



Conseil national
de l'information statistique

COMITÉ DU LABEL
DE LA STATISTIQUE PUBLIQUE

Montrouge, le 18 février 2021
N°2021_4130_DG75-L002

AVIS D'EXAMEN

Enquête EPIdémiologie et Conditions de Vie liées au COVID 19

EpiCov - vague T2

Service producteur : Co-Maîtrise d'ouvrage Direction de la recherche, des études, de l'évaluation et des statistiques (Drees) - Inserm (Centre d'épidémiologie et de santé des populations (CESP), Inserm U 1018, 94800 Villejuif).

Opportunité : Décision d'opportunité émise le 14 avril 2020, dans le cadre de la procédure d'urgence prévue en application de l'article 2 du décret n° 2009-318 du 20 mars 2009 relatif au Conseil national de l'information statistique, au comité du secret statistique et au comité du label de la statistique publique – commission « Démographie et questions sociales ».

Consultation télématique du Comité du label du 17 août 2020 au 11 septembre 2020 – commission « Ménages ».

| | |
|---|-------------------------------|
| Commission | Ménages |
| Type d'avis | Avis d'examen |
| Label d'intérêt général et de qualité statistique | Non |
| Caractère obligatoire | Non |
| Période de validité | 4 ^e trimestre 2020 |
| Publication JO | Oui |

Descriptif de l'opération

Le projet EpiCov est porté par l'Inserm et la Direction de la recherche, des études, de l'évaluation et des statistiques (Drees) du ministère des Solidarités et de la Santé, en lien avec leurs partenaires (Insee, Santé publique France, CNRS, Ined, Université Paris-Saclay). Il est constitué d'une enquête statistique en plusieurs vagues, visant à fournir une cartographie du statut immunitaire de la population et de sa dynamique, sur l'ensemble du territoire, via la collecte d'échantillons biologiques couplée à des questionnaires.

L'objectif du projet EpiCov est de fournir d'une part, une cartographie précise du statut immunitaire de la population sur des paramètres de santé, de conditions de vie et d'inégalités sociales, et d'autre part, un suivi de la dynamique épidémique à court, moyen et long terme. Le déploiement dans des délais exceptionnellement rapides d'une cohorte de surveillance épidémique à grande échelle et

statistiquement représentative à l'échelon départemental pourra nourrir les modélisations de l'épidémie. La taille de l'échantillon et le questionnaire doit permettre d'identifier et d'étudier la situation vis-à-vis de l'épidémie pour différents domaines de diffusion d'intérêt particulier : zones géographiques (départements ou métropoles particulièrement exposées à la maladie par exemple), groupes sociaux, etc.

La **vague 1** (dite T1) du projet repose sur une grande enquête nationale auprès d'un échantillon représentatif, sélectionné aléatoirement par l'Insee, d'environ 350 000 personnes de 15 ans ou plus (avec un objectif initial de 200 000 répondants), résidant sur le territoire. L'ensemble des départements français (à l'exception de la Guyane et de Mayotte) sont inclus dans l'enquête.

Les personnes sont invitées à répondre à un questionnaire (en ligne ou téléphonique) d'une durée de 20 à 30 minutes et, en parallèle, pour une partie d'entre elles (sous réserve d'acceptation), à réaliser à leur domicile un prélèvement de quelques gouttes de leur sang, qui sera renvoyé par la poste pour déterminer s'il y a eu contact avec le virus. Les réponses seront traitées de manière confidentielle par les équipes de recherche, dans le respect de la réglementation en vigueur (secret statistique, CPP, Cnil, RGPD).

La première vague de collecte a eu lieu entre le 2 mai et le 1^{er} juin. Les premiers résultats à partir des réponses aux questionnaires et des sérologies ont été publiés à l'échelle nationale en octobre 2020, plus tardivement que prévu en raison du délai nécessaire pour obtenir les résultats des analyses sérologiques et pour interpréter convenablement les données des questionnaires.

La **deuxième vague** de l'enquête (dite T2) aura lieu du 26 octobre à fin novembre 2020, en réinterrogeant les personnes ayant répondu en vague 1 (135 000 répondants). L'ensemble des personnes interrogées se verra proposer un kit d'auto-prélèvement pour réaliser une analyse sérologique, et une partie d'entre elles, échantillonnée aléatoirement à partir du plan de sondage réalisé pour la vague 1, recevra un kit pour chaque membre de son foyer âgé de 6 ans ou plus : cette analyse croisée des sérologies au sein d'un ménage permettra d'estimer l'ampleur des contaminations intra-familiales.

Cette deuxième vague de l'enquête EpiCov a pour objectif la mesure précise de la prévalence de la maladie sur le territoire français à l'automne 2020, avec une précision départementale. Cette photographie du statut épidémique sera ensuite croisée avec les différentes thématiques traitées par le projet, permettant d'apporter des éclairages complémentaires sur la façon dont l'épidémie affecte de manière différenciée la population selon l'âge, la profession, le niveau de vie, les conditions d'habitation, etc. L'impact des mesures sanitaires sur la santé physique et mentale ainsi que sur les comportements fera également l'objet d'analyses approfondies.

