

**L'EFFET DU RAPPEL DES REFUSANTS ET DU RANGE D'APPEL
SUR LA REPRESENTATIVITE D'UNE ENQUÊTE TELEPHONIQUE.
L'EXPERIENCE DE L'ENQUÊTE FECOND**

Stéphane Legleye, Géraldine Charrance, Nicolas Razafindratsima et l'équipe INSERM FECOND

I. Introduction

En théorie, le téléphone est un moyen très efficace de réaliser des enquêtes en population générale : En France, en 2011, plus de 89% de la population des ménages ordinaires sont équipés d'une ligne fixe, et 85% d'un téléphone portable, laissant seulement 1% de la population injoignable par téléphone (Bigot et Croutte, 2011). L'offre téléphonique s'est énormément diversifiée ces dernières décennies. La part des personnes disposant d'un double équipement (fixe et mobile) a rapidement augmenté : près de 74% des personnes sont équipés d'une ligne fixe et d'un téléphone portable (contre seulement 50% en 2003). De plus, beaucoup de ménages utilisent dorénavant un boîtier multiservice (box) pour téléphoner (58%). Ces évolutions ont contribué à une meilleure couverture de la population générale, mais ont complexifié la constitution d'échantillons représentatifs. En effet, il n'existe pas de répertoire exhaustif unique des souscripteurs à une ligne fixe ou mobile dans lequel un tirage aléatoire pourrait être réalisé. Par ailleurs, certains individus peuvent être tirés dans différentes sources (Brick et al., 2011 ; Gautier et al., 2006).

Malgré tout, la principale difficulté à laquelle sont confrontés les producteurs d'enquête reste la baisse générale des taux de participation, observée pour tous les modes de collecte et particulièrement pour les enquêtes téléphoniques (Lan, 2009 ; Schouten et al., 2009). Le biais de non-réponse résulte d'une combinaison de deux facteurs : la proportion de non-répondants et les différences entre les répondants et les non-répondants quant aux variables d'intérêt. Des taux de non-réponse élevés sont problématiques dans la mesure où la différence entre répondants et non-répondants est assez grande pour affecter un estimateur donné (Groves, 2006).

Malheureusement, une estimation correcte du biais de non-réponse requiert la collecte d'informations sur les variables d'intérêt de l'enquête auprès des non-répondants ainsi que leurs caractéristiques sociodémographiques (que l'on suppose liées aux comportements étudiés). La première condition est impossible à satisfaire, par définition. Pour la seconde, elle est difficile à atteindre dans le cadre d'une enquête téléphonique par génération aléatoire de numéros, dans la mesure où les numéros ne fournissent aucune information sociodémographique, ni même géographique, sur leurs possesseurs (seule la moitié

des souscripteurs de ligne fixe sont présents dans l'annuaire et il n'existe pas d'annuaire pour les possesseurs de téléphone portable). La correction du biais de non-réponse consiste à établir un système de pondérations qui permet de caler l'échantillon sur la structure de la population de référence. Cette pratique repose sur l'hypothèse selon laquelle à caractéristiques sociodémographiques égales, les répondants peuvent se substituer aux non-répondants. Dans l'idéal, nous devrions disposer d'un échantillon de non-répondants pour comparer, dans les deux échantillons, les associations entre les caractéristiques sociodémographiques et les variables d'intérêt de l'enquête.

Plusieurs études ont montré que la relation entre le taux de non-réponse et l'ampleur du biais de non-réponse est faible (Curtin, Presser et Singer, 2000 ; Keeter et al., 2000). Mais les études sur ce thème restent assez peu nombreuses et les résultats observés sont potentiellement sensibles au contexte national et historique ainsi qu'au sujet de l'enquête, ce qui doit conduire les concepteurs d'enquête à faire preuve de prudence lors de la présentation des résultats à l'échelle internationale.

L'intégration au processus de collecte d'interventions méthodologiques, dont l'objectif est de s'assurer que les ménages/les personnes qui sont les plus difficiles à interviewer soient inclus dans l'échantillon, devient de plus en plus courante. Ce principe repose sur l'hypothèse qu'il existe un continuum entre les répondants qui sont les plus « faciles » à enquêter, ceux qui nécessitent davantage d'efforts et les non-répondants. En d'autres mots, les individus difficiles à enquêter sont proches des non-répondants.

Etant donné que les non-répondants peuvent être divisés en deux sous-populations (les non-contactés et les refusants), différentes opérations doivent être considérées pour réduire le taux de non-réponse. Le taux de non-contact peut être réduit en augmentant le nombre de tentatives d'appel avant d'abandonner un numéro. Des études sur les bénéfices de cette pratique ont déjà été menées (Firdion, 1993 ; Beck, Legleye et Peretti-Watel, 2005 ; Razafindratsima, 2008), tout comme d'autres sur l'utilité d'intégrer les répondants tardifs (après 31 jours) aux autres répondants dans les enquêtes de surveillance mensuelles (Qayad, Chowdhury, Hu et Balluz, 2010). Elles concluent que les profils sociodémographiques des répondants difficiles à contacter diffèrent substantiellement de celui des autres répondants. Des différences de comportements (usage de drogues, comportements sexuels, attitudes politiques, niveau et comportements de santé) sont observées, même après contrôle des caractéristiques sociodémographiques. Dans ces enquêtes, une limite était toutefois imposée sur le nombre total d'appels.

Pour réduire le taux de refus, deux méthodes prédominent : le rappel des refusants et les incitations financières (Groves, 2006 ; Groves, Singer et Corning, 2000). Le recours aux incitations financières est assez marginal dans le domaine de la recherche publique en France, en raison de la forte tradition qui valorise la participation volontaire à la recherche scientifique et à la statistique publique, la considérant comme une contribution au bien public, par opposition à la pratique consistant à recourir à des incitations financières dans les enquêtes commerciales conduites par les entreprises privées. Tenter de convaincre les refusants de finalement participer à l'enquête commence à être pratiqué en France (Beck et Guilbert, 2007).

Cependant, aucune étude n'a été réalisée en France pour évaluer l'intérêt de cette pratique. Basée sur un corpus de données issues de six enquêtes menées en Grande-Bretagne durant les années 1990, une étude a montré que les efforts visant à convertir les refusants initiaux sont moins efficaces que ceux visant à atteindre les répondants difficiles à contacter dans les enquêtes en face-à-face (Lynn et Clarke, 2002). Aucune étude similaire n'a été réalisée sur les enquêtes téléphoniques.

L'objectif de cet article est, en se basant sur une enquête téléphonique menée auprès de la population française en 2010/2011, de caractériser les deux catégories de répondants difficiles à enquêter (dû soit à leur refus initial, soit à leur indisponibilité) en termes de profil sociodémographique et des comportements étudiés. Une comparaison des déterminants sociodémographiques de quelques comportements sexuels sera faite pour tester l'hypothèse de parfaite substituabilité entre répondants faciles à enquêter et répondants difficiles à enquêter. Dans cette enquête, aucune limite n'était fixée pour le nombre total d'appels et les refusants étaient rappelés jusqu'à deux fois : les questionnaires obtenus par ces moyens n'auraient pas été collectés dans la grande majorité des enquêtes téléphoniques et peuvent alors être considérés comme des questionnaires de personnes absentes de l'échantillon dans d'autres enquêtes. Enfin, nous comparerons les coûts au sens large de chaque combinaison possible de ces deux interventions (augmentation des tentatives d'appel et rappel des refusants).

II. Méthodes

Présentation de l'enquête

L'enquête FECOND (Fécondité, Contraception et dysfonctions sexuelles) a pour objectif d'analyser les enjeux contemporains en matière de santé sexuelle et reproductive et permettre le développement d'une approche compréhensive explorant simultanément divers sujets, du point de vue des hommes et des femmes (*Population et Sociétés*. N°492, septembre 2012.)

Cette enquête a été menée par téléphone entre Juin 2010 et Janvier 2011, avec un mois d'interruption en Août 2010. L'échantillon a été obtenu par sondage stratifié à deux degrés (ménage et individu) selon le type de téléphone (ligne fixe ou mobile).

L'unité primaire d'échantillonnage est le ménage. Pour cela, une génération aléatoire de numéros de ligne fixe et mobiles a été réalisée : les lignes mobiles représentent 16% des numéros générés, correspondant à la proportion de possesseurs de mobile exclusifs dans la population cible (au début du questionnaire, les personnes disposant d'une ligne fixe, en plus de leur ligne mobile, étaient éliminées). Il n'y a pas eu de stratification géographique car une telle information n'est pas disponible dans le cas d'une génération aléatoire de numéros. La base générée a d'abord été croisée avec celle des racines attribuées par l'ARCEP

(Autorité de Régulation des Communications Électroniques et des Postes) pour éliminer les numéros non valides, puis avec l'annuaire téléphonique pour éliminer les entreprises et, pour finir, les numéros étaient composés par un automate d'appels afin de vérifier que les lignes soient en service. Ce protocole permet de couvrir 99% de la population cible. Pour les ménages dont on disposait d'une adresse postale (grâce à l'annuaire téléphonique), une lettre-avis leur était envoyée quelques jours avant le début du terrain.

