilisera les données fournies dans le fichier REDRESSEMENT pour corriger la non-réponse, et alléger la charge d'enquête en conservant les informations suivantes contenues dans la base de sondage : revenus, coordonnées géographiques XY, type de logement, type de ménage, niveau de vie.

Les coordonnées géographiques (géolocalisation) contenues dans ce fichier permettront de construire une variable non directement identifiante représentant la distance du logement de l'individu à l'incendie, et qui sera conservée dans la base d'analyses. D'autre part, elles seront utilisées dans le cadre de la constitution d'une base distincte, anonymisée, visant à analyser la cartographie des expositions, qui ne contiendra que les informations suivantes : coordonnées XY et nature de l'exposition : visuelle, auditive, retombées de débris, retombées de suies, panache et odeurs. Ces données ne seront conservées que dans cette base d'analyse spécifique, qui ne contiendra aucune information directement identifiante.

- **Bases de collecte (Ipsos) :**

A l'issue de la collecte et après livraison des bases à Santé publique France, Ipsos s'engage à conserver de façon sécurisée l'ensemble des fichiers constitués dans le cadre du marché pour

une durée de 12 mois, et à détruire l'ensemble de ces fichiers à la demande de Santé publique France et à fournir une attestation de destruction.

➤ **Base de gestion et de suivi :**

Cette base est créée par Ipsos après réception et consolidation et enrichissement du fichier CONTACT. Elle contient les informations suivantes : identifiant IPSOS (Id-Ipsos), données identifiantes (identifiant enquête, nom, prénom, adresse, numéro de téléphone, numéro de l'enquêteur), suivi de l'issue de chaque contact : envoi de courrier, heure d'appel, rendez-vous, questionnaire complété, refus de participer à l'enquête ou opposition à l'appariement aux consommations de soins.

Elle est transmise par Ipsos à Santé publique France après suppression de toutes les données directement identifiantes des personnes contactées : identifiant Insee, issue de chaque contact, durée de questionnaire, mode de réponse (CATI/CAWI) etc.

➤ **Base d'identifiants** qui fait le lien entre l'identifiant du fichier CONTACT (Id) et l'identifiant généré par Ipsos (Id_Ipsos). L'identifiant généré par Ipsos (Id_Ipsos) servira uniquement à la gestion du terrain, tandis que l'identifiant initial du fichier CONTACT (Id) sera utilisé pour la livraison des données. Les données nominatives seront ainsi enregistrées dans une base distincte des bases du questionnaire contenant les réponses aux questions.

➤ **Bases questionnaires (CATI / CAWI) :**

Elle comprend l'Id de chaque individu et les données recueillies dans les questionnaires CATI ou CAWI (questionnaire comprenant eux l'Id_Ipsos) ; Elle est :

- Transmise par Ipsos à Santé publique France
- Conservée par Santé publique France pendant 10 ans; ce délai de conservation correspond à l'estimation du temps d'analyse pour le suivi de long terme de l'état de santé des populations ;

➤ **Base coordonnées répondants :**

Elle contient l'Id, les données d'identification (nom, prénom, adresse, téléphone, mail) mises à jour auprès des personnes contactées et enregistrées lors de la réalisation du questionnaire. Elle indique notamment si la personne a accepté d'être recontactée.

Elle est transmise par Ipsos à Santé publique France.

Ces données seront conservées dans un délai de 2 ans après le recueil des données, puis elles seront détruites.

▪ **La base d'analyse « Cartographie des expositions »**

Elle est construite à partir des réponses aux questionnaires concernant la description des personnes de leur exposition à l'incendie (visuelle, auditive, retombées de débris ou de suies, odeurs, panache de fumée) et des coordonnées X Y de leur logement. Cette base de données comprendra un nouveau numéro d'identifiant généré aléatoirement par Santé publique France (Id_SPF_carto)

▪ **La base d'analyses (Santé publique France) :**

Elle est construite à partir des réponses aux questionnaires, enrichies de certaines variables des bases d'échantillonnage (niveau de vie, type de ménage, type de logement, coordonnées géographiques, composition du ménage). Cette base de données comprendra un nouveau numéro d'identifiant généré aléatoirement par Santé publique France (Id_SPF). Elle contiendra des données pseudonymisées, toutes les données à caractère personnel directement identifiantes étant supprimées. Le sexe et l'âge des personnes seront conservés, ainsi que la strate d'appartenance et la variable 'distance à l'incendie'. Les coordonnées géographiques, de même que la commune de résidence, ne figureront pas dans cette base.

Seuls les agents de Santé publique France seront habilités à la constituer à partir des différents fichiers et à l'analyser.

La Base d'analyse pseudonymisée pourra être mise à disposition d'organismes de recherche dans le respect des dispositions de la loi informatique et libertés.

Tableau 7 : Description des données, stockage et durée de conservation

Catégorie de données	source	justification	Destinataires (qui accède aux données et pourquoi)	Durée de conservation
<u>Fichier Contact</u> (Données d'identification directe) : Données d'identification directe	Insee (Fideli)	Contacter les individus (Ipsos) ;	Santé publique France, puis prestataire de collecte Ipsos, pour contacter les individus sélectionnés	Par Ipsos: supprimé une fois le recueil terminé
<u>Fichier redressement</u> (poids de sondage, variables socio-économiques non identifiantes)	Insee (Fideli)	Redresser les données	Santé publique France, pour les traitements aval.	6 ans après la fin de la collecte
<u>Base de gestion et de suivi</u> (du terrain d'enquête) : Données d'identification directe	Ipsos	Gestion et suivi du terrain d'enquête	Transmission à Santé publique France après anonymisation . Utilisation des parodonnées (nombre d'appels, durée d'appel, information enquêteurs, mode de collecte etc.), de l'issue de chaque contact (refus, hors-champ, collecté etc.).	Base de gestion et de suivi d'Ipsos : conservation 12 mois après la fin de la collecte ; Base pseudonymisée de gestion et de suivi de Santé publique France : 10 ans
<u>Base questionnaire</u> : données non identifiantes (données socio-démographiques agrégées, réponse aux questions)	Ipsos	Recueil des données de l'enquête (CATI/CAWI)	Transmission à Santé publique France pour les analyses des données. Aucune donnée directement identifiante ne sera transmise dans la base questionnaire.	10 ans
<u>Base coordonnées répondants</u> (Données d'identification directe)	Ipsos	Enregistrement des refus de participation, des oppositions à des enquêtes secondaires ou à des appariements aux consommations de soins;	Santé publique France	Conservées pendant une durée 2 ans après la fin de la collecte puis détruite
<u>Base d'analyse « Cartographie des expositions »</u> (Coordonnées X Y + expositions)	Santé publique France	Réaliser une cartographie des différentes expositions déclarées	Santé publique France	10 ans
<u>Base consolidée pour analyse</u>	Santé publique France	Base contenant les réponses au questionnaire et les pondérations. Aucun données identifiantes n'est conservée	Elle pourrait être diffusée à certains destinataires sous réserve d'avoir préalablement signé une convention de mise à disposition des données définissant les droits et obligations quant à l'utilisation de ces données au regard de la loi informatique et libertés (base d'analyse pseudonymisée)	10 ans

6.8.7 Circulation des Bases de données pour le recueil par questionnaire

La circulation des données est faite en cinq étapes (figure 6)

Etape 1 : L'Insee envoie à Ipsos le fichier CONTACT.