Les 135 000 participants à la première vague seront recontactés afin de répondre une nouvelle fois à un questionnaire d'une trentaine de minutes, par Internet ou par téléphone. À l'issue de ce questionnaire, il leur sera proposé de recevoir du matériel pour réaliser un auto-prélèvement sanguin qui servira à tester leur statut immunitaire. L'ensemble des participants pourront réaliser ce test, et une partie d'entre eux auront l'occasion d'obtenir un kit pour tester la sérologie des autres membres de leur ménage (âgés de 6 ans ou plus). Ces prélèvements ne seront pas obligatoires.

Les résultats des analyses sérologiques portant sur cette deuxième vague d'enquête ne seront pas disponibles avant mars 2021, étant donné le volume très important d'analyses à réaliser (l'ordre de grandeur pour le nombre de prélèvements estimé est de 120 000).

L'opération EpiCov pourra être répétée une ou plusieurs fois encore pour suivre la dynamique de l'épidémie et l'évolution des conditions sanitaires et sociales dans le pays, dans tous les territoires, toutes les classes d'âge, tous les grands groupes sociaux.

* * * * *

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

- La deuxième vague de la cohorte EpiCov a été présentée au Comité du label dans les mêmes conditions d'urgence que la première vague et les remarques préalables faites dans l'avis d'examen relatif à la vague 1 peuvent être répétées dans le présent avis. Dans ce contexte, le Comité délivre à l'enquête EpiCov vague 2 **un simple avis d'examen** et non pas un avis de conformité plein et entier.
- Devant l'annonce de vagues ultérieures qui seraient présentées dans les mêmes conditions, le Comité s'interroge sur le fait de maintenir pour ces futures vagues le statut « d'enquête de la statistique publique », le traitement dérogatoire aux procédures standard d'examen de ces enquêtes ne pouvant se justifier indéfiniment.
- Le Comité du label note que la publication des premiers résultats relatifs à la première vague a été notablement retardée, en raison des délais d'obtention des résultats des tests et de la nécessité d'assurer l'expertise des résultats, qui s'est avérée plus complexe que prévu initialement. L'objectif initial d'utiliser l'enquête en lien direct avec les décisions publiques de confinement a été *de facto* abandonné.
- Le Comité félicite les services de prévoir une mise à disposition large des micro-données dans le délai de 9 mois après la disponibilité des résultats des tests, dans le respect des procédures *ad hoc*. Il encourage la maîtrise d'ouvrage à en assurer la publicité et mettre à profit ce délai pour assurer les contrôles aval et la documentation des données. **Il souhaite être informé des modalités précises de cette mise à disposition.**
- Le Comité observe que les analyses déjà réalisées sur la première vague montrent de forts risques de biais liés à la sélection des répondants en lien avec le thème de l'enquête, et la nécessité de redresser de façon parfois substantielle les résultats. L'ampleur de l'effet du redressement sur les résultats de la première vague justifie qu'on y consacre tous les moyens nécessaires. Le Comité note que l'Insee ne devrait plus contribuer substantiellement à ces travaux lors de la deuxième vague et demande aux services de **veiller à poursuivre les investissements sur la correction de la non-réponse.**
- Le Comité note que les protocoles multimodes d'articulation CATI / CAWI seront modifiés par rapport à la première vague, afin de prendre comme méthodologie de référence le multimode séquentiel. Le séquençage sera échelonné de plusieurs manières, afin de conserver des possibilités de correction s'appuyant sur les différences selon les lots de taux de réponse par mode (un lot monomode CAWI sera notamment inclus). Ces évolutions doivent permettre la modélisation pour le traitement de la non-réponse.
- Le Comité du label encourage vivement les services à réaliser l'appariement des données EpiCov avec le Système national des données de santé (SNDS), qui pourra fournir de nouvelles possibilités de correction de la non-réponse (ou de validation des redressements effectués), en fournissant des données à la fois sur les personnes interrogées et non-interrogées. Il demande à ce que **l'ensemble de ces travaux méthodologiques fassent l'objet d'une documentation publique dans des supports appropriés.**
- S'agissant du questionnaire de la deuxième vague, les suggestions émises par le Comité du label dans son rapport de prélabel ont pour partie été prises en compte et les réponses ont permis de documenter l'origine des questions. Le Comité souligne cependant que son examen a été réalisé *a minima*, sur la base de documents provisoires fournis dans des délais très serrés.
- Le Comité du label prend note des arguments des services pour le maintien du module "Suicide" introduit dans le questionnaire de vague T2, tout en regrettant que ce point n'ait pas fait l'objet d'une discussion au sein du Conseil scientifique. Il prend note que les questions spécifiques aux comportements suicidaires sont identiques à celles des baromètres Santé de

Santé Publique France. Il prend note du protocole de formation des enquêteurs pour accompagner la passation de ces questions délicates. Le Comité reste cependant circonspect sur la capacité de corriger les biais de sélection sur ce sujet dès lors que les redressements de la non-réponse en première vague n'ont pas intégré cet objectif. Il demande aux services de veiller particulièrement à l'interprétation des résultats sur ce thème particulier et à confronter ses analyses aux données externes disponibles.

- En raison des délais d'examen réduits, la Cnil a formulé ses observations directement aux services dans le cadre de la procédure de demande d'autorisation.

Le Comité du label de la statistique publique émet un avis d'examen favorable pour l'enquête EpiCov T2 (vague 2) de la Drees et de l'Inserm, avec inscription au programme des enquêtes de la statistique publique.

Cet avis est valable pour le quatrième trimestre 2020.

La présidente du comité du label de la
statistique publique

Signé : Nicole ROTH