La deuxième étape de l'échantillonnage consiste à sélectionner une personne de 15 à 49 ans dans le ménage par la méthode Kish (description des membres du foyer en termes de sexe et d'âge puis sélection aléatoire d'un éligible par l'enquêteur lors de la prise de contact). Les femmes avaient une probabilité plus importante d'être sélectionnées afin de les surreprésenter dans l'échantillon final.

Durant la collecte, deux interventions ont été mises en œuvre pour améliorer le taux de réponse : rappeler les refusants et autoriser un très grand nombre de tentatives d'appel. Ces interventions ont déjà été utilisées dans de précédentes enquêtes mais dans de moindres mesures. Nous avons choisi d'appliquer ces deux opérations quasiment sans restriction dans une approche expérimentale : en conséquence, les questionnaires recueillis par ces moyens peuvent être considérés comme des questionnaires de personnes qui auraient été non-répondants autrement. Cette approche nous permet de déterminer, en pratique, quelle intervention améliore le plus la qualité des données.

Les refusants ont été rappelés au plus tôt sept jours après leur refus. La plupart de ces refus sont des refus dits de niveau ménage, c'est-à-dire qu'ils ont lieu avant la présentation de l'enquête et la sélection du répondant (au moins pour les lignes fixes) : une nouvelle tentative est une chance supplémentaire de trouver un interlocuteur moins occupé avec lequel la sélection du répondant pourra être réalisée. Les enquêteurs les plus expérimentés ont été mobilisés pour cette opération. La prise de contact et l'argumentaire développé par l'enquêteur étaient adaptés au comportement de l'interlocuteur et aux raisons de son refus initial (les refus « violents » n'étaient pas rappelés). Deux vagues de rappel ont été réalisées pour les lignes fixes et une seule pour les lignes mobiles. Contrairement aux lignes fixes, les mobiles sont généralement utilisés par une seule personne, les chances de tomber sur le même interlocuteur et donc d'essuyer un nouveau refus étaient nettement plus grandes. Les résultats obtenus à l'issue de la première vague de rappel ont également contribué à prendre la décision de ne pas réaliser une seconde vague de rappel auprès des mobiles.

Parmi les 92 516 numéros valides appelés en première vague, le taux de réponse initial, calculé selon les recommandations de l'American Association for Public Opinion Research (AAPOR 2011, Standard definitions, page 44), est de 35,3%, le taux de non-contact de 25,3% et le taux de refus de 34,7%. Les 4,7% restants ont été considérés comme incapables de répondre à l'enquête soit en raison d'une absence

prolongée, soit de problèmes de langue. Au total, 16 726 refusants ont été rappelés (lors d'une seconde vague) : 15,1% ont accepté de participer à l'enquête, 68,2% ont refusé ou abandonné le questionnaire, 12,8% n'ont pu être joints et 3,9% se sont révélés impossible à enquêter (absence, langue). Une dernière vague de rappel a été menée auprès de 6 552 refusants contactés par une ligne fixe : 6,9% ont finalement accepté de répondre au questionnaire et 73,0% ont maintenu leur refus.

Le protocole de collecte n'imposait pas un nombre maximal de tentatives d'appel par numéro. La règle était la suivante : un numéro était appelé jusqu'à 20 fois si aucun contact n'était établi. Au-delà de cette limite, le numéro était classé comme injoignable. Dès qu'un contact était établi, un quota de 20 appels supplémentaires était fixé pour réaliser le questionnaire. Notons également que le compteur d'appels était réinitialisé à chaque vague. Le nombre d'appels réalisés durant la vague précédente n'était pas pris en compte dans la gestion des appels de la vague en cours. Le nombre maximum d'appel pour un numéro est de 373 (toutes issues) et 158 pour une interview pour les lignes fixes et 73 et 58 respectivement pour les lignes mobiles. 14,9% des lignes fixes et 25,3% des lignes mobiles ont été abandonné après 30 appels.

Le taux de participation final est de 44,8%. Près de 30% des numéros n'ont jamais pu être joints, 20,2% des numéros contactés ont maintenu leur refus et 5,1% étaient impossibles à enquêter (problème de langue ou absence prolongée). Le tableau 1 restitue ces informations.

Répondants difficiles à contacter et refusants initiaux

Dans l'échantillon final, 9,9% des questionnaires (n=856, 16,6% des répondants de l'échantillon mobile, 9,3% de l'échantillon fixe) ont été réalisés après le vingtième appel lors d'une vague donnée (ces individus représentent les « difficiles-à-contacter »). Par ailleurs, 19,0% des questionnaires (n=1 645, 13,3% des répondants de l'échantillon mobile, 20,0% de l'échantillon fixe) ont été réalisés après un premier refus (ces répondants seront appelés « refusants initiaux »). Ces deux sous-échantillons sont considérés comme les « difficiles à enquêter », à opposer au reste de l'échantillon considéré comme « facile à enquêter » (figure 1).

Analyses statistiques

Dans un premier temps, nous avons comparé les distributions des sous-échantillons de répondants (faciles versus difficiles à interroger) pour chacune des deux interventions (rappel des refusants et nombre de tentatives d'appel élevé) sur quelques caractéristiques sociodémographiques (sexe, âge, niveau de diplôme, etc.), en utilisant les pourcentages et les tests du χ^2 de Pearson.

Tableau 1. Nombre de questionnaires réalisés à chaque étape de l'enquête et nombre total pour chaque vague

	Vague 1	Vague 2 (1ère vague de rappel)	Vague 3 (2 ^{nde} vague de rappel)	Total
Nombre de questionnaires réalisés				
Rang d'interview ≤ 20 (faciles à contacter)	6 232	1 256	301	7 789
Lignes fixes	5 263	1 097	301	6 430
Téléphones portables	969	159	-	1 128
Rang d'interview > 20 (difficiles à contacter)	772	83	1	856
Lignes fixes	610	68	1	679
Téléphones portables	162	15	-	177
Total	7 004	1 339	302	8 645
Lignes fixes	5 873	1 165	302	7 340
Téléphones portables	1 131	174	-	1 305
Issues des appels				
Nombre de numéros appelés	92 516	16 726	6 552	92 516
Eligibles (1)	11 664	4 215	1 604	12 751
- Questionnaires réalisés (1.a)	7 004	1 339	302	8 645
- Refus (1.b)	2 691	2 028	1 004	1 703
- Abandons (1.c)	654	209	35	341
- Répondants injoignables (1.d)	822	400	181	1 403
- Interviews impossibles* (1.e)	310	154	70	534
- Interviews échouées (problèmes techniques) (1.f)	183	85	12	125
Eligibilité inconnue (2)	30 696	8 799	4 150	25 101
- Toujours occupé/sans réponse (2.a)	16 088	1 397	869	16 523
- Problème de langue au niveau ménage (2.b)	1 061	193	54	1 302
- Refus immédiat (2.c)	13 547	7 209	3 227	7 276
Non éligibles (3)	33 012	3 712	798	37 522
- Entreprises (3.a)	17 348	419	24	17 791
- Pas en France Métropolitaine (3.b)	495	62	1	558
- Résidences secondaires / mobiles non-exclusifs	4 876	959	6	5 841
- Répondant non éligibles (3.d)	10 008	2 235	760	13 003
- Hors quotas (3.e)	285	37	7	329
Faux numéros / Fax	17 144	-	-	17 144
Indicateurs				
- Taux de réponse	35.3%	15.1%	6.9%	44.8%
- Taux de refus/abandons	34.7%	68.2%	73.0%	20.2%
- Taux de non-contact	25.3%	12.8%	17.4%	29.9%
- Taux d'interviews impossibles	4.8%	3.9%	2.7%	5.1%

Deux taux d'éligibilité ont été appliqués à la catégorie "Eligibilité inconnue" selon si le numéro a été identifié comme étant un ménage ou non.