Etape 2 : L'Insee envoie à Santé publique France le fichier REDRESSEMENT

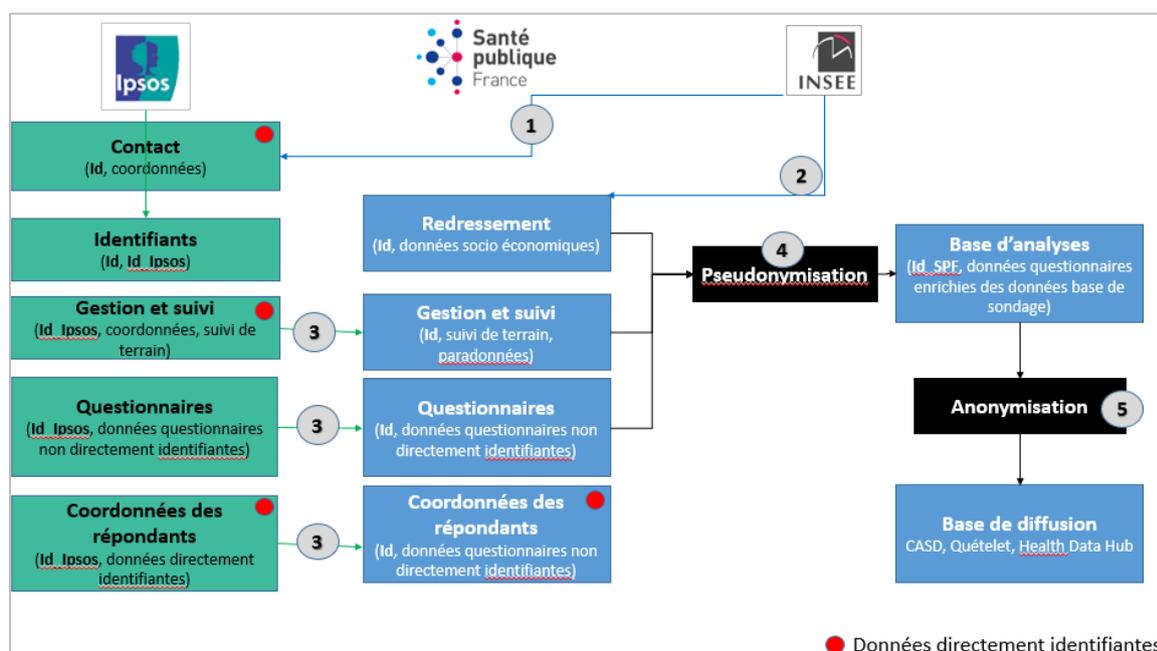
Etape 3 : Ipsos consolide et enrichit le fichier CONTACT. Il constitue pendant la collecte les bases suivantes : **Base de gestion et de suivi**, Base questionnaire, **Base coordonnées répondants**. **Ipsos transfère à Santé Publique France, à la fin de la collecte :**

- La base de gestion et de suivi sans données directement identifiantes, et intégrant des parodonnées sur le questionnaire (mode de collecte, nombre d'appels, jour d'appel etc.)
- Les bases questionnaire (CATI, CAWI) sans données directement identifiantes ;
- La base coordonnées répondants, avec données directement identifiantes ;

Etape 4 : Santé publique France stocke sur serveur sécurisé l'ensemble des bases reçus par Ipsos d'une part, l'Insee d'autre part.

Etape 5 : Après une étape d'anonymisation supplémentaire, une base de données de diffusion sera mise à disposition sur les plateformes publiques du type Quêtelet, CASD, Health Data Hub.

Figure 6 : Schéma de circulation des données



6.8.8 Mode de stockage et transmission des données

Les transmissions de fichiers comprenant des données à caractère personnel seront effectuées après chiffrement desdits fichiers, à l'aide de logiciel à l'état de l'art.

Santé publique France chiffrera ses fichiers avec le logiciel « Zed ! » et proposera cette solution à ses partenaires. Avec le logiciel « Zed ! », les données sont contenues dans une valise virtuelle cryptée dont l'ouverture nécessite des identifiants (login et mot de passe) communiqués par SMS au destinataire. Les transmissions des fichiers chiffrés seront effectuées *via* une plateforme sécurisée.

Santé publique France stockera l'ensemble des fichiers reçus par l'Insee et par Ipsos sur un serveur de données sécurisé à Santé publique France (lecteur S).

La base des coordonnées sera stockée dans un dossier séparé des bases questionnaires avec un accès limité à deux personnes. Un chiffrement (Zed !) des bases de données comprenant des données d'identification directe sera mis en place.

Santé publique France consolide les bases et constitue une base d'analyses pseudonymisée. Les données d'enquêtes, stockées sur des serveurs sécurisés, ne pourront être transmises que via des procédures sécurisées (chiffrement préalable (Zed !) et dépôt sur une plateforme sécurisé). Les fichiers comportant des données identifiantes seront conservés par santé publique France, chiffrée (Zed !). Aucune de ces données ne sera transférée en dehors de l'Union Européenne.

Ipsos, est chargé du recueil de données auprès des personnes concernées. A ce titre, il réalise les opérations suivantes :

- Hébergement et consolidation du fichier contact
- prise de contact et des relances des personnes
- création, hébergement et gestion de la Base identifiant
- Elaboration et hébergement des applications internet et interne permettant le recueil des données des questionnaires
- Gestion des droits des personnes pendant le recueil.
- Hébergement et gestion des bases de gestion et de suivi, et base répondants.

Conformément à la clause de sous-traitance reprenant les obligations de l'article 28 du RGPD et insérée dans le marché qui le lie à santé publique France, s'est engagé à réaliser l'enquête dans le respect de la réglementation RGPD. Ipsos s'engage notamment à exécuter ses obligations en termes de sécurité des systèmes d'information selon un Plan d'Assurance Qualité Sécurité, qui détaille les solutions mises en œuvre concernant le flux de données (gestion des identifiants, questionnaires sécurité des données, confidentialité, conservation, transfert de données sécurisés par des outils de transferts et de cryptage spécifiques), l'infrastructure informatique, les système de sauvegarde et de réplication et les différentes mesures de protection informatique. Notamment, les réponses par Internet sont collectées sur un site sécurisé auquel les personnes enquêtées ont accès à partir d'un identifiant et d'un mot de passe personnels qui leur sont adressés dans la lettre-avis. Ipsos a par ailleurs mis en place un système de backup quotidien avec externalisation sur un 2^{ème} datacenter dans une base dédiée et transfert sécurisé via une liaison spécialisée sécurisée. Les transferts de fichiers respectent un protocole sécurisé, recourant à la plateforme sécurisée LiquiFiles dédiée (eutransfer.ipsos.com, cryptage HTTPS via certificat GeoTrust RSA CA 2018) ou à un compte SFTP (cryptage SSH) dédié au projet sur un serveur propriété d'Ipsos et hébergé à Gentilly. Aucun fichier nominatif ne sera transmis par mail.

