

Mailys Goetschy, Université de Strasbourg, IDUS

Philippe Cordazzo, Université de Strasbourg, UMR 7363 SAGE

Topic. The impact of reporting and coding errors on the study of same-sex couples in France: illustration through the survey EPIC.

Short summary:

There are many methodological problems with studying same-sex couples in general population surveys. For example, reporting or coding errors concerning the sex of the respondent or his/her partner lead to "false" same-sex couples. From the EPIC survey (*Etude des Parcours Individuels et Conjugaux*), this study seeks to determine the extent to which people with misidentified couples impact outcomes in the same-sex couples study. To do this, after identifying people affected by these reporting or coding errors, by cross-checking the original variables, the imputed variables, and the variables after reconciliation, we will carry out a comparative analysis of the socio-demographic characteristics of misidentified individuals with respect to individuals in same-sex couples. Thus, we will observe how the results of studies on same-sex couples can be biased. We will also compare the results obtained with the main results of research on same-sex couples.

Sujet. L'impact des erreurs de déclaration et de codage sur l'étude des couples de même sexe en France : illustration au travers de l'enquête EPIC.

Résumé court :

L'étude des couples de même sexe à partir d'enquêtes en population générale pose de nombreux problèmes méthodologiques. Par exemple, les erreurs de déclaration ou de codage concernant le sexe du répondant ou de son(sa) conjoint(e) induisent de « faux » couples de même sexe. A partir de l'enquête EPIC (*Etude des Parcours Individuels et Conjugaux*), cette étude cherche à déterminer dans quelle mesure les individus dont le couple a mal été identifié impactent les résultats dans l'étude des couples de même sexe. Pour ce faire, après avoir identifié les individus concernés par ces erreurs de déclaration ou de codage, en recoupant les variables d'origine, les variables brutes, ainsi que les variables après apurement, nous réaliserons une analyse comparative des caractéristiques socio-démographiques des individus mal identifiés, au regard des individus en couple de même sexe. Ainsi, nous observerons comment les résultats des études sur les couples de même sexe peuvent être biaisés. Nous mettrons également en regard les résultats obtenus avec les principaux résultats des recherches sur les couples de même sexe.

Résumé long :

Sujet : L'impact des erreurs de déclaration et de codage sur l'étude des couples de même sexe en France : illustration au travers de l'enquête EPIC.

Contexte et enjeux :

Depuis les années 1970 en France, et plus largement en Europe, de profondes mutations ont lieu au sein des structures familiales (Régnier-Loilier, 2018). Les situations conjugales et familiales se diversifient et se complexifient. Le dénombrement des couples de même sexe, bien qu'il soit une opération particulièrement récente en démographie (Banens et Le Penven, 2013), s'inscrit pleinement dans ce contexte de transformations. La demande d'informations sur les couples de même sexe s'accroît considérablement depuis plusieurs années, liée à une reconnaissance juridique et sociale grandissante, et la littérature s'enrichit de connaissances.

La manière la plus fréquente d'identifier les couples de même sexe dans les enquêtes se fait par le recoupement du sexe du répondant et du sexe de son(sa) conjoint(e). Cependant, deux erreurs majeures subsistent : les erreurs de déclaration ou de codage sur le sexe qui induisent des « *faux* » couples de même sexe, ainsi que les individus qui ne se déclarent pas comme tel (Banens et Le Penven, 2013). Ces problèmes remettent en cause la robustesse des données, notamment en raison des effectifs particulièrement restreint que représentent les couples de même sexe (Régnier-Loilier, 2018).

Objet de la recherche :

Cette étude a pour objectif d'analyser dans quelle mesure ces individus en « *faux* » couple de même sexe impactent l'étude globale des couples de même sexe. Lorsque les erreurs de déclaration ou de codage conduisent à de nombreux couples mal identifiés, la surestimation des individus en couple de même sexe est importante (Festy, 2007 ; Algava et Hallépée, 2018). D'après Maks Banens et Eric Le Penven, et à partir de la comparaison du recensement français et de l'enquête *Famille et Logements* de 2011, 42 % des couples apparaissant comme couples de même sexe étaient en réalité des couples de sexe différent (Banens et Le Penven, 2018).

Au-delà du fait de surestimer le nombre d'individus en couple de même sexe, cette proportion considérable de couples mal identifiés induit nécessairement des biais dans les résultats. L'objet de cette étude sera donc d'analyser et de comparer les caractéristiques des individus réellement en couple de même sexe, par rapport aux individus en « *faux* » couple de même sexe, au regard de l'ensemble des individus *a priori* en couple de même sexe et en couple de sexe différent.