- Taux de ménages parmi tous les numéros ($a = ((1)+(3.b)+(3.c)+(3.d)+(3.e))/((1)+(3))$)

- Taux de ménages éligibles parmi tous les ménages ($b = (1) / ((3.b)+(3.c)+(3.d)+(3.e))$)

Pour les catégories (2.a) et (2.c), les deux taux ont été appliqués successivement aux effectifs. Pour la catégorie (2.b), seul le second taux (b) a été appliqué. Le dénominateur pour chacun des taux est le suivant : $(1)+a*b*((2.a)+(2.c))+b*(2.b)$. Pour le taux de réponse, le numérateur correspond aux effectifs de questionnaires réalisés (1.a). Pour le taux de refus/abandons, sont présents aux numérateurs les catégories suivantes : $(1.b)+(1.c)+a*b*(2.c)$. Pour le taux de non-contact, le numérateur est : $(1.d)+(a*b*2.a)$ et pour le taux d'interviews impossible : $(1.e)+(1.f)+(b*2.b)$.

* Un questionnaire est considéré comme impossible à réaliser s'il y a un problème de langue (avec le répondant ou l'interlocuteur), si la personne est physiquement ou mentalement incapable d'être enquêté ou si elle est absente pendant toute la durée du terrain.

Figure 1: Description des sous-échantillons obtenus

	Rangs ≤ 20 (1)+(3) Faciles à contacter	Rangs > 20 (2)+(4) Difficiles à contacter
Vague 1 Echantillon initial (1)+(2)	(1) Faciles à enquêter : Répondants initiaux Faciles à contacter <i>n=6232</i>	(2) Difficiles à enquêter : Répondants initiaux Difficiles à contacter <i>n=772</i>
Vague 2 et 3 Echantillon des refusants initiaux (3)+(4)	(3) Difficiles à enquêter : Refusants initiaux Faciles à contacter <i>n=1557</i>	(4) Difficiles à enquêter : Refusants initiaux Difficiles à contacter <i>n=84</i>

Nous avons ensuite étudié la contribution de chaque catégorie de répondants difficiles à enquêter à la réduction de la distorsion de la distribution de l'échantillon par rapport à la population cible (données du recensement sur la population générale vivant en France en 2008). Le biais est mesuré par la distance du Khi^2 , calculée comme une statistique classique du Khi^2 , à l'exception du fait que chaque échantillon considéré a une taille fixée arbitrairement à 2000 afin de neutraliser l'effet de la taille de l'échantillon sur la statistique de test. Cette méthode permet de comparer des distributions des variables dans des échantillons de taille différente. Le but était d'observer l'évolution de la distance entre l'échantillon initial des répondants faciles à enquêter et la population cible d'une part, et entre l'échantillon finale et la population cible, d'autre part, pour chaque opération.

Nous avons ensuite mesuré et comparé les prévalences de plusieurs comportements sexuels dans les sous-échantillons en utilisant le test du Khi^2 de Pearson. Ne disposant pas de source externe pour ces variables, nous n'avons pu mesurer l'éventuelle réduction du biais associée à chaque intervention méthodologique.

Enfin, nous avons tenté d'estimer les effets possibles des deux interventions sur les déterminants sociodémographiques des variables d'intérêt étudiées, en utilisant des régressions logistiques pour tester les différences d'associations entre caractéristiques sociodémographiques et comportements sexuels dans chaque sous-échantillon. Des tests d'égalité des coefficients ont été réalisés pour déterminer si les déterminants étaient identiques ou significativement différents dans les différents sous-échantillons. Pour des raisons de lisibilité, seuls les modèles sur le nombre de partenaires sexuels au cours de la vie et au cours des douze derniers mois sont présentés dans les tableaux.

En dernier lieu, nous avons comparé l'impact des deux interventions méthodologiques et leurs combinaisons sur la taille de l'échantillon, le biais sociodémographique et les coûts financiers de collecte. Parmi les mesures du biais de non-réponse, nous avons utilisé le ratio entre les poids après calage et les poids de sondage.

Les données à l'échelle micro, disponibles grâce à l'enquête annuelle du recensement (INSEE), ont été utilisées pour obtenir les répartitions de quelques variables sociales et démographiques dans la population

cible. La post-stratification, ou calage, a été réalisée grâce à la macro CALMAR créée par l'INSEE (Sautory, 1993). Toutes les analyses ont été faites avec le logiciel SAS V9.3.2.

Hypothèses méthodologiques

Nous avons étudié l'effet du rang d'interview toutes vagues confondues, pour deux raisons. D'une part, le profil des répondants difficiles à contacter interrogés en vague 1 et celui des personnes interrogées en vagues 2 et 3 sont très proches malgré quelques différences observées sur le niveau de diplôme, le mode de vie ou encore la région de résidence ; de plus, ces caractéristiques sont contrôlées dans les analyses multivariées. Deuxièmement, une seule vague de rappel a été réalisée auprès des lignes mobiles. Pour les mêmes raisons, nous avons agrégé les deux vagues de rappel des refusants et considéré les lignes fixes et mobiles conjointement dans les analyses.

III. Résultats

Refusants initiaux

Le tableau 2 (partie gauche) présente une comparaison des profils sociodémographiques des répondants selon la vague de l'interview (exploitation principale versus vagues de rappel). Parmi les refusants initiaux, on compte une plus grande proportion de personnes appartenant aux classes d'âge extrêmes (15/24 ans et 40/49 ans), de personnes peu diplômées, de chômeurs ou inactifs, de personnes vivant soit avec en couple, soit avec leurs parents, dans des ménages de trois personnes ou plus.

Afin de déterminer si l'inclusion des refusants initiaux dans l'échantillon conduit à une réduction du biais sociodémographique, nous avons comparé l'échantillon initial (vague 1) et l'échantillon final (vagues 1 à 3) aux données du recensement de 2008 (tableau 2, partie droite). L'inclusion des refusants initiaux tend à augmenter la distorsion des distributions de l'échantillon (mesuré par une augmentation des valeurs du Khi^2) comparé à l'échantillon initial. A contrario, les résultats ne montrent pas de différences significatives dans les prévalences des variables d'intérêt entre les répondants de la première vague et ceux des vagues de rappel (tableau 3, partie gauche).

Dans l'optique d'explorer les effets de la stratégie de rappel des refusants sur les résultats de l'enquête, nous avons étudié les différences d'effets des caractéristiques sociodémographiques des individus sur les comportements sexuels entre les sous-échantillons (exploitation principale versus vagues de rappel) (tableau 4). Pour le nombre de partenaires sexuels (au cours de la vie et au cours des douze derniers mois), l'hypothèse d'égalité des effets dans les deux sous-échantillons est rejetée chez les femmes (au seuil de 5%) : en d'autres termes, les déterminants sont significativement différents selon la vague d'interview.

Tableau 2. Profil sociodémographique des répondants par vague d'interview (% , p-value du χ^2 et valeur du χ^2)

	Echantillon initial (vague 1)	Echantillon des vagues de rappel (vagues 2 et 3)	Echantillon initial (vague 1)	Echantillon final (vagues 1, 2 et 3)	Population cible (recensement)
Sous-échantillons	(1)+(2)	(3)+(4)	(1)+(2)	(1)+(2)+(3)+(4)	
Sexe		ns			
			(12.75)	(13.64)	
Hommes	45.5	45.2	45.5	45.3	49.5
Femmes	54.5	54.9	54.5	54.7	50.5
Age		***			
			(11.47)	(22.75)	
15-17	9.1	13.4	9.1	9.9	8.1
18-24	20	20.6	20	20.1	18.4
25-29	12.2	9.7	12.2	11.8	13.4
30-34	12.4	10.2	12.4	11.9	14.0
35-39	15	13.8	15	14.8	15.3
40-44	15.4	15.4	15.4	15.4	15.5
45-49	15.9	17	15.9	16.1	15.2
Niveau de diplôme		***			
			(149.55)	(117.55)	
Aucun diplôme, CEP, BEPC	16.7	22.4	16.7	17.8	25.9
BEP, CAP	21.9	23.8	21.9	22.2	24.0
Baccalauréat	23.2	23.1	23.2	23.1	20.4
Bac+1 à Bac+3	22	18.1	22	21.2	14.7
Bac+4 et plus	16.3	12.7	16.3	15.6	15.0
Situation professionnelle		***			
			(0.77)	(2.91)	
Actifs occupés	68.3	63.9	68.3	67.4	69.2
Inactifs/au chômage	31.7	36.1	31.7	32.6	30.8
Pays de naissance		ns			
			(36.69)	(38.99)	
France	92.6	93.3	92.6	92.8	88.3
Autres	7.4	6.7	7.4	7.2	11.7
Mode de vie		***			
			(13.01)	(15.96)	
En couple	60.1	59.5	57.7	60	56.3
Avec les parents	20.8	25.2	21.4	21.7	21.9
Autres situations	19.1	15.3	20.9	18.3	21.8
Région de résidence		ns			
			(17.75)	(15.71)	
Ile-de-France, Bassin parisien	32.8	34.2	33	33.1	37.3
Nord, Est, Ouest	31	30.3	30.9	30.9	28.4
Centre et Sud	36.2	35.6	36	36	34.3
Taille du ménage		***			
			(17.09)	(30.28)	
Personnes seules	11.7	7.6	11.7	10.9	13.0
Deux personnes	16.9	13.8	16.9	16.3	19.8
Trois ou quatre personnes	52.8	56.7	52.8	53.5	50.3
Cinq personnes ou plus	18.6	21.9	18.6	19.3	16.9
Total	100	100	100	100	
Répondants	7004	1641	7004	8645	

***, **, *, ns : p-value du χ^2 <0.01, <0.05, <0.1 et non significatif quand on compare deux sous-échantillons (colonnes 2 et 3). Entre parenthèses et en italique : distance à la population cible (valeur du χ^2); (voir *Analyses statistiques* pour plus de détails).

