
UN PROTOCOLE INNOVANT, MULTIMODE EN DEUX PHASES, POUR MESURER LA VICTIMATION : RÉSULTATS ET PRINCIPAUX ENSEIGNEMENTS DE LA COLLECTE DE L'ENQUÊTE GENRE ET SECURITE (GENESE)

Hélène GUEDJ, Axelle QUIVIGER, Sandra ZILLONIZ (*), Tiaray RAZAFINDRANOVA (**)

(*) SSMSI, section Enquêtes statistiques

(**) SSMSI, section Méthodes statistiques et analyses conjoncturelles

helene.guedj@interieur.gouv.fr, axelle.quiviger@interieur.gouv.fr,
sandra.zilloniz@interieur.gouv.fr, tiaray.razafindranovala@interieur.gouv.fr

Mots-clés (6 maximum) : Protocole multimode, collecte, relance, internet, coordonnées de contact, parodonnées

Domaine concerné : 4. Collecte Protocole, conception des enquêtes,... / Multimode

Résumé (813 mots)

Au sein du ministère de l'Intérieur, le Service statistique ministériel de la sécurité intérieure (SSMSI) est en charge du domaine couvrant la victimation, la délinquance et l'insécurité. Très sollicité sur la thématique des violences faites aux femmes, le SSMSI a obtenu en novembre 2019 un financement européen pour conduire en 2021, au niveau national, l'enquête Genre et Sécurité (Genese), portant sur les violences liées au genre et basée sur le questionnaire *Gender-Based Violence* développé par Eurostat. Cette première enquête en population générale conduite par le SSMSI a été conçue pour répondre à deux ambitions majeures : éclairer le débat public à l'échelle européenne avec des données inédites sur les parcours de violences sexistes et sexuelles ; réaliser une expérimentation multimode grandeur réelle pour mesurer les victimations et les perceptions en matière de sécurité en vue de la refonte de l'enquête « Cadre de vie et sécurité » à l'horizon 2022.

La mesure de la victimation pose deux difficultés majeures liées à : i/ la faible fréquence des atteintes mesurées, en particulier les violences les plus graves, qui impose des tailles d'échantillons conséquentes ; ii/ le caractère potentiellement sensible des sujets abordés qui rend particulièrement délicat l'arbitrage entre les avantages et inconvénients des différents modes de collecte. En s'inspirant d'enquêtes existantes comme l'enquête Vie quotidienne et santé (VQS), le SSMSI a conçu, en lien avec l'Insee, un protocole multimode strictement nominatif en deux phases. La première phase consiste en une enquête de victimation « généraliste » courte (~20 mn) basée sur le recueil des caractéristiques sociodémographiques, des questions de recensement des atteintes et des opinions en matière de sécurité auprès d'un échantillon de 169 060 personnes âgées de 18 à 74 ans vivant en France métropolitaine tiré dans les fichiers fiscaux (Fidéli) par l'Insee. La collecte, assurée par un institut de sondage, est concurrentielle différée : elle commence par internet puis prévoit une relance téléphonique auprès d'une sélection aléatoire de 25 000 non-répondants et l'envoi d'un auto-questionnaire papier à différents stades de la collecte en

fonction des coordonnées disponibles pour contacter l'enquêté. La deuxième phase consiste en une enquête de suivi thématique plus longue (~30 mn) basée sur le questionnaire conçu par Eurostat qui permet un recueil d'informations très détaillées sur les violences sexistes et sexuelles subies tout au long de la vie. Elle est proposée par internet et téléphone à un échantillon de 15 000 répondants de phase 1 dans lequel les victimes de violences physiques et sexuelles sont surreprésentées. Enfin au terme de la phase 1, un volet méthodologique en face à face est conduit auprès de tous les non-répondants totaux de Gironde et d'une sélection aléatoire de non-répondants d'Île-de-France. L'enquête Genese a obtenu en janvier 2021 l'avis de conformité du Comité du label avec caractère obligatoire.

Le terrain de la phase 1 a démarré le 1^{er} mars 2021 et s'est achevé le 16 mai après 11 semaines de collecte. Une lettre-avis et trois relances postales ont été envoyées à intervalle régulier ainsi qu'un mail-avis et six relances mail. Au total, près de 109 000 individus (64 %) ont répondu à la phase 1 dont 51 % par internet, 3 % par téléphone et 10 % par papier. En boostant la réponse par internet et en personnalisant l'envoi du questionnaire papier en fonction des coordonnées de contact disponibles, le schéma de relance conçu par le SSMSI a permis de dépasser l'objectif de participation (50 %). La phase 2 s'est déroulée du 1^{er} mars au 10 juillet : sur 15 000 individus sollicités, plus de 10 000 (68 %) ont répondu : par internet et par téléphone. Enfin plus de 1 000 entretiens en face à face ont été réalisés auprès des 3 000 non-répondants de phase 1 sélectionnés (34 %).

