

Participation et pondération de l'enquête nationale 10 ans ½

Afin d'analyser au mieux les données recueillies dans le cadre de l'enquête 10 ans 1/2, des variables de participation et de pondération ont été constituées.

Participation

5 variables de participation ont été constituées selon la participation des familles Elfe aux enquêtes CATI et CAPI.

Les variables **A10X_PARTIMEN10A_CATI** et **A10X_PARTIMEN10A_CAPI** renseignent simplement sur la participation ou non du ménage.

A10X_PARTIMEN10A_CATI	Fréquence	Pourcentage
0,0=le ménage ne fait pas partie de la base enquêtable	5242	28,6
1.0=le ménage est non participant (CATI) ET n'a pas fait le Q minimal	2622	14,31
1.1=le ménage est non participant (CATI) - Q minimal rempli	658	3,59
2.0=le ménage participe (CATI commencé)	9807	53,51

A10X_PARTIMEN10A_CAPI	Fréquence	Pourcentage
0.0=le ménage ne fait pas partie de la base enquêtable	5242	28,6
1.0=le ménage est non participant (CAPI)	5338	29,12
1.1=le ménage participe mais PAS de consentement conforme	261	1,42
2=le ménage participe (CAPI commencé)	7488	40,85

La variable **A10X_PQ_PARTIMEN10A** tente de caractériser les causes de non-participation (c'est le statut du dernier appel tenté pour joindre un ménage n'ayant pas participé au CATI). Cette variable est bien évidemment remplie uniquement pour les ménages éligibles non participants.

A10X_PQ_PARTIMEN10A	Fréquence	Pourcentage
ABANDON	1	0,01
ABANDON DU QUESTIONNAIRE	145	0,79
FAUX NUMÉRO	512	2,79
HC RÉPOND CAWI	5	0,03
NON UTILISÉ	37	0,2
NRP/OCCUPÉ	1397	7,62
NUISANCE	2	0,01
RDV	358	1,95
STOP ABANDON ELFE/EPIPAGE	412	2,25
STOP ACCEPT=3 REFUS ENTRETIEN TÉL	5	0,03
STOP AUTRE PARENT NE VOIT PLUS L'ENFANT	8	0,04
STOP ENFANT ELFE/EPIPAGE DÉCÉDÉ	3	0,02
STOP ENFANT PLACÉ	7	0,04
STOP INTER OPPOSITION ENTRETIEN TÉL	2	0,01
STOP NONOPPOS=3	8	0,04
STOP PARENT DÉCÉDÉ ABANDON ELFE	1	0,01
STOP PARENT IMPOSSIBLE À INTERROGER	121	0,66
STOP RAISON(2)=2 RECHERCHE MANUELLE À EFFECTUER (NSP)	6	0,03
STOP RAISON(2)=3 Recherche manuelle à effectuer (refus)	2	0,01
STOP REFUS 10,5 ans seulement	83	0,45
STOP REFUS DE NOUS PASSER LE PARENT (CONTACT=12)	164	0,89
STOP REFUS DE NOUS PASSER LE PARENT (RAISON=5)	1	0,01

Enfin, les variables **A010X_REFCOMP10A_CATI** et **A010X_REF2COMP10A_CATI** caractérisent la participation et la complétude des 2 parents au CATI.

	A010X_REFCOMP10A_CATI	A010X_REF2COMP10A_CATI
0=Sans objet, le ménage ne participe pas	8522	8522
1=Questionnaire REF complet	9663	613
2=Questionnaire REF incomplet	144	21
3=Questionnaire REF non fait		222
4=Sans objet, pas de REF2 (efvit ne 4)		8951

Pondération

Les pondérations ont pour objectif de permettre d'inférer les estimations produites à partir des données fournies par les seuls répondants à la population cible (dans le cas de la cohorte Elfe, la population d'inférence est celle des nourrissons nés durant l'année 2011 en France métropolitaine, issus d'un accouchement au plus gémellaire, hors grands prématurés, ayant une mère majeure, en mesure de donner un consentement éclairé notamment dans l'une des langues proposées, nés dans une maternité métropolitaine et dont les parents ne résidaient pas en métropole seulement de façon temporaire). Le total de la population cible est fixé à 753 500 familles (et 764 000 enfants). Ainsi, pour corriger le biais introduit par la non-réponse (sauf à supposer que les répondants et les non répondants ont des caractéristiques identiques), on utilise dans les enquêtes Elfe la repondération.