Tableau 3. Prévalences de quelques variables d'intérêt selon la vague d'interview, le rang d'interview et le sexe (% et p-value du Khi²)

	Echantillon initial (vague 1)	Echantillon des vagues de rappel (vagues 2 et 3)	P	Rang d'interview 1-20 (toutes vagues)	Rang d'interview 21 et plus (toutes vagues)	P
Sous-échantillons	(1)+(2)	(3)+(4)		(1)+(3)	(2)+(4)	
Activité sexuelle au cours des douze derniers mois						
Hommes	94.6	95.1	ns	94.5	96.3	ns
Femmes	93.6	94.5	ns	93.7	94.8	ns
Cinq partenaires ou plus du sexe opposé au cours de la vie						
Hommes	58.6	56.7	ns	57.4	66.5	**
Femmes	35.2	36.2	ns	35.4	35.3	ns
Deux partenaires ou plus du sexe opposé au cours des douze derniers mois						
Hommes	17	16.2	ns	16.1	22.9	***
Femmes	8.9	8.2	ns	8.6	11.3	*
Utilisation de la pilule						
Femmes	46.4	47.2	ns	46.2	50.6	ns

Les prévalences sont calculées à partir de données pondérées (la pondération tient compte de la probabilité individuel d'être sélectionné et corrigée par la proportion de détenteurs de mobiles exclusifs par sexe dans la population cible.

***, **, *, ns: p-value du Khi² <0.01, <0.05, <0.1 et non significatif quand on compare les colonnes 2 et 3, et 4 et 5 respectivement.

Pour le nombre de partenaires au cours de la vie, ceci est dû aux différences d'effets de l'âge et du niveau de diplôme. Pour le nombre de partenaires au cours des douze derniers mois, l'âge et le type de téléphone semblent influencer d'une façon différente le comportement étudié. Nous avons également constaté que les facteurs associés à l'utilisation de la pilule au moment de l'enquête divergeaient entre les deux sous-échantillons : dans l'échantillon de primo-répondants, un niveau de diplôme élevé est positivement associé à l'utilisation de la pilule, alors que parmi les refusants initiaux, les femmes ayant un niveau de diplôme élevé sont moins enclines à être des utilisatrices de pilule.

Pour les hommes, les déterminants sociodémographiques des comportements sexuels étudiés ne diffèrent pas significativement entre les sous-échantillons. En outre, dans tous les modèles, une atténuation de la significativité des coefficients parmi les refusants initiaux est observée, ce qui pourrait refléter une plus grande homogénéité des comportements mais peut également résulter du fait que les échantillons sont de plus petite taille.

Tableau 4. Déterminants des comportements sexuels étudiés selon la vague (résultats de modèles de régression logistique)

Sous-échantillons	≥ 5 partenaires du sexe opposé au cours de la vie				≥ 2 partenaires du sexe opposé au cours des douze derniers mois			
	Hommes		Femmes		Hommes		Femmes	
	Vague 1 (1)+(2)	Vagues 2 et 3 (3)+(4)	Vague 1 (1)+(2)	Vagues 2 et 3 (3)+(4)	Vague 1 (1)+(2)	Vagues 2 et 3 (3)+(4)	Vague 1 (1)+(2)	Vagues 2 et 3 (3)+(4)
Age	***	***	***	***	*	ns	***	***
18-24	-0.95 ***	-0.79 *	-0.83 ***	-0.33	0.25	0.15	0.73 ***	2.39 ***
25-29	-0.34 **	0.22	-0.39 ***	-0.45	0.31	0.80	0.08	-0.06
30-34	Ref.	Ref.	Ref.	Ref.	Ref.	Ref.	Ref.	Ref.
35-39	-0.20	0.37	0.08	0.40	-0.45 *	-0.10	-0.04	0.62
40-44	-0.03	0.65 *	-0.05	0.48 *	-0.14	0.25	-0.62 **	0.00
45-49	0.06	0.74 **	0.01	-0.22	0.08	0.84	-0.40	-1.28
TEC: Khi² (p-value)	5.21 (0.391)		11.24 (0.047)		2.29 (0.807)		13.02 (0.022)	
Niveau de diplôme	***	ns	***	*	ns	ns	ns	ns
Aucun diplôme, CEP, BEPC	-0.23	-0.19	-0.09	-0.23	0.04	0.35	0.25	0.07
BEP, CAP	0.04	-0.16	-0.10	-0.64 ***	0.27	-0.35	-0.01	-1.03 **
Baccalauréat	Ref.	Ref.	Ref.	Ref.	Ref.	Ref.	Ref.	Ref.
Bac+1 à Bac+3	-0.12	0.04	0.32 ***	-0.24	0.04	-0.19	0.29 *	-0.42
Bac+4 et plus	-0.46 ***	-0.06	0.43 ***	-0.40 *	0.09	0.62	0.21	-1.08 **
TEC: Khi² (p-value)	3.57 (0.467)		12.16 (0.016)		5.43 (0.246)		7.56 (0.109)	
Situation professionnelle	ns	ns	ns	ns	ns	ns	*	ns
Actifs occupés	Ref.	Ref.	Ref.	Ref.	Ref.	Ref.	Ref.	Ref.
Inactifs/au chômage	-0.21	-0.06	-0.04	0.30	0.18	0.12	-0.27 *	-0.20
TEC: Khi² (p-value)	0.27 (0.602)		2.54 (0.111)		0.02 (0.878)		0.03 (0.862)	
Pays de naissance	ns	ns	ns	ns	ns	ns	ns	ns
France	Ref.	Ref.	Ref.	Ref.	Ref.	Ref.	Ref.	Ref.
Autres	-0.02	0.01	-0.48 ***	-0.14	-0.04	-0.19	-0.07	0.19
TEC: Khi² (p-value)	0.00 (0.954)		1.16 (0.281)		0.04 (0.842)		0.16 (0.692)	
Mode de vie	***	**	***	***	***	***	***	***
En couple	Ref.	Ref.	Ref.	Ref.	Ref.	Ref.	Ref.	Ref.
Avec les parents	0.18	0.96 **	0.17	-0.28	2.58 ***	3.54 ***	2.02 ***	0.90 *
Autres situations	0.80 ***	0.54 *	0.80 ***	0.63 ***	2.71 ***	3.54 ***	2.51 ***	2.74 ***
TEC: Khi² (p-value)	4.79 (0.091)		1.57 (0.455)		3.35 (0.187)		5.55 (0.062)	
Taille du ménage	ns	ns	***	***	*	ns	***	ns
1-2	Ref.	Ref.	Ref.	Ref.	Ref.	Ref.	Ref.	Ref.
3-4	-0.03	-0.10	-0.32 ***	-0.39 **	-0.22	-0.03	-0.20	0.00
5+	-0.31 *	-0.23	-0.85 ***	-0.88 ***	-0.63 **	0.63	-0.95 ***	0.13
TEC: Khi² (p-value)	0.26 (0.880)		0.12 (0.939)		4.64 (0.099)		3.08 (0.214)	
Lieu de résidence	***	ns	***	ns	ns	ns	ns	ns
IdF, bassin parisien	Ref.	Ref.	Ref.	Ref.	Ref.	Ref.	Ref.	Ref.
Nord, Est, Ouest	-0.42 ***	-0.06	-0.12	0.07	-0.20	-0.55	-0.10	-0.14
Centre et Sud	0.13	0.09	0.23 ***	0.23	-0.10	-0.36	0.10	0.04
TEC: Khi² (p-value)	2.45 (0.294)		1.13 (0.568)		0.76 (0.685)		0.02 (0.991)	
Téléphone	***	ns	ns	ns	ns	**	**	*
Fixe	Ref.	Ref.	Ref.	Ref.	Ref.	Ref.	Ref.	Ref.
Mobile	0.34 ***	0.30	0.17	-0.11	0.12	0.86 **	0.37 **	-1.25 *
TEC: Khi² (p-value)	0.01 (0.911)		0.80 (0.3708)		2.56 (0.109)		5.85 (0.016)	
Constante	0.71 ***	-0.01	-0.48 ***	-0.16	-2.88 ***	-3.73 ***	-3.41 ***	-3.69 ***
TEC: Khi² (p-value)	2.71 (0.100)		0.75 (0.386)		1.04 (0.307)		0.11 (0.734)	
Tous les coefficients sauf la constante								
TEC: Khi² (p-value)	17.89 (0.463)		34.59 (0.011)		15.13 (0.653)		33.91 (0.013)	
Tous les coefficients y compris la constante								
TEC: Khi² (p-value)	18.63 (0.481)		37.34 (0.007)		15.64 (0.681)		33.56 (0.016)	

Ref: Catégorie de référence; TEC: Test d'égalité des coefficients; IdF: Île-de-France;

***, **, *, ns: p-value du Khi² de Wald <0.01, <0.05, <0.1 et non significatif: quand on compare un coefficient à la référence.