6.8.9 Information et droits des personnes

L'information individuelle sera réalisée dans le courrier d'invitation au moyen d'une lettre-avis et d'une plaquette d'information, ainsi que du courriel-avis. Ces informations seront également disponibles sur le site Internet de l'étude et dans les différents supports de communication envisagés, et rappelées lors de la passation du questionnaire par l'enquêteur.

L'information consistera en une présentation des objectifs et du déroulement de l'étude. Les personnes seront informées de la sélection par tirage au sort, du champ des personnes concernées, des modalités de participation ou pour manifester un refus. Il sera notamment précisé que des informations concernant les consommations de soins en ville et à l'hôpital disponible auprès de la Caisse nationale d'assurance maladie (Cnam) seront collectées sous réserve de leur non-opposition, puis analysées et traitées de manière confidentielle.

Les personnes seront informées de leur liberté de participer et de leurs droits quant à l'accès et la rectification des données recueillies.

Les refus pourront être exprimés de manière explicite au **téléphone** (numéro vert), par **mail**, par **courrier ou sur le site internet de l'étude**. Les informations utiles pour exprimer un refus seront mentionnées dans la lettre-avis, le courriel-avis, la plaquette d'information et le site internet de l'étude.

Les demandes d'exercice des droits seront adressées par courrier électronique ou par courrier au DPO d'Ipsos puis au DPO de Santé publique France.

S'agissant des enfants, le recueil des données est opéré auprès du titulaire de l'autorité parentale qui est interrogé sur l'état de santé et le ressenti de l'enfant sous sa responsabilité. Conformément à l'article 70, l'information est réalisée auprès du titulaire de l'autorité parentale ainsi contacté, titulaire qui sera informé des droits dont il dispose sur les données relatives à son enfant qu'il nous communique.

6.8.10 Utilisation des données par un prestataire extérieur

La collecte sera réalisée par l'institut de sondage Ipsos Observer, titulaire d'un marché pluriannuel concernant la réalisation d'enquêtes en méthode aléatoire pour Santé publique France. Ipsos assurera notamment :

- La mise en place de la méthode de recrutement choisie pour cette étude,
- Le suivi des inclusions,
- Le développement des masques de saisie des questionnaires,
- Le recrutement et la formation des enquêteurs de terrain,
- Le recueil des refus de participation ou d'opposition aux appariements envisagés
- Les recueils des données (questionnaires CATI et CAWI),
- Le transfert sécurisé des données à Santé publique France.

Ipsos dispose de nombreuses références dans le domaine de la statistique publique, des enquêtes de santé, et du respect des règles de sécurité dans le traitement de données à caractère personnel. La sous-traitance sera déclarée au Comité du secret au cours de la séance exceptionnelle fin mai 2020, comprenant différentes opérations préalables à la collecte : enrichissement téléphonique du fichier de contacts, impression et mise sous pli des courriers postaux...

Les conditions de réalisation du marché par le titulaire sont encadrées par un Cahier des clauses administratives particulières qui précise, dans les articles 15 et 16 notamment, l'obligation de confidentialité, de sécurité et de protection des données.

Article 15 : Confidentialité

« Le titulaire s'engage à considérer comme confidentielles toutes les informations qui lui sont communiquées comme telles par Santé publique France dans le cadre de l'exécution du marché, et s'interdit en conséquence pendant toute la durée de celui-ci, sans limitation de durée après son expiration – à condition que les informations susvisées ne soient pas tombées dans le domaine public – de les divulguer à quelque personne que ce soit, sans l'accord préalable de Santé publique France, et s'engage à faire signer par chacun de ses personnels un engagement de stricte confidentialité.

L'engagement de confidentialité porte notamment sur les fichiers et données traitées par Santé publique France. Le Titulaire s'interdit toute communication écrite ou verbale sur les sujets objet du marché, sans l'accord préalable de Santé publique France.

Toutefois, Santé publique France autorise le titulaire à faire mention auprès de tiers de la mission faisant l'objet du présent marché à titre de référence. Le non-respect de ces dispositions est considéré comme une faute de nature à conduire Santé publique France à résilier le marché aux torts du Titulaire,

aux frais et risques de ce dernier, sans préjudice des réparations éventuelles demandées par Santé publique France au Titulaire, au titre de l'article 1384 du code civil. »

Article 16 : SECURITE ET PROTECTION DES DONNEES - Cadre Légal

Le Titulaire s'engage à assurer la protection et la sécurité des Données recueillies dans le cadre du marché dans le respect des obligations posées par la loi n°78-17 du 6 janvier 1978 relative à l'informatique, aux fichiers et aux libertés, ainsi que par les législations et recommandations applicables à Santé publique France en matière de protection des données (en particulier des données de santé à caractère personnel) et de sécurité des systèmes d'information.

Le Titulaire apporte son appui à Santé publique France dans la mise en conformité du traitement de données aux obligations légales et réglementaires auxquels ce dernier est soumis.

A cet effet, il s'engage à:

- Effectuer une veille régulière pour suivre les évolutions législatives et réglementaires et les mettre en œuvre, notamment le règlement Européen n° 2016/679 sur la protection des données personnelles qui sera applicable à compter du 25 mai 2018.

- Tenir à la disposition de Santé publique France les documents relatifs à la sécurité de ses données comprenant notamment la documentation technique nécessaire, les analyses de risques et analyses d'impact relatives à la protection des Données produites et la liste détaillée des mesures de sécurité mises en œuvres.

Il est rappelé qu'en cas d'atteinte à la confidentialité et à la sécurité des données à caractère personnel, la responsabilité du Titulaire peut être engagée sur la base des dispositions des articles 226-17 et 226-5 du code pénal.

Les données à caractère personnel transmises au Titulaire ou collectées par le Titulaire lors de l'exécution du Marché sont couvertes par le secret professionnel (article 226-13 du code pénal). »

6.9 Lettres-avis, notices ou plaquettes d'information

6.9.1 Courriers, mails et SMS aux enquêtés de l'échantillon principal

Les différents modes d'informations sur l'enquête (lettre, mail, SMS) des individus sélectionnés aux ont été déclinés pour tenir compte de chaque situation :

- Connaissance ou non de l'identité de la personne sélectionnée ;
- Information initiale ou information de relance ;
- Individu sélectionné dans la zone d'étude principale (Rouen et ses environs) ou dans la zone témoin (Le Havre et ses environs)

Les informations sont dès que possible personnalisées tant que faire se peut : civilité, prénom, mot de passe d'accès privé au site internet. Lorsque l'identité de l'individu sélectionné ne sera pas connue, il sera demandé à la personne contactée de transmettre l'information (lettre, mail ou SMS) à la personne sélectionnée.