Contrairement aux étapes précédentes des enquêtes, la pondération longitudinale (pondération concernant les enfants ayant participé à TOUTES les enquêtes) n'est plus proposée. Seule la transversale est réalisée.

En effet, la composition des enfants ayant participé à toutes les enquêtes risque de ne plus assurer la représentativité de la population des enfants nés en 2011. De plus, ces poids longitudinaux ne seraient utilisés que pour les analyses demandant d'exploiter des variables sur toutes les vagues d'enquête, ce qui a peu de chance de se produire. Ainsi, plutôt que de proposer une pondération aboutissant à des résultats difficilement exploitables voir inutilisables (le manque d'effectif sur certaines sous populations entrainerait le calcul de poids extrêmes ; tronquer ces poids pour éviter des poids trop important risquerait de biaiser fortement les analyses), l'équipe Elfe a choisi de ne pas proposer de pondération longitudinale. Des pondérations spécifiques pourront évidemment être réalisées à la demande en cas de besoin (population d'enfants participants à certaines vagues spécifiques par exemple).

CATI

Dans le cas de l'enquête 10 ans ½, le premier contact correspond, pour tous les enfants, à une enquête téléphonique réalisée par CATI auprès de leur parent.

On a donc choisi de débiter le processus de pondération en considérant uniquement les enfants dont au moins 1 parent avait rempli un questionnaire CATI complet, soit 9663 enfants. La première pondération mise à disposition pour ces enfants est la pondération transversale CATI (**A10E_PONDREF_CATI**).

Tout d'abord, une modélisation logistique a permis d'identifier les variables significatives dans le processus de non réponse (entre les enfants inclus dans la cohorte Elfe en maternité, soit 18329 enfants et les répondants complets au CATI 10 ans 1/2).

Les variables suivantes ont été retenues (*pvalue* en rouge) :

Effet	DDL	Khi-2	Pr > khi-2
		de Wald	
CS_1='Region de la maternité'	4	26,814	<,0001
CS_2='Mère Primipare'	1	22,9195	<,0001
CS_3='Etat matrimonial'	1	4,3117	0,0379
CS_4='Age mère '	3	111,745	<,0001
CS_5='Niveau etudes mère'	3	535,026	<,0001
CS_6='Statut immigré mère'	1	128,167	<,0001
CS_7='séance de préparation à la naissance'	1	302,166	<,0001
CS_8='Activité du père au moment de l'accouchement'	1	22,9876	<,0001
CS_9='Mère vivant en couple à la naissance'	1	52,5733	<,0001
CS_10='Age pere'	3	31,1643	<,0001
CS_11='Consommation d alcool pendant la grossesse'	1	30,0111	<,0001
CS_13='Activité de la Mère au moment de la grossesse'	1	96,6446	<,0001
CS_14='sexe'	1	0,8309	0,362

Seul le sexe de l'enfant n'est pas significatif. Les variables prises en compte sont donc les suivantes :

- Groupe de région du domicile en 5 groupes (Île-de-France, Centre, Picardie/Nord-Est/Nord-Ouest/Sud-Est/Sud-Ouest)
- Mère primipare (oui/non)
- État matrimonial (né dans le mariage/hors mariage)
- Âge de la mère (18-24/25-29/30-34/35 et +)
- Niveau études de la mère (non scolarisée, primaire, CAP, BEP/2nd, 1ère, terminale/bac à bac +3/bac +3 et plus)
- Statut mère immigrée (oui/non)
- Séances de préparation à l'accouchement (oui/non)
- Activité du père au moment de l'accouchement (en emploi/autre)
- Mère vivant en couple à la naissance (oui/non)
- Âge du père (18-24/25-29/30-34/35 et +)
- Consommation d'alcool pendant la grossesse (oui/non)
- Activité de la mère au moment de l'accouchement (en emploi/autre)

20 groupes de réponses homogènes ont été constitués (les probabilités de participer varient de 12.3 à 75.5%).