En gras : Test du Khi² significatif au seuil de 5% pour la différence entre les coefficients des deux sous-échantillons (par exemple, dans la modélisation de ≥ 5 partenaires du sexe opposé au cours de la vie pour les femmes titulaires d'un BEP ou CAP, -0.10 comparé à -0.64).

Répondants difficiles à contacter

Comparés aux répondants « faciles » à enquêter, les difficiles à contacter sont plus souvent âgés de 25 à 39 ans, en activité, nés à l'étranger, vivant dans des ménages de petite taille et vivent moins souvent en couple ou avec leurs parents (tableau 5, partie gauche). L'intégration de ces difficiles à contacter conduit à une nette amélioration de la structure sociodémographique de l'échantillon (mesuré par une chute de la valeur des Khi^2), hormis pour le niveau de diplôme (tableau 5, partie droite).

De plus, les difficiles à contacter rapportent plus souvent avoir eu plusieurs partenaires sexuels du sexe opposé au cours des douze derniers mois (tableau 3, partie droite). Les hommes difficiles à contacter déclarent également plus souvent plus de quatre partenaires au cours de la vie.

Dans l'ensemble, les associations entre les caractéristiques sociodémographiques et les comportements sexuels étudiés ne diffèrent pas significativement selon le rang d'interview (tableau 6). Cependant, les tests effectués modalité par modalité révèlent que certaines caractéristiques sont différemment associées aux comportements sexuels selon le rang. L'âge, le niveau de diplôme (pour les femmes) et le pays de naissance (pour les hommes) sont différemment liés au nombre de partenaires sexuels au cours de la vie selon le rang d'interview. Par ailleurs, de rares différences sont observées dans les déterminants du nombre de partenaires au cours des douze derniers mois. Aucune différence, a contrario, n'a été observé pour les déterminants de l'usage de la pilule et de l'activité sexuelle au cours des douze derniers mois.

Tableau 5. Profil sociodémographique des répondants selon le rang d'interview (% , p-valeur du Khi² et valeur du Khi²)

Sous-échantillons	Rangs		Echantillon		
	1-20 (1)+(3)	21 et plus (2)+(4)	1-20 (1)+(3)	final (tous rangs) (1)+(2)+(3)+(4)	Population cible (recensement)
Sexe	ns				
Hommes	45.1	47.6	(15.25)	(13.64)	49.5
Femmes	54.9	52.4	54.9	54.7	50.5
Age	***				
			(28.63)	(22.75)	
15-17	10.1	8.6	10.1	9.9	8.1
18-24	20.1	19.7	20.1	20.1	18.4
25-29	11.5	14.2	11.5	11.7	13.4
30-34	11.6	14.5	11.6	11.9	14
35-39	14.7	15.8	14.7	14.7	15.3
40-44	15.5	14.9	15.5	15.4	15.5
45-49	16.5	12.4	16.5	16.1	15.2
Niveau de diplôme	ns				
			(116.38)	(117.55)	
Aucun diplôme, CEP, BEPC	18	16.4	18	17.8	25.9
BEP, CAP	22.1	23.9	22.1	22.2	24
Baccalauréat	23.2	22.3	23.2	23.2	20.4
Bac+1 à Bac+3	21.2	21.3	21.2	21.2	14.7
Bac+4 et plus	15.5	16.1	15.5	15.6	15
Situation professionnelle	***				
			(5.19)	(2.91)	
Actifs occupés	66.8	73	66.8	67.4	69.2
Inactifs/au chômage	33.2	27	33.2	32.6	30.8
Pays de naissance	***				
			(46.27)	(38.99)	
France	93.2	89	93.2	92.8	88.3
Autres	6.8	11	6.8	7.2	11.7
Mode de vie	***				
			(20.79)	(15.96)	
En couple	60.2	57.8	60.2	60	56.3
Avec les parents	22.1	18	22	21.6	21.9
Autres situations	17.8	24.2	17.8	18.4	21.8
Région de résidence	*				
			(18.34)	(15.71)	
Ile-de-France, Bassin parisien	32.7	36.5	32.7	33.1	37.3
Nord, Est, Ouest	31	30.2	31	30.9	28.4
Centre et Sud	36.3	33.3	36.3	36	34.3
Taille du ménage	***				
			(40.15)	(30.28)	
Personnes seules	10.6	13.8	10.6	10.9	13
Deux personnes	15.8	21.1	15.8	16.3	19.8
Trois ou quatre personnes	53.9	50.2	53.9	53.5	50.3
Cinq personnes ou plus	19.7	14.8	19.7	19.3	16.9
Total	100	100	100	100	
Répondants	7789	856	7789	8645	

***, **, *, ns : p-valeur du Khi² <0.01, <0.05, <0.1 et non significatif quand on compare deux sous-échantillons (colonnes 2 et 3).
Entre parenthèses et en italique : distance à la population cible (valeur du Khi²); (voir *Analyses statistiques* pour plus de détails).

Tableau 6. Déterminants des comportements sexuels étudiés selon le rang d'interview (résultats de modèles de régression logistique)