Par ailleurs, le pli de la lettre-avis, dans la zone exposée seulement, contiendra également une plaquette de présentation de l'enquête.

La lettre-avis de l'enquête est disponible en Annexe 4.

6.9.2 Outils de communication de l'enquête

Un plan de communication globale de l'enquête a été établi. Il s'appuie sur les objectifs de faire connaître l'enquête et ses objectifs en ciblant les **personnes relais auprès des populations** : maires, professionnels de santé et leur représentants, associations Rouen Respire, Métropole, Conseil départemental, PMI, maisons de quartier ;

En plus des envois aux personnes sélectionnées pour l'enquête, des courriers seront envoyés aux communes :

- Un courrier spécifique sera adressé **aux mairies de la zone d'étude**, avec des informations concernant les outils mis à disposition : affiches et dépliants A4 mis à disposition, vignettes à télécharger, informations plus complètes sur le site Internet de l'agence ;
- Un envoi sera aussi fait à **l'intention des professionnels de santé des communes** : médecins, pharmaciens, PMI, mais aussi centres médico-sociaux, associations de quartier, etc.
- Un courrier spécifique sera adressé **aux mairies de la zone témoin** (courrier n°2 + dépliant 4 pages A4), avec des informations qui permettront au maire de répondre aux éventuelles questions des citoyens.

6.9.2.1 Dépliant d'information

Exercice de votre droit d'opposition, d'accès et de rectification

Votre participation est libre et volontaire. Vous pouvez à tout moment décider de l'arrêter. Pour exercer votre droit d'opposition, ainsi que vos droits d'accès, de rectification, de suppression des données vous concernant ou de limitation de leur traitement, adressez votre demande - accompagnée de la copie d'une pièce d'identité* - par courriel à : dpo@santepubliquefrance.fr ou par courrier signé à : Santé publique France - DPO - 12 rue du Val d'Osne, 94415 Saint-Maurice Cedex

* La justification préalable de l'identité du demandeur est une obligation légale - elle permet de s'assurer que seule la personne concernée dispose de ses données et en particulier accède aux données qu'elle nous a confiées.

Plus d'informations

Numéro vert : 0 800 942 590
[service et appel gratuit / du lundi au vendredi : 9h00-18h00 ; le samedi : 10h00-16h00]
www.sante-post-incendie-76.fr : le site dédié à l'étude
Mairie de votre commune, votre médecin, votre pharmacien
www.santepubliquefrance.fr avec des informations sur les autres études du dispositif Santé Post Incendie 76



Santé Post Incendie 76

AOÛT - SEPTEMBRE

UNE ÉTUDE À L'ÉCOUTE DE VOTRE SANTÉ

Incendie industriel du 26 septembre 2019



Santé Post Incendie 76

Santé publique France mène une étude sur la santé et la qualité de vie des habitants des communes concernées par l'incendie des entrepôts de Lubrizol et de NL Logistique. Elle fait partie d'un ensemble d'études scientifiques mis en place par Santé publique France pour évaluer l'impact sanitaire de cet accident industriel : le dispositif « Santé Post Incendie 76 ».



Santé Post Incendie 76

Une étude à l'écoute de votre santé

Quels sont les objectifs de l'étude ?

Santé publique France a pour mission d'évaluer l'impact sur la santé physique et mentale des personnes qui ont pu être exposées à l'accident industriel. Cette étude va s'intéresser aux **symptômes et problèmes de santé que vous avez pu ressentir, à votre accès à l'information** pendant l'accident, ainsi qu'à **vos santé et votre qualité de vie actuelles**. Cette enquête de santé va être menée auprès d'un échantillon représentatif d'habitants des communes touchées par l'accident. Pour comparer et interpréter les résultats obtenus, l'étude est aussi réalisée dans une zone témoin située au Havre et ses environs.

Pourquoi cette étude ?

L'ampleur de l'incendie des entrepôts de Lubrizol et NL Logistique et ses conséquences potentielles pour la santé justifient de mettre en place des réponses en termes d'études de santé, au-delà des études et données environnementales. Dans ce cadre, Santé publique France a été missionnée par la Direction Générale de la santé pour évaluer l'impact sanitaire. De nombreux habitants et travailleurs de Rouen et ses environs ont pu être exposés aux bruits de l'incendie et des explosions, à la vision des flammes et du panache de fumées, à des odeurs désagréables, à des suies, à des débris de fibrociment, à des substances chimiques, au stress et à l'inquiétude, etc. Les résultats de l'étude permettront de **mieux connaître les conséquences sanitaires de cet accident industriel**.

Qui participe à la conception de cette étude ?

Les études du dispositif Santé Post Incendie 76 sont discutées au sein du **Groupe Santé**, composé d'associations locales, de professionnels de santé et d'élus de la région. Cette concertation permet d'adapter les travaux réalisés aux préoccupations de la population et de tenir compte du contexte local. Les études bénéficient également des conseils d'un **Comité d'appui thématique**. Mis en place au niveau national, ce comité indépendant est constitué de spécialistes de plusieurs disciplines scientifiques : épidémiologie, toxicologie, pathologie professionnelle, sociologie, etc.

Comment se passe l'étude ?

Une étape préparatoire

Une pré-enquête a été menée au mois de février 2020 dans 4 communes : Le Petit-Quevilly, Bois-Guillaume, Préaux et Buchy. Son but était d'ajuster le questionnaire de l'étude de santé en identifiant les sujets d'inquiétude et les attentes des personnes concernées par l'accident.

L'étude de santé et de qualité de vie

La collecte des données de l'étude est prévue en août et septembre 2020. L'étude s'adresse aux habitants des communes de Seine-Maritime qui ont été exposées à l'incendie. Elle est menée auprès de 4 000 adultes et de 1 200 enfants qui ont été tirés au sort, comme vous.

Il est possible de répondre à l'enquête de deux manières. En ligne, sur le site Internet www.sante-post-incendie-76.fr et à l'aide du mot de passe qui vous a été envoyé, ou par téléphone. Un enquêteur de la société d'études Ipsos prendra contact avec vous. Vous pouvez également convenir d'un rendez-vous sur le site Internet de l'étude ou en appelant le n° vert. Le questionnaire dure environ 30 minutes. Toutes les données recueillies seront analysées de façon anonyme. Tous les résultats de l'étude seront rendus publics.

Nous vous remercions par avance pour votre participation. Toutes vos réponses seront utiles pour améliorer les connaissances concernant les conséquences sanitaires de cet incendie.

Qui organise cette étude ?