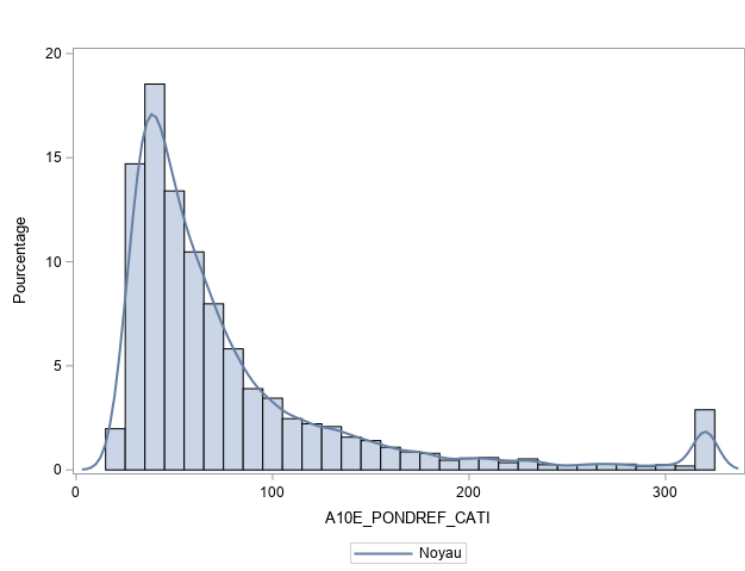
Le poids initial est ainsi construit en multipliant le poids de tirage (qui prend simplement en compte la taille de la maternité en 5 strates et la vague de naissance en 4 strates) par l'inverse de la probabilité de participer du GRH auquel appartient l'enfant participant.

Ce poids initial est ensuite redressé par la méthode CALMAR de l'INSEE, le calage étant réalisé sur les 12 variables sélectionnées précédemment auxquelles on a ajouté le sexe de l'enfant et le statut gémeleire de la naissance (oui/non).

Pour ces nouvelles variables, les données manquantes (5% pour l'âge du père, moins de 2% pour les autres) ont été imputées avant le calage. L'ensemble des totaux sur la population Elfe sont issus de l'état civil ou de l'ENP de 2010.

Enfin, les poids extrêmes (supérieurs à la moyenne +2 écarts type, soit 300 ont été tronqués – 3.5% des poids), l'ensemble étant ensuite redressé pour conserver le total de 764000 enfants.

Variable d'analyse : A10E_PONDREF_CATI													
N	Moyenne	Maximum	Minimum	Intervalle	Somme	5ème ctl	10ème ctl	25ème ctl	50ème ctl	75ème ctl	90ème ctl	95ème ctl	99ème ctl
9 663	79,06	320,52	19,48	301,04	764 000	27,97	31,39	39,22	56,20	89,90	155,23	231,10	320,52



CAP1

Dans le cas de l'enquête 10 ans ½, on a également proposé aux ménages joints lors de l'enquête téléphonique (même si celle-ci n'était pas considérée comme complète) de recevoir un enquêteur au domicile pour proposer d'autres questions à l'enfant. On propose donc une seconde pondération en considérant les enfants ayant rempli le CAPI et pour lesquels les parents ont rempli le consentement demandé pour autoriser la visite à domicile, la conservation et l'utilisation de ces données, soit 7488 enfants. La pondération mise à disposition pour ces enfants est là encore la pondération transversale (**A10E_PONDREF_CAPI**).

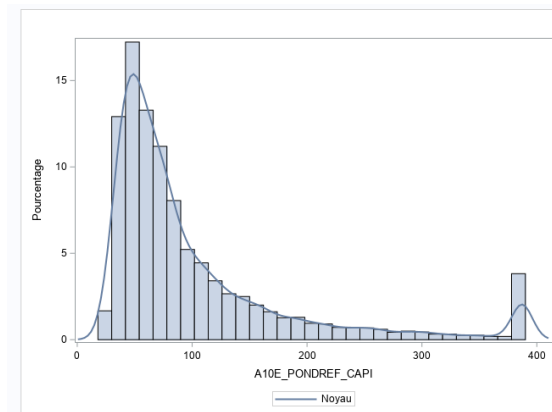
Comme pour le CATI, une modélisation logistique a permis d'identifier les variables significatives dans le processus de non réponse (entre les enfants inclus dans la cohorte Elfe en maternité, soit 18329 enfants et les répondants complets au CAPI 10 ans 1/2). Les mêmes variables significatives ont été retenues que pour le CATI, à savoir là encore :