Sous-échantillons	≥5 partenaires du sexe opposé au cours de la vie				≥2 partenaire du sexe opposé au cours des douze derniers mois			
	Hommes		Femmes		Hommes		Femmes	
	Rangs 1-20 (1)+(3)	Rangs 21+ (2)+(4)	Rangs 1-20 (1)+(3)	Rangs 21+ (2)+(4)	Rangs 1-20 (1)+(3)	Rangs 21+ (2)+(4)	Rangs 1-20 (1)+(3)	Rangs 21+ (2)+(4)
Age	***	ns	***	ns	ns	ns	***	ns
18-24	-0.99 ***	-0.555	-0.81 ***	-0.12	0.09	1.21	1.09 ***	0.52
25-29	-0.27 *	-0.058	-0.37 ***	-0.79 *	0.18	1.44 **	0.24	-0.44
30-34	Ref.	Ref.	Ref.	Ref.	Ref.	Ref.	Ref.	Ref.
35-39	-0.11	0.040	0.16	-0.09	-0.41	-0.55	0.21	-0.77
40-44	0.08	0.441	0.09	-0.39	-0.28	1.03	-0.55 **	-0.68
45-49	0.18	0.731	0.00	-0.74 *	0.09	0.58	-0.45 *	-0.23
TEC: Khi² (p-value)	1.25 (0.940)		11.53 (0.041)		6.14 (0.293)		4.19 (0.523)	
Niveau de dipl	**	ns	***	ns	ns	ns	ns	ns
Aucun diplôme, CEP,								
BEPC	-0.32 **	0.45	-0.02	-1.03 **	-0.09	1.62 **	0.24	-0.20
BEP, CAP	-0.05	0.09	-0.22 **	-0.22	0.08	0.88	-0.14	-0.53
Baccalauréat	Ref.	Ref.	Ref.	Ref.	Ref.	Ref.	Ref.	Ref.
Bac+1 à Bac+3	-0.16	0.39	0.27 ***	-0.66 **	-0.12	0.95 *	0.11	0.43
Bac+4 et plus	-0.39 ***	-0.41	0.32 ***	-0.43	0.07	1.02	-0.02	0.32
TEC: Khi² (p-value)	3.48 (0.481)		13.78 (0.008)		6.72 (0.151)		2.24 (0.691)	
Situation professionnelle	ns	*	ns	ns	*	ns	ns	ns
Actifs occupés	Ref.	Ref.	Ref.	Ref.	Ref.	Ref.	Ref.	Ref.
Inactifs/au chômage	-0.10	-0.73 *	0.06	-0.26	0.28 *	-0.46	-0.20	-0.82
TEC: Khi² (p-value)	2.29 (0.131)		0.91 (0.341)		1.91 (0.166)		1.31 (0.252)	
Pays de naissance		**	***	**	ns	ns	ns	ns
France	Ref.	Ref.	Ref.	Ref.	Ref.	Ref.	Ref.	Ref.
Autres	0.11	-1.01 **	-0.36 ***	-0.87 **	-0.04	-0.25	-0.02	-0.71
TEC: Khi² (p-value)	4.69 (0.030)		1.40 (0.237)		0.08 (0.779)		0.93 (0.335)	
Mode de vie	***	ns	***	***	***	***	***	***
En couple	Ref.	Ref.	Ref.	Ref.	Ref.	Ref.	Ref.	Ref.
Avec les parents	-0.15	0.22	0.16	-0.31	2.69 ***	2.86 ***	1.80 ***	1.99 ***
Autres situations	-0.74 *	0.08	0.78 ***	0.81 ***	2.76 ***	3.12 ***	2.52 ***	2.39 ***
TEC: Khi² (p-value)	3.53 (0.171)		0.59 (0.744)		0.48 (0.788)		0.19 (0.908)	
Taille du ménage	ns	ns	***	ns	ns	ns	**	ns
1-2	Ref.	Ref.	Ref.	Ref.	Ref.	Ref.	Ref.	Ref.
3-4	0.00	-0.45	-0.33 ***	-0.25	-0.20	-0.44	-0.14	-0.26
5+	-0.25 *	-0.69	-0.83 ***	-0.81 *	0.42 *	-0.18	-0.84	0.75
TEC: Khi² (p-value)	1.57 (0.457)		0.10 (0.950)		0.46 (0.794)		5.68 (0.0585)	
Lieu de résidence	***	ns	***	ns	ns	ns	ns	ns
IdF, bassin parisien	Ref.	Ref.	Ref.	Ref.	Ref.	Ref.	Ref.	Ref.
Nord, Est, Ouest	-0.38 ***	-0.24	-0.06	-0.44	-0.28 *	-0.14	-0.15	0.13
Centre et Sud	0.11	0.19	0.24 ***	0.14	-0.13	-0.28	0.17	-0.54
TEC: Khi² (p-value)	0.15 (0.929)		1.42 (0.492)		0.37 (0.832)		4.17 (0.124)	
Telephone	***	ns	ns	ns	ns	ns	ns	ns
Fixe	Ref.	Ref.	Ref.	Ref.	Ref.	Ref.	Ref.	Ref.
Mobile	0.40 ***	-0.13	0.15	-0.18	0.24	0.03	0.22	-0.15
TEC: Khi² (p-value)	2.39 (0.122)		0.89 (0.346)		0.22 (0.636)		0.50 (0.477)	
Constant	**	*			**	**	**	
Constant	0.53 *	1.20 *	-0.48 ***	0.38	-2.82 *	-3.98 *	-3.53 *	-2.61 ***
TEC: Khi² (p-value)	1.46 (0.226)		3.62 (0.057)		1.79 (0.180)		1.53 (0.216)	
Tous les coefficients sauf la constante								
TEC: Khi² (p-value)	20.01 (0.332)		26.37 (0.092)		16.09 (0.586)		20.96 (0.281)	
Tous les coefficients y compris la constante								
TEC: Khi² (p-value)	26.76 (0.110)		26.44 (0.118)		18.96 (0.459)		24.28 (0.186)	

Légende : cf. Tableau 4

Comparaison des deux opérations: taille d'échantillon, biais et coûts de collecte

Pour définir l'allocation optimale des efforts de collecte et faciliter les décisions relatives au protocole de collecte de futures enquêtes, nous avons comparé les résultats obtenus selon différents scénarios. Dans le scénario 1, qui correspond au cas où aucune intervention n'est mise en œuvre pour réduire le taux de non-réponse, seuls les répondants interrogés durant les vingt premiers appels lors de la première vague sont considérés (échantillon 1). Dans le scénario 2, on ajoute les répondants difficiles à contacter de la première vague pour évaluer l'effet du nombre de tentatives élevé. Le scénario 3 combine toutes les vagues, mais ne sont pris en compte que les répondants interrogés lors des 20 premiers appels, pour évaluer la portée de la réduction du taux du refus. Le scénario 4 offre un aperçu de ce qu'aurait été l'échantillon si l'on avait appliqué la règle des tentatives d'appel « illimitées » uniquement lors de la première exploitation de la base (vague 1) puis rappelé les refusants en nous limitant à 20 appels. Le dernier scénario (scénario 5) combine tous les répondants et considère l'intégralité de l'échantillon recueilli. Le tableau 7 propose une synthèse des différents indicateurs retenus pour comparer les scénarios entre eux.

L'évolution de la taille des échantillons selon les scénarios montre que par rapport au scénario 1, le scénario qui consiste à rappeler les refusants mais ne pas augmenter le nombre de tentatives d'appels (scénario 3) conduit à une augmentation de la taille de l'échantillon de 25,0%, alors que l'augmentation de la taille de l'échantillon associée au scénario 2 (augmenter les tentatives d'appels sans rappeler les refusants) est seulement de 12,4%. La combinaison des deux opérations dans le scénario 4 conduit à augmenter la taille de l'échantillon de 37,4%. Par rapport au scénario 4, le scénario 5 (augmentation des tentatives d'appel sur toutes les vagues) ne conduit qu'à une augmentation de un point de pourcentage de la taille de l'échantillon.

Pour étudier l'impact de ces opérations sur le biais sociodémographique, nous avons comparé la structure de l'échantillon pour chaque scénario avec celle de la population cible pour les 8 variables utilisées dans les tableaux précédents (sexe, âge, niveau de diplôme, situation professionnelle, pays de naissance, mode de vie, région de résidence et taille du ménage et qui ont servi pour le calage sur marges). Pour ce faire nous avons comptabilisé le nombre de variables, parmi ces huit, pour lesquelles la distance du Khi^2 à la population cible était diminuée. Alors que l'augmentation des tentatives d'appel seule réduit considérablement le biais de sélection sur les caractéristiques sociodémographiques, le rappel des refusants conduit à une plus grande distorsion de l'échantillon que si aucune mesure n'est mise en place. La combinaison des deux interventions conduit à un résultat intermédiaire. Une autre façon d'illustrer le biais est d'utiliser les rapports de poids (poids après calage / poids de sondage) : plus le ratio est éloigné de 1, plus la structure de l'échantillon pondéré par le poids de sondage est différente de celle de la population cible (ce qui signe un biais de sélection important). Les statistiques sur les rapports de poids (coefficient de variation, maximum et minimum) peuvent être utilisées pour évaluer la distance de l'échantillon à la population cible. Par rapport au scénario 1, l'augmentation des tentatives d'appel (scénario 2) améliore

légèrement le coefficient de variation du rapport de poids (réduction de 1,5 point). Le rapport du ratio maximum sur le ratio minimum passe de 36,2% à 34,9%. A l'inverse, l'inclusion des refusants initiaux (scenario 3) n'améliore pas le coefficient de variation des rapports de poids.

Tableau 7. Comparaison des interventions méthodologiques

	Scenario 1	Scenario 2	Scenario 3	Scenario 4	Scenario 5 (échantillon final)
Sous-échantillon ^(a)	(1)	(1)+(2)	(1)+(3)	(1)+(2)+(3)	(1)+(2)+(3)+(4)
Taille des échantillons					
Taille de l'échantillon	6232	7004	7789	8561	8645
Effectifs de répondants supplémentaires par rapport au scénario 1: n (%)	-	772 (12.4%)	1557 (25.0%)	2329 (37.4%)	2413 (38.7%)
Mesures du biais sociodémographique					
<i>Comparaison avec la structure de la population cible</i>					
		(1)+(2) vs (1)	(1)+(3) vs (1)	(1)+(2)+(3) vs (1)	(1)+(2)+(3)+(4) vs (1)
Variables plus proches de la distribution de la population cible ^(b)	ref.	7 variables /8	2 variables /8	4 variables /8	4 variables /8
<i>Statistiques sur les rapports de poids (poids après calage / poids de sondage)</i>					
Ecart-type	0.51	0.50	0.51	0.49	0.48
Coeff. de variation	49.82	48.27	48.70	47.29	46.00
Min	0.17	0.17	0.16	0.16	0.17
Max	6.15	5.93	6.67	6.43	6.71
Max/min	36.18	34.88	41.69	40.19	39.47
Ecart interquartile	0.45	0.42	0.47	0.45	0.43
Coûts financiers de collecte					
		(1)+(2) vs (1)	(1)+(3) vs (1)	(1)+(2)+(3) vs (1)	(1)+(2)+(3)+(4) vs (1)
Coûts supplémentaires (%)	Ref.	24.56%	29.67%	54.23%	57.06%
Coût moyen par questionnaire					
- Par sous-échantillon	1	(2)=1.98	(3)=1.19	(2)=1.98; (3)=1.19	(2)=1.98; (3)=1.19; (4)=2.10
- Au total	1	(1)+(2)=1.11	(1)+(3)=1.04	(1)+(2)+(3)=1.12	(1)+(2)+(3)+(4)=1.13

Ref.: référence

(a) Voir la figure 1 pour la description des sous-échantillons

(b) Basé sur la différence entre la distance du Chi^2 entre le scénario étudié et la population cible d'une part, et le scénario 1 et la population cible d'autre part.