Données :

Les analyses s'appuieront sur les données de l'enquête EPIC (*Etude des Parcours Individuels et Conjugaux*), une enquête française réalisée par l'INSEE et l'INED entre la fin de l'année 2013 et le début de l'année 2014, auprès d'individus âgés de 26 à 65 ans et résidant en ménage ordinaire. L'enquête permet de reconstituer les parcours individuels et de les analyser, en saisissant la diversité des formes de la conjugalité ou encore les choix qui amènent les individus à vivre leur union sous une forme plutôt qu'une autre.

Elle offre une acception large et subjective de ce qu'est le couple : « *La difficulté à définir aujourd'hui ce qu'est un couple suggérerait plutôt de laisser les personnes interrogées décrire leur vie conjugale et amoureuse dans une acception large et volontairement subjective, en dehors de toutes notions préconçues : pas de limite de durée, pas de restrictions liées à la vie commune, mais simplement les histoires qui font sens à leurs yeux* » (Rault et Régnier-Loilier, 2019). A partir de ce positionnement, ce sont 5 607 répondants à l'enquête qui déclarent être en couple ou en relation amoureuse importante, dont 68 individus, *a priori*, en couple de même sexe. De plus, l'enquête EPIC dispose d'une table « *Histoire* » dont l'unité statistique est la relation, permettant ainsi d'augmenter la puissance statistique.

Méthodes de recherche :

Par un travail méthodologique de recoupement d'informations, et de comparaison entre les variables d'origine, les variables imputées et les variables après apurement, nous identifierons les individus en couple de sexe différent qui apparaissent en couple de même sexe dans l'enquête en raison d'erreurs de déclaration, de codage ou lors de l'apurement.

Dans un second temps, il s'agira d'effectuer une analyse comparative entre les individus qui sont réellement en couple de même sexe et les individus qui sont en « *faux* » couple de même sexe, et de déterminer dans quelle mesure ils sont susceptibles de venir biaiser les résultats, notamment concernant les caractéristiques socio-démographiques. Ainsi, nous utiliserons essentiellement des méthodes de statistiques descriptives. Enfin, nous mettrons en regard nos résultats avec les principaux résultats des recherches sur les couples de même sexe.

Résultats attendus :

Si l'on sait que les couples de même sexe ont des caractéristiques socio-démographiques distinctes des couples de sexe différent, la présence d'individus en « *faux* » couple de même sexe dans les enquêtes est susceptible de biaiser les résultats.

Ainsi, nous supposons que les caractéristiques des individus dont le couple a mal été identifié se distinguent des caractéristiques des individus en couple de même sexe, se rapprochant davantage des caractéristiques des individus en couple de sexe différent, et induisant ainsi des résultats biaisés.

Par exemple, une première analyse a permis de mettre en évidence que les individus en « *faux* » couple de même sexe impactent les résultats obtenus quant aux individus en couple de même sexe en termes d'âge : la proportion d'individus en « *faux* » couple de même sexe âgée de 60 ans ou plus est de 25 % tandis qu'elle n'est que de 1,9 % chez les individus en « *vrai* » couple de même sexe. La présence d'individus dont le couple a mal été identifié biaise également les résultats concernant le statut matrimonial : la moitié des individus en « *faux* » couple sont mariés, tandis que ce n'est le cas que de 9,6 % des individus réellement en couple de même sexe. Ainsi, la proportion d'individus mariés parmi l'ensemble des individus *a priori* en couple est de 16,2 %, soit une surestimation de 6,6 points de pourcentage. Il en est de même pour la plupart des caractéristiques socio-démographiques des individus.

Bibliographie :

Algava E., Hallépée S., 2018, Estimer les effectifs de couples de personnes de même sexe au recensement : expérimentation d'une solution de validation du sexe par le prénom, *Collection Documents de Travail*, Insee, 51 p.

Banens M., Le Penven E., 2018, « Etudier les couples de même sexe avec le recensement de la population », in Cauchi-Duval Nicolas, *Observer, décrire et analyser les structures familiales*, Paris, AIDELF, pp. 5-19.

Banens M., Le Penven E., 2013, « Erreurs de sexe, estimation des couples de même sexe et test d'une méthode de correction », pp.1-18.

Festy P., 2007, « Enumerating Same-Sex Couples in Censuses and Population Registers », *Demographic Research*, 17, pp. 339-368.

Rault W., Régnier-Loilier A., 2019, « Étudier les parcours individuels et conjugaux en France. Enjeux scientifiques et choix méthodologiques de l'enquête Épic », *Population*, 74(1-2), pp. 11-40.

Régnier-Loilier A., 2018, « Are the Generations and Gender Surveys Well Suited for Studying Same-Sex Couples? », *European Journal of Population*, 34(4), pp. 567-78.