- Groupe de région du domicile en 5 groupes (Île-de-France, Centre, Picardie/Nord-Est/Nord-Ouest/Sud-Est/Sud-Ouest)
- Mère primipare (oui/non)
- État matrimonial (né dans le mariage/hors mariage)
- Âge de la mère (18-24/25-29/30-34/35 et +)
- Niveau études de la mère (non scolarisée, primaire, CAP, BEP/2nd, 1ère, terminale/bac à bac +3/bac +3 et plus)
- Statut mère immigrée (oui/non)
- Séances de préparation à l'accouchement (oui/non)
- Activité du père au moment de l'accouchement (en emploi/autre)
- Mère vivant en couple à la naissance (oui/non)
- Âge du père (18-24/25-29/30-34/35 et +)
- Consommation d'alcool pendant la grossesse (oui/non)
- Activité de la mère au moment de l'accouchement (en emploi/autre)

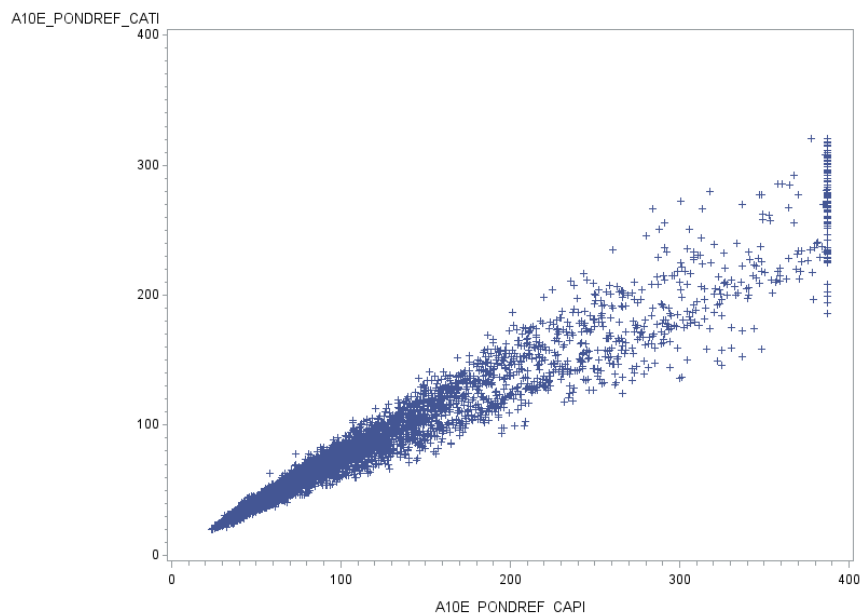
20 groupes de réponses homogènes ont été constitués, et les poids initiaux ont été redressés par la méthode CALMAR de l'INSEE, le calage étant réalisé sur les 12 variables sélectionnées précédemment auxquelles on a ajouté le sexe de l'enfant et le statut gémellaire de la naissance (oui/non).

Là encore, les poids extrêmes (supérieurs à la moyenne +2 écarts type, soit 350 ont été tronqués – 3.3% des poids), l'ensemble étant ensuite redressé pour conserver le total de 764000 enfants.

Variable d'analyse : A10E_PONDREF_CAPI A10E_PONDREF_CAPI													
N	Moyenne	Maximum	Minimum	Intervalle	Somme	5ème ctl	10ème ctl	25ème ctl	50ème ctl	75ème ctl	90ème ctl	95ème ctl	99ème ctl
7488	102,03	387,35	23,39	363,95	764 000,00	34,25	38,74	48,93	70,67	117,48	213,98	318,16	387,35

