

UNE ENQUÊTE EN CRISE : LE PROTOCOLE DE L'ENQUÊTE EPICOV

Guillaume BAGEIN (*), Nathalie BAJOS (**), Muriel BARLET (*), François BECK (***/****),
Thomas DERUYON (*), Aude LEDUC (*/*/*/*), Nicolas PALIOD (****), Philippe RAYNAUD (*),
Patrick SILLARD (****), Josiane WARSZAWSKI (**), équipe EpiCov

(*) DREES ; (**) Inserm ; (***) Santé Publique France ; (****) INSEE

guillaume.bagein@sante.gouv.fr

Mots-clés : Protocole, Covid-19, population générale

Domaine concerné : Collecte ; Mesures et impact de la pandémie de Covid-19

Résumé :

L'enquête EpiCov est une opération statistique comprenant une enquête en panel de quatre interrogations (mai 2020, novembre 2020, juillet 2021 et la dernière prévue en mars 2022) portant sur un échantillon initial de 371 000 personnes résidant en France, ainsi qu'un volet biologique permettant de caractériser le statut sérologique de certains répondants de l'enquête à l'issue des interrogations. Le présent article vise à décrire le protocole de cette opération.

La mise en place de cette cohorte s'est faite dès le début de la crise sanitaire. L'échantillon de la cohorte a été constitué à partir des sources fiscales (Fidéli), et comportait 350 000 individus de métropole, ainsi que 7 000 de Martinique, de Guadeloupe et de La Réunion. Cet échantillon se compose de 20 lots, de structure identique, permettant des nuances dans le protocole de collecte. En vague 2 comme en vague 3, n'étaient interrogées que les personnes ayant complété le questionnaire de la vague précédente.

- La première vague d'enquête s'est déroulée (pour la partie questionnaire) du 2 mai au 1^{er} juin 2020. Au cours de ce terrain, l'ensemble des participants a reçu un lien de connexion pour répondre en ligne au questionnaire, et 20 % des individus ont par ailleurs été appelés, en cas de non-réponse. Le début des appels a été légèrement échelonné sur quelques jours pour les trois premiers lots, et le quatrième lot a commencé à être appelé au bout de trois semaines seulement. Des tests sérologiques ont été proposés à une partie de l'échantillon métropolitain, entre 1 et 16 lots selon le département. Un questionnaire légèrement plus long était administré aux lots 1 et 20.
- La deuxième vague d'enquête a eu lieu entre le 26 octobre et le 14 décembre 2020. Les appels téléphoniques et les tests sérologiques ont cette fois été proposés à l'ensemble des participants. Un sous échantillon de quatre lots s'est vu proposer des tests sérologiques pour l'ensemble des membres du ménage (à partir de 6 ans), afin de pouvoir étudier les contaminations intrafamiliales.

- La troisième vague d'enquête a eu lieu entre le 24 juin et le 6 août 2021. Cette vague n'a comporté aucun test sérologique. La distinction entre questionnaires court et long n'avait pas d'objet pour cette vague.
- Une quatrième et dernière vague est enfin prévue pour mi-mars 2022. Cette interrogation sera l'occasion de mettre en place une réalimentation de la cohorte, en n'interrogeant pas que les seuls répondants de la vague 3, mais l'ensemble des 371 000 personnes initialement interrogées, à l'exception de celles ayant explicitement refusé de participer à l'enquête. Des tests sérologiques seront à nouveau proposés en vague 4.

Le taux de réponse obtenu au cours de la vague 1 a été de 36 % ; en ré-interrogation, la vague 2 et la vague 3 ont obtenu des taux de réponse aux alentours de 80 %. Une analyse spécifique des taux de réponse en vague 1 est proposée par l'article [2].

Les taux d'acceptation et de renvoi des prélèvements sérologiques ont été de 87 % des répondants acceptant en vague 1 de recevoir le matériel, et de 81 % parmi eux ayant effectivement renvoyé un kit (soit 70 % de ceux auxquels un prélèvement a été proposé) ; en vague 2, ces taux sont légèrement plus faibles : 81,5 % ont accepté de participer au volet biologique, et 77,1 % d'entre eux ont bien renvoyé un kit, soit 62,8 %. Les cohabitants des personnes interrogées auxquels un kit a été envoyé ont été 63,5 % à renvoyer un kit. Une analyse spécifique de ces taux d'acceptation et de renvoi en vague 1 est proposée par l'article [3].

Les traitements post-collecte menés ont principalement consisté à mener des traitements de correction de la non-réponse totale : des critères propres à chaque vague ont été déterminés pour juger de la complétude des réponses, puis une repondération par constitution de groupe de réponse homogène ainsi qu'un calage sur marges ont été menés, pour chacun des différents échantillons distingués au sein du protocole. L'observation des prévalences obtenues pour la description des symptômes pouvant évoquer une infection au SARS-CoV-2 a motivé une analyse plus en profondeur des effets du mode de collecte (téléphone ou internet), et une pondération spécifique a été produite en conséquence pour neutraliser cet effet : la constitution de cette pondération est décrite dans l'article [4].

L'enquête est enfin enrichie de plusieurs manières : une reconstitution des professions et des secteurs d'activité à partir des libellés collectés dans l'enquête, une actualisation des données fiscales de la base de sondage (Fidéli 2018) grâce aux données fiscales de 2020, et un appariement prévu aux données du système national des données de santé (SNDS), contenant les données sur les consommations de soins des individus rassemblées par l'Assurance maladie.

Un retour critique sur le protocole d'enquête, sur les forces et sur les limites de l'enquête permettra de conclure cet article.

Bibliographie

[1] Warszawski, J. et al., A national mixed-mode seroprevalence random population-based cohort on SARS-Cov-2 epidemic in France : the socio-epidemiological EpiCov study, 2021, preprint MedRxiv.

[2] Bondon M., Charrance G., Cochet P., Leduc A., Merly-Alpa T., « Apprendre des paradonnées pour améliorer les protocoles de collecte : l'exemple d'EpiCov » (à paraître, présentation aux JMS 2022)

[3] Duchesne L., Rahib D., Lydié N., « Renvoi des kits d'auto-prélèvement et caractéristiques des participants au volet séro-épidémiologique de l'enquête EpiCov » (à paraître, présentation aux JMS 2022)

[4] Castell L., « Biais de mesure et biais de sélection : comment distinguer les biais d'une enquête multimode ? L'exemple de l'enquête EpiCov » (à paraître, présentation aux JMS 2022)