Note: le sous-échantillon 4 ne peut être considéré seul, il dépend des deux interventions méthodologiques (voir figure 1).

Pour finir, nous avons évalué les coûts financiers de chaque scénario. Le rappel des refusants (dans les scénarios 3, 4 et 5) conduit à une augmentation plus importante des coûts de collecte que l'augmentation des tentatives d'appels (scénario 2) mais conduit à inclure davantage de répondants. Au total, les coûts supplémentaires associés à ces interventions méthodologiques s'élèvent à plus de 57%. L'examen des coûts par questionnaire amène à des conclusions contraires : les questionnaires réalisés auprès des répondants difficiles à contacter (échantillons 2 et 4) sont significativement plus coûteux que les autres. Les refusants initiaux qui répondent dans les 20 premiers appels (échantillon 3) présentent des coûts par questionnaire substantiellement plus bas (1,19).

IV. Discussion

Résultats

Dans l'optique de réduire le biais potentiel de non-réponse, des opérations sont envisagées pour atteindre les individus les plus difficiles à enquêter (parce qu'ils sont injoignables ou réticents à répondre). Cette pratique repose sur l'hypothèse qu'il existe un continuum allant des individus les plus faciles à interroger aux non-répondants, les répondants difficiles à interroger devant ressembler davantage aux non-répondants. Dans cette approche, les répondants difficiles à interroger peuvent représenter les non-répondants. Mais cette hypothèse ne peut être vérifiée. Afin de déterminer la validité de cette hypothèse, l'étude distingue deux types (non-exclusifs) de répondants difficiles à enquêter : les difficiles à contacter (plus de 20 tentatives d'appel pour réaliser un questionnaire) et les refusants initiaux (ceux qui n'acceptent de répondre qu'après un ou plusieurs refus préalables) . Nous avons comparé les deux catégories afin de déterminer laquelle contribue le plus à la réduction du biais de l'échantillon. Nous avons trouvé que l'inclusion dans l'échantillon des refusants initiaux conduit à accroître la distorsion de l'échantillon, contrairement à celle des répondants difficiles à contacter qui tend à réduire le biais de sélection. Nous avons ensuite observé des différences significatives de comportements sexuels entre les répondants difficiles à contacter et les autres. A contrario, aucun écart significatif de comportements n'a été observé entre les répondants de la première vague et ceux des vagues de rappel. Concernant les déterminants sociodémographiques des comportements sexuels étudiés, ils diffèrent davantage entre les refusants initiaux et les autres qu'avec le rang d'interview, surtout chez les femmes. Les conclusions n'allant pas toutes dans le même sens, il n'est pas possible de conclure si une catégorie est plus proche des non-répondants que l'autre. Cependant, nous pouvons affirmer que les deux catégories sont différentes l'une de l'autre et ne sont pas substituables l'une à l'autre. Il semble donc important de maintenir les deux interventions pour assurer une bonne qualité de l'échantillon. Elles ont un impact complémentaire sur la qualité de l'échantillon. Par conséquent, la correction de la non-réponse par post-stratification est d'autant plus mauvaise que les répondants difficiles à interroger ne sont pas intégrés à l'échantillon (leur absence conduit à des biais importants).

Le taux de réponse de la seconde vague de rappel des refusants étant très bas (6,9%), un choix raisonnable serait de conduire une première exploitation de la base (vague 1) avec un nombre de tentatives d'appels élevé, puis mener une seule vague de rappel auprès des refusants mais avec un nombre de tentatives plus limité. Les coûts de collecte associés à cette recommandation ne devraient pas être supérieurs aux coûts actuels, voire inférieurs. Cependant, l'objectif de l'enquête (estimer des prévalences ou modéliser des comportements) joue également un rôle dans la décision finale.

Comparaison avec d'autres enquêtes

Les résultats obtenus dans cette étude sont dans la lignée de ceux observés sur d'autres enquêtes téléphoniques (Lynn and Clarke, 2002 ; Qayad, Chowdhury, Hu et Balluz, 2010), mais également sur des enquêtes en face-à-face (Chiu, Riddick et Hardy, 2001 ; Heerwegh, Koen et Loosdveldt, 2007). Si l'on se ramène à un contexte national, on peut comparer l'enquête FECOND au Baromètre Santé de 2000 et à l'enquête CSF (Contexte de la sexualité en France) réalisée en 2006. Dans ces deux enquêtes, une étude des répondants difficiles à contacter a été menée. Les auteurs concluaient que des différences significatives existaient entre les répondants difficiles à contacter et ceux faciles à enquêter (Beck, Legleye et Peretti-Watel, 2005 ; Razafindratsima, 2008). De plus, dans l'enquête CSF, portant sur des thèmes proches de ceux évoqués dans FECOND, les répondants difficiles à contacter déclaraient un nombre moyen de partenaires au cours de la vie plus élevé que les autres répondants de l'échantillon (Razafindratsima, 2008), ce qui corrobore les résultats obtenus ici. Lors de ces deux enquêtes, les refusants n'étaient pas rappelés, population pourtant grandissante dans les enquêtes nationales par téléphone.

Une étude de cette population et une comparaison avec les répondants difficiles à contacter a été menée sur six enquêtes menées par le gouvernement britannique (Durrant et Steele, 2009). Les auteurs ont montré les caractéristiques corrélées à la non-réponse étaient très différentes entre les deux sous-populations de difficiles à enquêter. Le fait d'être difficile à contacter semble plutôt lié à la tendance d'être chez soi, et de ce fait, associé aux caractéristiques du ménage et du mode de vie. Le fait de refuser de répondre à l'enquête reflète un phénomène social plus complexe, davantage déterminé par des caractéristiques individuelles comme le niveau d'éducation, le statut socioéconomique, l'âge... Ils ont également montré que la combinaison des deux interventions méthodologiques conduisait à balancer les effets et masquer les différences de profils des deux populations de non-répondants. Nous confirmons cette affirmation avec les données de l'enquête FECOND : les refusants initiaux et les répondants difficiles à contacter diffèrent significativement en termes d'âge, de situation professionnelle, de taille de ménage et chacune des deux catégories nécessite d'être incluse dans l'échantillon pour améliorer la qualité des données recueillies. Toutefois, leurs effets ne se contrebalancent pas autant que dans l'étude anglaise. Lynn et Clarke indiquaient également que le profil des individus difficiles à contacter est plus stable que celui des refusants (non lié au sujet de l'enquête).

Considérations éthiques

En menant cette étude, notre objectif était de proposer une description des coûts et bénéfices de l'inclusion des répondants difficiles à contacter ou refusants initiaux. Par rapport à l'argumentaire classique sur la déontologie, qui préconise de ne pas insister auprès des refusants, nous avons trouvé quelques arguments (taille de l'échantillon et différences dans les déterminants de quelques variables de

comportements) en faveur d'efforts visant à tenter de convaincre les refusants initiaux de participer. Cependant, même si les questionnaires réalisés auprès des refusants initiaux sont moins coûteux (à l'unité) que ceux des répondants difficiles à contacter, l'effet à long terme d'une telle pratique peut contribuer à réduire la confiance dans la recherche publique. Augmenter le nombre de tentatives est également une pratique pouvant être sujet à critique : appeler un grand nombre de fois un foyer peut en effet être perçu comme une forme de harcèlement. Ces efforts de collecte peuvent être vus comme ayant un « coût » pour les enquêtés et un coût sur le long terme pour les producteurs d'enquêtes et la recherche publique.