Santé publique France est une agence publique nationale qui a pour mission la surveillance de la santé des populations, la promotion de la santé et la réponse aux crises sanitaires.

AOÛT - SEPTEMBRE

UNE ÉTUDE À L'ÉCOUTE DE VOTRE SANTÉ

Incendie industriel du 26 septembre 2019



À votre écoute

Santé publique France mène une étude auprès des habitants des communes concernées par l'incendie des entrepôts de **Lubrizol** et de **NL logistique** en les invitant à répondre à un questionnaire sur leur santé et leur qualité de vie.



Comment se déroule l'étude

La collecte des données est réalisée en août et septembre 2020 auprès d'un **échantillon représentatif** de 4 000 adultes et 1 200 enfants. Les participants à l'étude, **tirés au sort** en août, vont recevoir un courrier de Santé publique France.

Cette enquête fait partie d'un ensemble d'études mises en place pour évaluer l'impact sanitaire de l'accident industriel : le dispositif « Santé Post Incendie 76 ».

Plus d'informations

Numéro vert : 0 800 942 590
(service et appel gratuit / de lundi au vendredi : 9h00-18h00 ; le samedi : 10h00-16h00)
www.santé-post-incendie-76.fr : le site dédié à l'étude
Mairie de votre commune, votre médecin, votre pharmacien
www.santepubliquefrance.fr avec des informations sur les autres études du dispositif « Santé Post Incendie 76 ».



7 Coût des enquêtes

Des moyens internes sont déjà mis à disposition et le resteront pendant 1 an : 3 ETP, couvrant l'ensemble des opérations de mise en œuvre et de suivi de la collecte, de traitements post-collecte et d'analyses, mais également d'animation des différents groupes de concertation qui se poursuivront après la collecte. Outre le coût à prévoir pour la réalisation du tirage de l'échantillon (~8 000 euros), le coût lié à l'opération de collecte est évalué, dans le devis reçu, à 210 000 euros HT.

8 Liste des annexes

8.1 Avis du comité du secret statistique



N° dossier scss-3486-1

Fin d'habilitation :

Comité du Secret Statistique
Consultation électronique du 22 juin 2020
Point M1113

Demande d'accès à des données confidentielles pour un nouveau projet

Projet : Lubrizol : enquête de santé déclarée

Chercheurs :

- RICHARD Jean-Baptiste, Directeur, SpF (*responsable scientifique*)
- BLINEAU VALERIE, Directeur, IPSOS OBSERVER
- SAOUDI Abdessatar, Chargé d'études / Chargé de mission, Santé publique France
- ZEGHNOUN Abdelkrim, Chargé d'études / Chargé de mission, Santé publique France

Source obtenue :

DGFIP/INSEE	Echantillon Fideli (Echant_Fideli)
-------------	------------------------------------

Le comité du secret donne un avis favorable à la demande M1113.

8.2 Avis d'examen du comité du Label



COMITÉ DU LABEL
DE LA STATISTIQUE PUBLIQUE

Montrouge, le 16 juillet 2020
N°2020_17077_DG75-L002

AVIS D'EXAMEN

Enquête « Santé post incendie 76 : une étude à l'écoute de votre santé »

Service producteur : Santé publique France

Opportunité : avis favorable émis le 28 novembre 2019 par la Commission « Démographie et questions sociales »

Réunion du Comité du label du 11 juin 2020 - commission « Ménages » (séance en webconférence)

Commission	Ménages
Type d'avis	Avis d'examen
Label d'intérêt général et de qualité statistique	Non
Caractère obligatoire	Non
Période de validité	Année 2020
Publication JO	Oui

Descriptif de l'opération

En octobre 2019, le directeur général de la santé a saisi Santé publique France afin d'évaluer l'impact à moyen terme sur la santé, de l'incendie survenu le 28 septembre 2019 sur les sites de l'usine Lubrizol et Normandie Logistique situés à Rouen.

En réponse, Santé publique France propose de mettre en œuvre un dispositif de surveillance des effets sanitaires potentiels à court, moyen et long terme s'appuyant sur 4 approches différentes et complémentaires :

1/ Réaliser une étude de santé déclarée en population, menée auprès d'un échantillon représentatif de la population exposée aux conséquences de l'incendie, dont l'objectif est de décrire les impacts liés à l'événement, et consécutifs à l'ensemble des expositions, nuisances et stimuli subis par la population pendant et dans les suites de l'incendie. Seront étudiés les effets immédiats (événements de santé, nuisances, anxiété, etc.) ou à court terme, et/ou n'ayant pas forcément donné à lieu à une prise en charge médicale.

2/ Mettre en place pour la population exposée un suivi à moyen et long terme de l'état de santé, au travers d'une analyse à intervalles réguliers d'indicateurs de santé, destiné à décrire l'évolution dans le temps, de la fréquence des pathologies pouvant être en lien avec les effets à moyen et long termes des nuisances associées à l'incendie.

3/ Proposer en lien avec les services de médecine du travail une méthodologie pour le suivi longitudinal des salariés des deux entreprises et des professionnels intervenus sur site lors de l'incendie et des jours qui ont suivi l'incendie pour le nettoyage, les prélèvements, etc.

4/ Mener une étude d'imprégnation si les résultats des mesures de polluants dans l'environnement montrent un risque d'exposition de la population à certaines substances toxiques dispersées par le panache.

RÉPUBLIQUE FRANÇAISE

Institut National de la Statistique et des Études Économiques – 88 avenue Verdier – CS 70058 - 92541 MONTROUGE CEDEX - FRANCE - www.insee.fr
Tél : 01.87.89.50.00 - N° SIRET : 120 027 018 00583 - Code APE : 8411Z - Service Insee Contact : 09 72 72 4000 - (tarification "appel local")

Les objectifs spécifiques de cette enquête sont d'analyser :

- les symptômes et nuisances ressentis par la population pendant et dans les suites de l'incendie ;
- les indicateurs de santé perçue (la qualité de vie, l'anxiété et les troubles du sommeil) ;
- la symptomatologie clinique déclarée (irritations, maux de tête, nausée...) selon différentes zones de gradient d'exposition définies à partir des prélèvements environnementaux ;
- la santé perçue dans la population exposée à l'incendie à des références nationales ;
- la santé perçue selon différentes zones d'exposition aux nuisances générées par l'accident industriel. Ces différentes zones d'exposition sont définies en tenant compte de (1) la modélisation du panache des fumées, (2) la perception des odeurs et (3) le bruit et des stimuli visuels ;
- la relation entre l'intensité de l'exposition et les événements de santé déclarés ;
- les attitudes et perceptions de la population face aux recommandations des pouvoirs publics en situation de crise pour identifier les points d'amélioration dans la prise en charge, l'information ou la compréhension des messages.

La population cible est celle des communes du département de Seine-Maritime exposées à l'incendie (123 communes environ 360 000 personnes). Le champ de l'enquête recouvre les adultes et les enfants résidant sur la zone exposée.