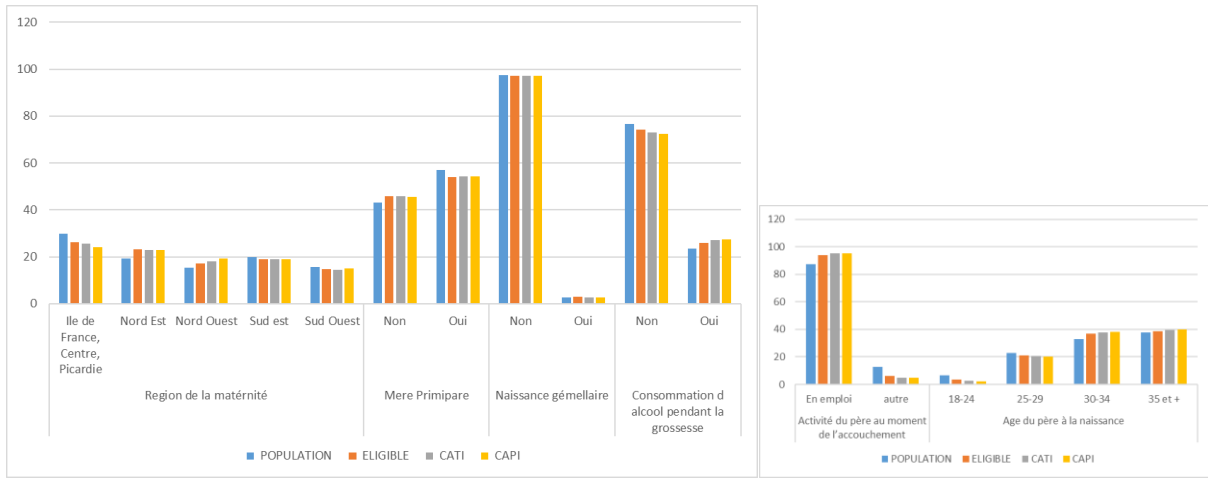


Les variables retenues pour les 2 pondérations étant identiques, et comme il y a peu de différences majeures entre les 2 échantillons de répondants, on note une corrélation très forte entre les poids calculés **A10E_PONDREF_CAPI** et **A10E_PONDREF_CATI**.

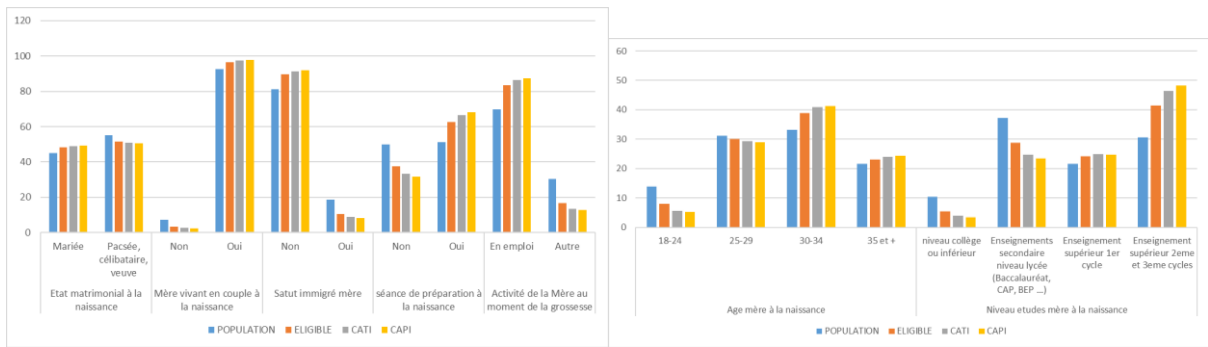


Qui sont les répondants à cette phase d'enquête

Pour essayer de caractériser les répondants aux enquêtes CATI et CAPI du 10 ans 1/2, on peut comparer, pour l'ensemble des variables de calage, la répartition sur la population cible (mesurée comme étant égale aux marges de calage), sur les éligibles aux enquêtes (c'est-à-dire sur ceux encore suivis à cet âge - permettant d'estimer les perdus de vue avant cette phase) et les répondants aux 10 ans ½ (afin d'estimer l'attrition spécifique à la dernière enquête).