Une description détaillée du protocole ainsi que les bilans de collecte des deux phases de Genese et du volet méthodologique seront présentés. L'étude examinera de manière approfondie : la qualité des données de contact fournies par l'Insee (adresses postales, mails, numéros de téléphone), l'apport de l'enrichissement de coordonnées réalisé par le prestataire de collecte ainsi que l'efficacité des différents modes de contact en fonction des coordonnées disponibles pour l'individu cible. Les travaux analyseront les gains de représentativité offerts par chacun des modes sur l'échantillon de répondants de phase 1. L'articulation entre les deux phases sera évaluée, en particulier l'attrition des répondants de phase 1 et l'intérêt du protocole lissé proposé aux individus automatiquement sélectionnés pour la phase 2. Ces résultats seront enrichis de l'analyse des nombreuses parodonnées recueillies dans le dispositif concernant la connexion au questionnaire (type de support, jours et heures, etc.) et des données de consultation du site internet et de contacts à la hotline.

Les conclusions de ces travaux doivent permettre d'adopter le protocole le plus pertinent en termes de coût-efficacité pour la nouvelle enquête annuelle de victimation conduite par le SSMSI dont la première édition sera sur le terrain en mars 2022.

Bibliographie

- [1] De Leeuw, E. D. (2018). « Mixed-mode: Past, present, and future. » In *Survey Research Methods* (Vol. 12, No. 2, pp. 75-89).
- [2] Schouten, B., van den Brakel, J., Buelens, B., Giesen, D., Luiten, A., et Meertens, V. (2021). « Mixed-Mode Official Surveys: Design and Analysis », CRC Press.

**JMS 2022 : Nomenclature des thématiques
pour la classification des communications**

Thématique	Sous-thématique
1. Théorie des sondages amont	Échantillonnage Échantillonnages particuliers : spatial, équilibré, sur population continue .. Bases de sondage Unités statistiques Couplage avec données administratives
2. Théorie des sondages aval	Redressement, pondération et repondération, calage sur marges Calcul de précision, estimation de variance Appariements statistiques
3. Contrôle et redressement des données, data editing	Non-réponse Imputation Identification, traitement des valeurs atypiques ou extrêmes (cas du milliardaire dans une enquête..) ou des valeurs influentes Codification automatique
4. Collecte	Protocole, conception des enquêtes, couverture de populations particulières Conception de questionnaire Multimode Effets de mode Effets d'oubli / de mémoire Paradoxaux
5. Intégration de données	Appariements et fusion de sources (<i>record linkage</i> , appariement « exact » de fichiers) Signes de vie Registres d'individus et de logements Nettoyage des données administratives.
6. Statistique spatiale	Économétrie spatiale Statistiques locales Estimation sur petits domaines, carroyage Zonages
7. Econométrie	Théorique Appliquée (étude de cas) Évaluation des politiques publiques
8. Modélisation	Mathématique ou stochastique Microsimulations Algorithmes
9. Séries temporelles	Analyse des séries Désaisonnalisation

Nowcasting
Projections, prévisions

**10. Analyse des données
et *data science***

Statistique descriptive, analyse
factorielle
“*Big data*”, nouvelles données,
données massives
Machine learning / Apprentissage
statistique, classification
Intelligence artificielle, deep
learning, réseaux de neurones
NLP, analyse textuelle
Data viz

11. Institutionnel, *open science*

Histoire
Confidentialité
Anonymisation
Diffusion
Communication
Open data, documentation,
métadonnées, mise à disposition,
partage de codes, reproductibilité

12. Comptabilité nationale

Nationale
Estimations régionales
Indicateurs macro-économiques

13. Concepts et mesures

Indicateurs, échelles, indices
Incertitudes
Erreurs de mesures

**14. Mesures et impact de la
pandémie de Covid**

15. Données médicales

16. Démographie

Recensement
Indicateurs ou études
démographiques

17. Statistique d'entreprises

18. Enseignement, éducation

Évaluation des élèves
Autre