La collecte sera effectuée par la société Ipsos Observer auprès de 5 000 individus entre août 2020 et octobre 2020. La durée moyenne d'interrogation ne devrait pas excéder 30 minutes.

La collecte sera réalisée selon un multimode séquentiel : d'abord sur Internet puis par téléphone. Les données administratives fournies par l'Insee seront utilisées pour contacter les personnes à enquêter.

Le protocole de l'enquête, notamment le questionnaire et la zone d'étude, a été défini en concertation avec le « Groupe santé » issu du comité de transparence et de dialogue, et constitué de représentants d'associations, de professionnels de santé, d'élus, de collectivités locales et d'universitaires.

Les protocoles scientifiques sont validés par un comité d'appui thématique (conformément à la comitologie de Santé publique France) constitué de personnes qualifiées (toxicologue, sociologue, cliniciens, épidémiologiste, statisticien, métrologiste). Ce comité a pour rôle de valider les choix scientifiques, d'aider à l'interprétation des résultats et à la formulation de recommandations.

Un groupe de travail associant la Drees permet de définir et de valider la méthodologie de l'enquête.

Les premiers résultats sont prévus en début 2021.

L'avis d'opportunité favorable a été délivré pour la première approche de ce dispositif, à savoir la réalisation d'une étude sur la santé déclarée auprès d'un échantillon représentatif de la population exposée.

Le Comité du label de la statistique publique émet les recommandations ou observations suivantes :

Remarques générales

- Le Comité du label félicite le service pour les efforts déployés afin d'assurer la plus grande publicité à l'opération auprès des différents relais dans la population des zones concernées, dans le but notamment de favoriser la réponse à l'enquête (zone d'étude et zone témoin).
- La préparation de l'enquête, déjà contrainte par un calendrier très serré, a été affectée par la crise du Covid-19. Ces circonstances particulières expliquent en partie le caractère incomplet du dossier présenté, notamment l'absence de tests. Le Comité du Label attire l'attention du service sur les possibles biais de mémoire. Il suggère au service de veiller à ce point lors des tests, de façon à identifier les questions sujettes à des effets de mémoire qui rendraient leur exploitation délicate.

- Le Comité du label note que l'enquête de santé déclarée constitue une partie du dispositif d'évaluation mis en place pour évaluer les effets de l'incendie. L'avis d'examen porte sur la seule enquête détaillée dans le dossier, à l'exclusion des appariements et des post-enquêtes quantitatives envisagés. Le Comité attire l'attention du service sur les ambiguïtés du périmètre du dossier sur la question des appariements et la nécessité de clarifier ce point (cf. sur ce point le memento de la Cnil transmis au service).

Methodologie

- Le Comité du label estime que, compte tenu des objectifs de l'enquête et de la sensibilité de la question abordée pour la population locale, il est très important de garantir la représentativité de l'enquête et de contrôler les différents biais susceptibles de nuire à la crédibilité des résultats. Il n'est pas assuré que les traitements actuellement prévus dans le dossier rendront possibles l'atteinte de cet objectif. Le Comité invite donc le service à conduire une réflexion approfondie sur l'évaluation et le redressement du biais d'auto-sélection ou de non-réponse non ignorable, et un examen rigoureux des écarts éventuels entre l'échantillon d'étude et l'échantillon témoin afin de les contrôler au mieux. Il encourage le service à intégrer dans son Comité d'appui thématique toutes les compétences nécessaires à ces objectifs. L'usage des données du SNDS aurait pu constituer un moyen de contrôler ces biais, mais en l'état du dossier elles ne seront pas disponibles pour les premières exploitations de l'enquête.
- Le Comité du label note par ailleurs que certains traitements statistiques ambitieux sont envisagés, en termes de partage des poids dans le cas des enfants dont les parents sont séparés, et concernant l'existence d'effets de mode. Il encourage le service à réaliser ces travaux, tout en soulignant la nécessité de prioriser les travaux sur les effets d'auto-sélection mentionnés supra, ainsi que sur les calculs de précision effectifs après prise en compte des résultats de l'enquête en termes de nombre de répondants.

Collecte

- Le Comité du label note que le protocole de collecte semble avoir évolué récemment en privilégiant la séquence CAWI puis CATI, contrairement à la version initiale du dossier présentée au Comité.
- Le Comité du label formule quelques remarques complémentaires à celles du prélabel concernant les lettres-avis :
 - Il conviendrait de clarifier l'articulation des protocoles CAWI et CATI.
 - Dans l'esprit d'une ouverture plus large des données, il conviendrait d'élargir le périmètre des personnes ayant accès aux données (pas uniquement les professionnels de santé).
 - La question de l'appariement aux données de santé est prématurée dans la lettre-avis, ou alors il faudrait revoir le périmètre de l'opération. Il est rappelé qu'il restera à instruire comment cet accord d'appariement sera obtenu dans le cadre du dispositif plus complet.
 - La description de l'origine de l'échantillon (base administrative) est trop laconique : les lettres-avis des enquêtes utilisant Fidéli se réfèrent à la « base des logements constituée par l'Insee à partir de sources administratives ».
 - Les formulations retenues (comme « population de référence » pour l'échantillon témoin restent difficilement compréhensibles pour les non-spécialistes.
 - Enfin, la forme des courriers a été jugée un peu trop impersonnelle dans un contexte de forte attente de la population de la zone d'étude.

Diffusion

- Le Comité du label rappelle que, s'agissant d'enquêtes de la statistique publique, il convient d'assurer un accès aux données collectées le plus large possible, compatible avec les règles de protection des données. Il encourage le service à développer l'information auprès des chercheurs de toutes disciplines, notamment ceux des sciences sociales, sur l'existence de l'enquête et les procédures d'accès aux données. Au minimum, il convient de mettre en place sur des dispositifs généralistes comme le réseau Quetelet, ou le CASD, des liens renvoyant au site de l'enquête sur Santé publique France. Une mise à disposition des fichiers –

éventuellement après une anonymisation complémentaire – sur ces centres de données favoriserait cet usage inter-disciplinaire.

CNIL

- Les remarques écrites de la Cnil ont été évoquées en séance et transmises aux services. L'attention du service est attirée sur le fait que les informations transmises à la Cnil dans le cadre de la commission semblaient insuffisamment précises et susceptibles d'engendrer des confusions qui, en allongeant le traitement du dossier, pourraient entraîner des retards incompatibles avec le calendrier très ambitieux de l'enquête.

Le Comité du label de la statistique publique émet un avis d'examen favorable pour l'**enquête Santé post incendie 76**, et propose l'inscription de l'enquête au programme des enquêtes de la statistique publique, avec attribution d'un numéro de visa.

Cet avis est **valable pour 2020**.