Comparaison des différentes populations – données familiales et sur le père



Comparaison des différentes populations – données sur la mère

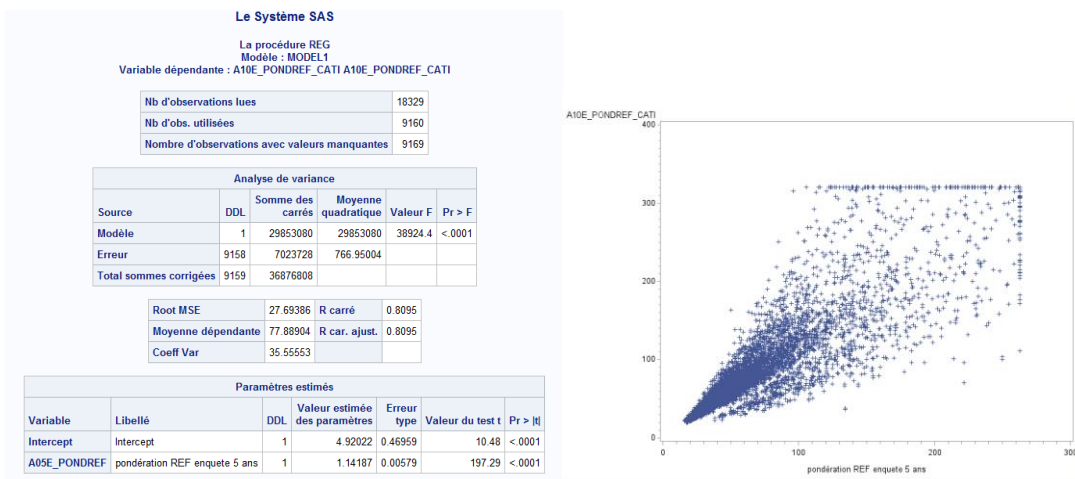
On peut constater avec les graphiques suivants quelques phénomènes bien connus :

- Une certaine population est de toute façon non répondante – les jeunes parents, immigrés, peu diplômés et sans emploi. Cette population a abandonné l'enquête au cours du temps (dès le recrutement ou au cours du temps lorsque l'attrition s'est accélérée).
- Puis enfin une certaine population a quitté l'enquête lors de cette seule dernière phase aux 10 ans ½ (lorsque les barres grise et jaune sont inférieures à la barre orange).

Pour tenter d'estimer cette dernière attrition, on peut se concentrer sur les différences entre les seules populations éligibles VS ayant accepté le 10 ans 1/2, on note peu de différences majeures.

		ELIGIBLE	CATI			ELIGIBLE	CATI
Region de la maternité	Ile de France, Centre, Picardie	26,1	25,6	Etat matrimonial à la naissance	Mariée	48,4	49,0
	Nord Est	23,2	22,8		Pacsée, célibataire, veuve	51,7	51,0
	Nord Ouest	17,1	18,0	Mère vivant en couple à la naissance	Non	3,4	2,5
	Sud est	18,8	19,1		Oui	96,6	97,5
	Sud Ouest	14,7	14,6	Satut immigré mère	Non	89,5	91,3
Mere Primipare	Non	46,0	45,7	Oui	10,5	8,7	
	Oui	54,1	54,3	Age mère à la naissance	18-24	8,1	5,7
Naissance gémellaire	Non	97,1	97,2		25-29	30,0	29,3
	Oui	3,0	2,8		30-34	38,9	40,9
					35 et +	23,1	24,1
Consommation d'alcool pendant la grossesse	Non	74,1	72,9		Niveau études mère à la naissance	niveau collège ou inférieur	5,5
	Oui	25,9	27,1	Enseignements secondaire niveau lycée (Baccalauréat, CAP, BEP...)		28,8	24,7
Activité du père au moment de l'accouchement	En emploi	93,8	95,0	Enseignement supérieur 1er cycle	24,2	25,0	
	autre	6,2	5,0	Enseignement supérieur 2eme et 3eme cycles	41,5	46,4	
Age du père à la naissance	18-24	3,6	2,5				
	25-29	21,2	20,5				
	30-34	36,7	37,6				
	35 et +	38,6	39,4				

À noter enfin, lorsqu'on compare, pour les 9160 enfants dont le parent référent a répondu aux enquêtes 5 ans ½ et 10 ans ½ la part de la population que chacun d'eux représente (le poids qu'on leur attribue), une corrélation assez forte de ces 2 pondérations, ce qui permet d'assurer une certaine continuité des analyses (une variabilité d'un phénomène mesurée aux 5 ans ou aux 10 ans des enfants résultera bien d'une variabilité du phénomène, et non d'une variabilité importante des pondérations).



Pondération de l'enquête nationale 10 ans ½ pour mise en commun avec les données recueillies auprès des enfants de la cohorte Epipage2

Le recueil des informations à 10 ans ½ ayant été réalisé en commun avec les enfants de la cohorte Epipage2, afin de rendre les populations cibles Elfe et Epipage2 totalement disjointes et complémentaires, on propose une pondération complémentaire en excluant des analyses Elfe les enfants nés avant 35 SA (on a donc un tout petit moins de participants « bruts » que d'individus pondérés, cela permettant facilement d'agréger les 2 échantillons).