Limites

Les analyses préliminaires ayant montré peu de différences selon l'équipement téléphonique (ligne fixe versus mobile), nous avons opté pour une analyse conjointe des deux types de téléphone. L'inclusion d'un échantillon assez important de détenteurs de mobiles exclusifs (n=1305) dans l'enquête FECOND représente une intervention supplémentaire visant à améliorer la représentativité de l'échantillon. Cette population était encore considérée comme inatteignable dans les enquêtes il y a quelques années (Beck, Legleye et Peretti-Watel, 2004 ; Beck, Legleye et Peretti-Watel, 2005 ; Gautier et al., 2006 ; Hu et al., 2011). Cependant, le taux de refus auprès des détenteurs de mobiles exclusifs est plus élevé que pour les lignes fixes (33,1% versus 30,6% pour la première vague, 21,7% contre 17,5% pour l'échantillon final).

Comme tous les refusants et les personnes non-contacts ne sont pas inclus dans l'échantillon, la question du biais de non-réponse résiduel reste ouverte. De plus, la génération aléatoire de numéros de téléphone rend impossible l'usage d'indicateurs synthétiques de représentativité comme l'indicateur R (Schouten, Cobben et Bethlehem, 2009), nécessitant des données sur les non-répondants.

En dernier lieu, rappelons que les résultats présentés dans cet article doivent être interprétés avec précaution parce qu'ils portent uniquement sur une partie de la population (15-49 ans) et l'enquête traite d'un sujet sensible : les conclusions ne peuvent donc pas nécessairement être généralisables à d'autres thèmes de recherche ou d'autres populations.

Recherches futures

Malgré ces limites, nous pensons que ces résultats peuvent fournir de nouveaux éclairages sur les coûts et les bénéfices de tels efforts pour inclure des répondants difficiles à contacter ou refusants initiaux, non seulement pour améliorer le taux de réponse, mais également pour réduire la distorsion de l'échantillon et améliorer les estimations. De tels résultats nécessiteraient d'être enrichis par une étude des mécanismes de prises de décision des personnes sollicitées (Adua et Sharp, 2010 ; Groves, Singer et Corning, 2000 ; Singer, 2011 ; Wenemark, Persson, Brage, Svensson et Kristenson, 2011) afin d'améliorer la connaissance

des raisons de refus et de mettre en place des stratégies adaptées pour améliorer la participation aux enquêtes et réduire le biais de sélection, sans compromettre l'acceptation des personnes sollicitées. L'allocation des ressources afin d'inclure un échantillon de personnes difficiles à interroger, par exemple en face-à-face ou en faisant varier les modes de contact, afin de vérifier si elles diffèrent des répondants ultimes ou non, permettrait d'identifier s'il existe un continuum entre les individus « faciles » à interroger et les sous-catégories de personnes difficiles-à-interroger (par exemple, combinaison d'un nombre de tentatives élevé et/ou refusants). Ces résultats pourraient être pris en compte pour améliorer le processus de post-stratification, comme par exemple, la méthode de Hansen-Hurwitz (1946) ou éventuellement de valider la pratique classique.

V. References

- AAPOR. 2011. *Standard Definitions: Final Dispositions Of Case Codes And Outcome Rates For Surveys*: AAPOR.
- Adua, Lazarus, and Jeff S. Sharp. 2010. "Examining survey participation and response quality : the significance of topic salience and incentives". *Survey methodology*, June 2010, vol. 36, No. 1, pp. 95-109
- Beck, François, and Philippe Guilbert. 2007. "Baromètres santé : un éclairage sur leur méthode et leur évolution." Pp. 27-43 in *Baromètre santé 2005 : Attitudes et comportements de santé, Baromètre santé*, edited by F. Beck, P. Guilbert, and A. Gautier. Saint-Denis: INPES.
- Beck, François, Stéphane Legleye, and Patrick Peretti-Watel. 2004. "Using the telephone in general population surveys on drugs." Pp. 113-140 in *European studies on drugs and drug policy. Proceedings of the 14th International Conference of the European Society for Social Drug Research (ESSD) in Ghent from 2-4 October 2003*, edited by T. Decorte and D. J. Korf. Brussels: VUB Press.
- . 2005. "Aux abonnés absents : liste rouge et téléphone portable dans les enquêtes en population générale sur les drogues." *Bulletin de méthodologie statistique* n°86:pp. 5-29.
- Bigot, R and P Croutte. 2011. *La diffusion des technologies de l'information et de la communication dans la société française*. Paris: CREDOC.
http://www.arcep.fr/fileadmin/uploads/tx_gspublication/rapport-credoc-diffusion-tic-2011.pdf
- Brick, J. Michael, Ismael Flores Cervantes, Sunghye Lee, and Greg Norman. 2011. "Nonsampling errors in dual frame telephone surveys". *Survey methodology*, June 2011, vol. 37, No. 1, pp. 1-12
- Chiu, Pei-Lu, Howard Riddick, and Ann M. Hardy. 2001. "A comparison of characteristics between late/difficult and non-late/difficult interviews in the national health interview survey." in *Annual Meeting of the American Statistical Association*.
- Curtin, Richard, Stanley Presser, and Eleanor Singer. 2000. "The effects of response rate change on the index of consumer sentiment." *Public Opinion Quarterly* 64:413-428.
- Durrant, Gabrielle B and Fiona Steele. 2009. "Multilevel Modelling Of Refusal And Noncontact In Household Surveys: Evidence From Six Uk Government Surveys." *Journal Of The Royal Statistical Society, Series A* 172:361-381.
- Firdion, Jean-Marie. 1993. "L'effet du rang d'appel et de la présence du conjoint dans une enquête par téléphone." *Population* 48:1281-1314.
- Gautier, A, F Beck, S Marder, S Legleye, B Riandey, A Gayet, N Beltzer, and P Guilbert. 2006. "Téléphones portables exclusifs : résultats d'une méthode de génération partielle de numeros." Pp. 60-64 in

- Methodes d'enquêtes et sondages - Pratiques européenne et Nord-Américaine, Québec*, edited by P. Lavallée and L. Rivest. Québec: Dunod.
- Groves, Robert M. 2006. "Nonresponse Rates And Nonresponse Bias In Household Surveys." *Public Opinion Quarterly* 70:646-675.
- Groves, Robert M, Eleanor Singer, and A Corning. 2000. "Leverage-saliency theory of survey participation." *Public Opinion Quarterly*:299-308.
- Hansen, Morris H and William N Hurwitz. 1946. "The problem of non response in sample surveys." *Journal of the American Statistical Association* 41:517-529.
- Heerwegh, Dirk, Abts Koen, and Geert Loosveldt. 2007. "Minimizing survey refusal and noncontact rates: do our efforts pay off?" *Survey Research Methods* 1:3-10.
- Hu, S. S., L. Balluz, M. P. Battaglia, and M. R. Frankel. 2011. "Improving public health surveillance using a dual-frame survey of landline and cell phone numbers." *Am J Epidemiol* 173:703-11.
- Keeter, Scott, Carolyn Miller, Andrews Kohut, Robert M Groves, and Stanley Presser. 2000. "Consequences of reducing nonresponse in a national telephone survey." *Public Opinion Quarterly* 64:125-148.
- Lan, Romuald Le. 2009. "Enquêtes ménages : vers la fin de la baisse des taux de réponse ?" *Courrier des statistiques* n°128.
- Lynn, Peter and Patricia Clarke. 2002. "Separating Refusal Bias And Non-Contact Bias: Evidence From Uk National Surveys." *Journal Of The Royal Statistical Society Series D (The Statistician)* 51:319-333.
- Qayad, Mohammed G, Pranesh Chowdhury, Shaohua Hu, and Lina Balluz. 2010. "Respondent differences and length of data collection in the Behavioral Risk Factor Surveillance System". *Survey methodology*, December 2010, vol. 36, No. 1, pp. 223-227.
- Razafindratsima, N. 2008. "Rang d'appel et profil de l'échantillon : l'intérêt d'un nombre élevé d'appels." pp. 82-84 in *Enquête sur la sexualité en France : pratiques, genre et santé*, edited by N. Bajos and M. Bozon. Paris: Editions la découverte.
- Sautory, Olivier. 1993. "La macro Calmar. Redressement d'un échantillon par calage sur marges." Paris: INSEE. <http://www.insee.fr/fr/methodes/outils/calmar/doccalmar.pdf>
- Schouten, Barry, Fannie Cobben, and Jelke Bethlehem. 2009. "Indicators for the representativeness of survey response." *Survey methodology* n°35:pp. 101-113.
- Singer, Eleanor. 2011. "Toward A Benefit-Cost Theory Of Survey Participation: Evidence, Further Tests, And Implications." *Journal Of Official Statistics* 27:379-392.
- Wenemark, Marika, Andreas Persson, helle Nordlind Brage, Tommy Svensson, and Margareta Kristenson. 2011. "Applying motivation theory to achieve increased response rates, respondent satisfaction and data quality." *Journal of Official Statistics* 27:393-414.