La présidente du comité du label de la
statistique publique



Nicole ROTH

8.3 Précisions sur le calcul du nombre de sujets nécessaires

De manière résumée, le calcul de la taille d'échantillon a été réalisé en tenant compte des éléments suivants :

- Mise en évidence d'écarts non ajustés, tenant compte d'un effet plan de 1,5 ;
- Application d'un facteur multiplicatif pour mettre en évidence des écarts ajustés ;
- Contraintes budgétaires ;

L'hypothèse d'un effet plan de 1,5 est en effet élevée si on ne prend en considération que le de plan de sondage de l'étude, sans l'éventuel impact de la dispersion des poids dû au phénomène de non réponse totale mais aussi du redressement. Nous n'avons pas d'élément à ce jour pour juger du taux de participation qui sera observé au sein des différentes strates, et surtout d'éventuels taux de participations différenciés selon différents segments de population : personnes peu diplômées, revenus les plus faibles, personnes les plus jeunes et plus âgées. Cet effet plan est celui estimé dans certaines enquêtes portées par Santé publique France, comme le Baromètre santé, dans laquelle ces catégories de population sont sensiblement sous-représentées.

L'estimation du nombre de sujets nécessaires a été réalisée en utilisant différentes publications scientifiques^{38, 39}. Nous avons d'abord calculé une taille d'échantillon n_1 pour une analyse bivariée permettant de mettre en évidence un écart de 10% entre les prévalences de 2 zones d'exposition ($n_1=586$ au maximum). Cette taille d'échantillon a été ensuite augmentée en multipliant par le facteur d'inflation de la variance (VIF) due à l'introduction d'autres covariables (de risque ou d'ajustement) dans le modèle multivarié afin d'obtenir la taille d'échantillon dans le cas ajusté ($n_2=838$). Nous avons fait l'hypothèse d'un VIF égal à 1,43 ($R^2=30\%$), ce qui est une valeur relativement faible, les VIF associés à certaines variables peuvent être plus élevés. Cet ajustement vise à prendre en compte l'inflation de la variance des paramètres des variables explicatives (i.e., les régresseurs), due à la colinéarité entre les variables explicatives. La taille de l'échantillon est augmentée dans le modèle multivarié pour pallier à cette inflation de variance.

Model bivarié : $Y = a_0 + b_1 X_1$

X_1 variable explicative d'intérêt

V_{1_bivar} = Variance (b_1) dans le modèle bivarié

Model multivarié : $Y = a_0 + b_1 X_1 + b_2 X_2 + \dots + b_p X_p$

$V_{1_multivar}$ = Variance (b_1) dans le modèle multivarié

L'ajustement sur les variables explicatives $X_2 \dots X_p$ engendre une inflation de la variance V_{1_bivar} du paramètre b_1 associé à variable explicative X_1 . Cette inflation de la variance est due à la colinéarité entre la variable X_1 et les autres variables explicatives ajoutées dans le modèle. La variance $V_{1_multivar}$

³⁸ Hsieh FY. Sample size tables for logistic regression. Stat Med 1989;8(7):795-802. Hsieh FY, Bloch DA, Larsen MD. A simple method of sample size calculation for linear and logistic regression. Stat Med 1998;17(14):1623-34.

³⁹ Hsieh FY, Lavori PW, Cohen HJ, Feussner JR. An overview of variance inflation factors for sample-size calculation. Eval Health Prof 2003;26(3):239-57.

du paramètre b_1 de la variable explicative X_1 dans le modèle multivarié peut être approchée en augmentant la variance V_{1bivar} par un facteur égal à $1/(1-R^2)$:

$$V_{1multivar} = V_{1bivar} \times 1/(1-R^2) \quad (1)$$

Le facteur $1/(1-R^2)$ est appelé facteur d'inflation de la variance ou VIF. R^2 représente la proportion de la variance de X_1 expliquée par le modèle de régression incluant les variables explicatives $X_2 \dots X_p$.

Le calcul de la taille de l'échantillon dans le cas d'un modèle multivarié utilise des formules complexes. Pour simplifier les calculs, la taille de l'échantillon n_{bivar} est calculée dans le cas bivarié. Ce qui revient à calculer la taille de l'échantillon dans le cas d'une comparaison de deux moyennes observées (X_1 quantitative) ou deux pourcentages observés (X_1 qualitative). Les formules utilisées dans ce cas de figure sont très simples et largement utilisées. La taille de l'échantillon dans le cas multivarié $n_{multivar}$ est alors approchée en augmentant la taille de l'échantillon dans le cas bivarié par le facteur d'inflation de la variance (VIF) due à l'ajustement par les variables explicatives X_2, \dots, X_p , comme dans l'équation (1) ci-dessus : $n_{multivar} = n_{bivar} \times VIF$.

Estimation d'une prévalence

Sur l'ensemble de la zone exposée, 3265 individus sont nécessaires pour estimer une prévalence entre 5 et 15% avec une précision relative située entre 10 et moins de 20% (tableau 1).

Prévalence attendue	Précision relative (%)				
	5%	10%	15%	20%	25%
5%	43 794	10 949	4 866	2 737	1 752
10%	20 745	5 186	2 305	1 297	830
15%	13 061	3 265	1 451	816	522
20%	9 220	2 305	1 024	576	369
25%	6 915	1 729	768	432	277
30%	5 378	1 345	598	336	215
35%	4 281	1 070	476	268	171
40%	3 457	864	384	216	138
45%	2 817	704	313	176	113
50%	2 305	576	256	144	92
55%	1 886	471	210	118	75
60%	1 537	384	171	96	61
65%	1 241	310	138	78	50
70%	988	247	110	62	40

Tableau 1 – Nombre de sujets nécessaire pour estimer une prévalence avec une précision souhaitée ($\alpha=5\%$, $1-\beta=80\%$, effet plan=1,5)

Comparaison des prévalences entre 2 zones d'exposition

Pour mettre en évidence un écart de 10% non ajusté entre les prévalences de deux zones, l'effectif nécessaire varie entre 57 et 586 individus par zone en fonction de la valeur de la prévalence attendue.

Prévalence attendue dans la zone i	Ecart entre la zone i+1 et la zone i				
	5%	10%	15%	20%	25%
5%	116				
10%	635	57			
15%	1 019	198	37		
20%	1 352	292	104	27	
25%	1 636	371	144	66	21
30%	1 872	437	178	88	47
35%	2 061	491	205	106	61
40%	2 203	533	227	120	71
45%	2 297	562	243	131	80
50%	2 344	580	253	139	86
55%	2 344	586	259	143	90
60%	2 297	580	259	145	92
65%	2 203	562	253	143	92
70%	2 061	533	243	139	90

Tableau 2 – Comparaison de la prévalence entre 2 zones ($\alpha=5\%$, $1-\beta=80\%$, effet plan=1,5)

8.4 Lettre-avis

Vos contacts

Tél. : 0800 942 590 (Numéro vert et gratuit)

Courriel : contact@sante-post-incendie-76.fr

Site internet : www.sante-post-incendie-76.fr

Votre identifiant : **<IDENTIFIANT>**

Votre mot de passe : **<MDP>**

<key>

<Prénom> <nom>

<Adr1>

<cp> <ville>

Objet : Santé Post Incendie 76 : une étude à l'écoute de votre santé

<civilité>,

L'agence Santé publique France, missionnée par le ministère chargé de la Santé, réalise une étude de santé auprès des habitants des communes touchées par l'incendie industriel de Lubrizol et NL Logistique, survenu le 26 septembre 2019. Cette enquête s'intéresse à votre perception des pollutions générées par l'incendie, aux symptômes et problèmes de santé que vous avez ressentis, à votre accès aux informations au cours de l'événement, ainsi qu'à votre santé actuelle et votre qualité de vie. **Vous avez été tiré au sort pour participer à l'étude**, et vous avez reçu ou allez peut-être recevoir un mail ou un SMS à ce sujet.