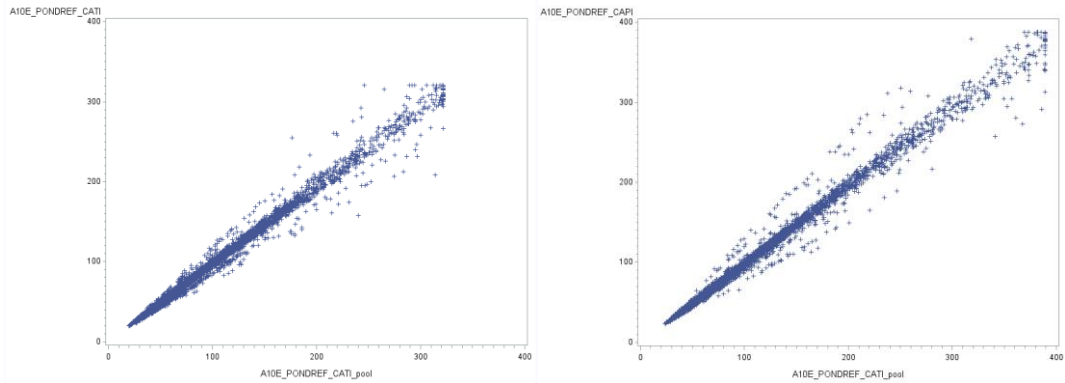
On dispose donc par Elfe d'une estimation des enfants nés en 2011 à 35 SA et plus, et via Epipage2 d'une estimation des enfants nés à 34 SA et moins.

Pour le CATI, ce sont donc 9412 enfants qui ont été considérés et 7307 pour le CAPI (soit 97.5% des enfants Elfe ayant participé aux enquêtes).

Les pondérations génériques et pour les bases communes sont évidemment très proches.

Variable d'analyse : A10E_PONDREF_CATI_pool A10E_PONDREF_CATI_pool													
N	Moyenne	Maximum	Minimum	Intervalle	Somme	5ème ctl	10ème ctl	25ème ctl	50ème ctl	75ème ctl	90ème ctl	95ème ctl	99ème ctl
9412	81,17	321,88	19,87	302,01	764 000,00	28,49	32,05	40,16	57,77	92,61	160,67	238,94	321,88

Variable d'analyse : A10E_PONDREF_CAPI_pool A10E_PONDREF_CATI_pool													
N	Moyenne	Maximum	Minimum	Intervalle	Somme	5ème ctl	10ème ctl	25ème ctl	50ème ctl	75ème ctl	90ème ctl	95ème ctl	99ème ctl
7307	104,56	389,13	23,87	365,26	764 000,00	35,11	39,74	50,43	72,79	121,70	215,95	333,86	389,13



La table PONDÉRATION comprend donc 4 jeux de poids :

A10E_PONDREF_CATI pour la pondération à appliquer sur les répondants au CATI si on utilise uniquement les bases Elfe

A10E_PONDREF_CAPI pour la pondération à appliquer sur les répondants au CAPI si on utilise uniquement les bases Elfe

A10E_PONDREF_CATI_POOL pour la pondération à appliquer sur les répondants au CATI si on utilise les bases Elfe et Epipage2

A10E_PONDREF_CAPI_POOL pour la pondération à appliquer sur les répondants au CAPI si on utilise les bases Elfe et Epipage2

Variable	Libellé	N	Minimum	Maximum	Moyenne	Somme	e_type
A10E_PONDREF_CATI	A10E_PONDREF_CATI	9 663,00	19,48	320,52	79,06	764 000,00	64,63
A10E_PONDREF_CATI_pool	A10E_PONDREF_CATI_pool	9 412,00	19,87	321,88	81,17	764 000,00	65,87
A10E_PONDREF_CAPI	A10E_PONDREF_CAPI	7 488,00	23,39	387,35	102,03	764 000,00	84,42
A10E_PONDREF_CAPI_pool	A10E_PONDREF_CATI_pool	7 307,00	23,87	389,13	104,56	764 000,00	85,67