Cette enquête est reconnue d'intérêt général par le Conseil national de l'information statistique (Cnis).

Votre participation est essentielle afin de permettre la prise en compte de la diversité des caractéristiques de la population interrogée et d'assurer ainsi la qualité des résultats de l'étude.

Vous pouvez répondre au questionnaire :

- **En ligne**, en vous connectant à l'adresse sécurisée ou en scannant le QR code ci-dessous.

<https://etude.sante-post-incendie-76.fr>

Mot de passe : **[MDP]**

QR
code

Cette adresse est à inscrire dans la barre d'adresse de votre navigateur. Puis vous saisissez le mot de passe indiqué ci-dessus (merci de faire attention à bien respecter les caractères : minuscules / majuscules)

- **Par téléphone**, un enquêteur de l'institut Ipsos prendra prochainement contact avec vous pour vous interroger, avec ce numéro : **02 36 80 10 52**. Vous pouvez également convenir d'un rendez-vous en appelant le **0800 942 590** ou sur le site **www.sante-post-incendie-76.fr**, en saisissant votre mot de passe.

A la suite de l'enquête, que vous ayez ou non répondu au questionnaire, sous réserve de votre non-opposition, des données sur vos consommations de soins qui sont disponibles auprès de la Caisse nationale d'assurance maladie (Cnam)

pourront être recueillies. Comme vos réponses à l'enquête, elles seront traitées de manière confidentielle, en respectant votre vie privée.

Vous trouverez au verso de cette lettre, des informations sur le traitement de données mis en œuvre dans le cadre de cette étude ainsi que les modalités d'exercice de vos droits.

Je vous donne l'assurance que, comme la loi en fait la plus stricte obligation, vos réponses resteront confidentielles et serviront uniquement à l'établissement de statistiques.

En vous remerciant par avance de votre participation, je vous prie d'agréer, <civilite>, l'expression de ma considération distinguée.

Dr Jean-Claude Desenclos
Directeur scientifique de Santé publique France



Comment vos coordonnées ont été obtenues ?

Vos coordonnées ont été obtenues à partir des bases de logements de l'Insee qui sont constituées à partir de sources administratives.



Que deviennent les informations que vous donnerez ?

Les résultats de l'enquête seront rendus publics à partir de fin 2020. Les données obtenues seront très précieuses pour évaluer l'impact de cet incendie industriel, prendre d'éventuelles mesures de santé publique et améliorer la gestion sanitaire d'autres accidents industriels.

Vu l'avis favorable du Conseil national de l'information statistique, l'enquête **Santé post incendie 76 : une étude à l'écoute de votre santé** a obtenu le visa n°2020X712SA du Ministre de l'économie et des finances, valable pour l'année 2020 – Arrêté en date du 03/07/2020.

Confidentialité et traitement de vos données à caractère personnel

Cette enquête met en œuvre un traitement de données à caractère personnel fondé sur la réalisation d'une mission d'intérêt public et dont Santé publique France a la responsabilité. Le règlement général 2016/679 du 27 avril 2016 sur la protection des données (RGPD), la loi n°78-17 du 6 janvier 1978 relative à l'informatique, aux fichiers et aux libertés, ainsi que la loi n°51-711 du 7 juin 1951 relative à l'obligation, la coordination et le secret en matière de statistiques, s'appliquent à la présente enquête.

À cette fin, les données suivantes vous concernant ou concernant votre enfant seront recueillies dans la mesure où ces données sont nécessaires aux objectifs visés : données démographiques et sociales, données de santé (incluant symptômes, recours aux soins), situations professionnelles et financières actuelles, consommation de tabac et d'alcool. Ces données ne feront pas apparaître vos noms et prénoms et seront associées à un identifiant anonyme.

Les réponses à ce questionnaire sont protégées par le secret statistique et destinées à Santé publique France.

Le prestataire Ipsos est seul destinataire des données d'identification (nom et coordonnées), habilité par le Comité du secret de la statistique publique, et les Archives de France, conformément aux dispositions de l'article L. 212-4 du code du patrimoine. Elles seront conservées par Ipsos pendant 3 mois à compter de la fin de la collecte, et par Santé publique France pendant 2 ans si vous acceptez de les communiquer pour être recontacté à des fins d'enquêtes ultérieures.

Cette enquête n'est pas obligatoire.

Sous réserve de votre non-opposition, seront collectées auprès de l'Assurance maladie des données concernant le suivi de votre santé (médicaments, consultations, hospitalisations, examens complémentaires). Ces données ainsi que vos réponses au questionnaire seront conservées pendant 10 ans à compter de la fin de la collecte pour les besoins de l'enquête par Santé publique France. Elles seront archivées au-delà de cette durée. À tout moment, leur usage et leur accès seront strictement contrôlés et limités à l'élaboration de statistiques ou à des travaux de recherche scientifique.

L'analyse de ces données sera effectuée par Santé publique France ou des organismes de recherche, à partir de données ne permettant pas de vous identifier directement (exclusion du nom, prénom, adresses).

Exercice de vos droits d'opposition, d'accès et de rectification

Conformément à la loi, vous pouvez exercer un droit d'accès, d'opposition, de rectification, d'effacement ou de limitation de traitement pour les données vous concernant ou concernant votre enfant pendant la période de conservation des données d'identification. Ces droits sont applicables à la collecte de données, dans le cadre de l'enquête, à l'appariement ultérieur des données recueillies avec des données de l'Assurance Maladie et aux utilisations ultérieures de ces données. Ces droits peuvent être exercés auprès de l'enquêteur qui vous contactera, ou en vous adressant par courriel à : dpo@santepubliquefrance.fr ou par courrier à : Santé publique France - DPO - 12 rue du Val d'Osne, 94415 Saint-Maurice Cedex. Pour toute question relative au traitement de vos données, vous pouvez contacter le délégué à la protection des données de Santé publique France à l'adresse dpo@santepubliquefrance.fr. Vous pouvez, si vous l'estimez nécessaire, adresser une réclamation auprès de la CNIL.