



Enquête nationale Pondérations aux 5 ans ½ de l'enfant

Thierry Siméon

Table des matières

1. Pondération de l'enquête nationale 5 ans ½	2
2. Pondération transversale	4
3. Pondération longitudinale.....	7

1. Pondération de l'enquête nationale 5 ans ½

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. Celle-ci provient de la méthode de pondération mise en œuvre pour l'enquête 3 ans ½ et les suivantes¹ (méthode dite calage simultané).

Dans le cas de l'enquête 5 ans ½, même si le mode de collecte est le même pour tous (téléphonique, par CATI), les questionnaires diffèrent selon la situation du couple et de l'enfant au moment de l'enquête (parents cohabitants ou non, parent détenant la garde de l'enfant, enfant placé ou non, enfant hospitalisé ou non) ainsi que sur certaines de leurs réponses aux enquêtes précédentes. Pour les pondérations, on distingue donc deux échantillons utiles : **les enfants dont au moins un parent a répondu (questionnaire référent complet) et les enfants dont les deux parents cohabitants ont répondu (questionnaires référent et cohabitant complets)**. Une pondération est applicable uniquement au 1^{er} échantillon (qui représente la totalité des répondants et de la population cible) nommée PONDREF.

Il faut signaler qu'une autre catégorie de répondants existe : les parents non-cohabitants (dans le cas d'une garde alternée par exemple). Il y a cependant trop peu de cas pour qu'une pondération de cette sous-population soit pertinente dans le cadre général évoqué ici.

Il faut enfin signaler que, contrairement aux pondérations réalisées avant l'enquête aux 3 ans ½ de l'enfant, et puisque la pondération de l'enquête nationale 5 ans ½ est basée sur la « nouvelle » méthode de pondération (le taux de jumeaux – 2,75% des enfants, donc 21 000 jumeaux, appartenant à 10 500 familles – étant fixé par les marges utilisées), il n'est pas utile de réaliser 2 pondérations distinctes pour les familles et les enfants. La pondération de base permet à la fois de travailler à ces 2 niveaux en sélectionnant simplement si on souhaite travailler au niveau famille un unique enfant dans chacune d'elle (avec la variable construite M00M1_RANGALEA par exemple).

On propose donc d'utiliser la méthode de calage simultané sur 13 variables.

Variables prises en compte :

6 variables dites « de contexte » (variables déjà prises en compte dans le principe de calage depuis l'enquête maternité) :

- 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 d'études de la mère (non scolarisée, primaire, CAP, BEP/2nd, 1^{ère}, terminale/études supérieures)
- Statut mère immigrée (oui/non)

¹Pour la description précise de la méthodologie, se reporter à [Enquête Elfe : Pondérations des enquêtes nationales – T. Siméon](#).

Auxquelles on ajoute les variables suivantes permettant de prendre en compte l'attrition et sélectionnées depuis les 3 ans ½ :

- 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)
- Naissance gémellaire (oui/non)
- Activité de la mère au moment de l'accouchement (en emploi/autre)

Signalons que 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. Comme pour les méthodes précédentes, l'ensemble des totaux sur la population Elfe sont issus de l'état civil ou de l'ENP de 2010.

Ainsi, 2 pondérations sont réalisées : 1 pondération transversale et 1 pondération longitudinale.

En fonction de la provenance des variables mobilisées dans l'analyse à réaliser, l'utilisateur choisira :

- **si les variables proviennent exclusivement d'une unique étape d'enquêtes, la pondération transversale associée à cette étape ;**
- **si les variables proviennent d'étapes d'enquêtes différentes, la pondération longitudinale associée à la dernière étape des enquêtes mobilisées.**

Remarque : il se peut que la base de données livrées contienne moins d'individus que ceux dénombrés dans ce document, les enquêtés ayant la possibilité de demander la destruction de leurs données. Le phénomène, rare, est supposé ne pas avoir d'impact sur la pondération.

2. Pondération transversale

La première pondération mise à disposition est la pondération **transversale** (A05E_PONDREF). Elle concerne les enfants dont le parent référent a répondu au questionnaire 5 ans ½ et dont le questionnaire est considéré comme complet, soit **11 220** enfants.

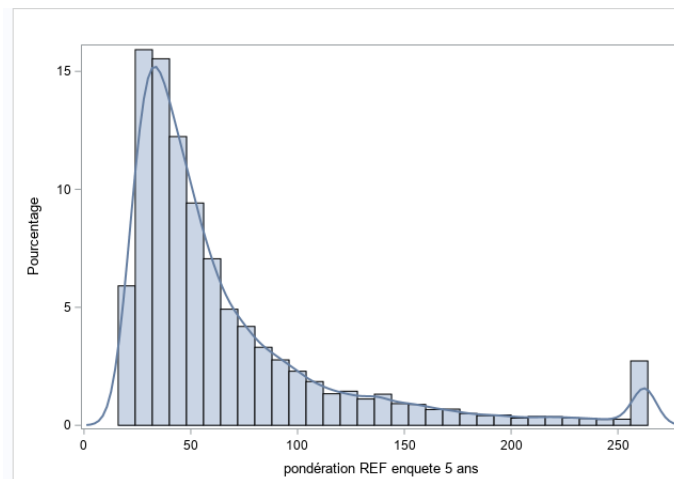
Afin de ne pas avoir de poids trop extrême, une troncature à 250 est mise en place (les poids supérieurs sont limités à 250). L'ensemble est ensuite redressé pour obtenir le bon total de 764 000 enfants).

Les résultats de cette pondération sont les suivants :

Mesures statistiques de base			
Location		Variabilité	
Moyenne	68,10	Ecart-type	54,28
Médiane	48,35	Variance	2 946,00
Mode	262,55	Intervalle	246,06
		Ecart interquartile	46,24

Quantiles (Définition 5)	
Niveau	Quantile
100Max 100%	262,55
99%	262,55
95%	198,09
90%	139,69
75% Q3	79,55
50% Médiane	48,35
25% Q1	33,31
10%	26,40
5%	23,31
1%	19,63
0% Min	16,50

À noter que la troncature, au-delà de 250, concerne 2,6% des individus.



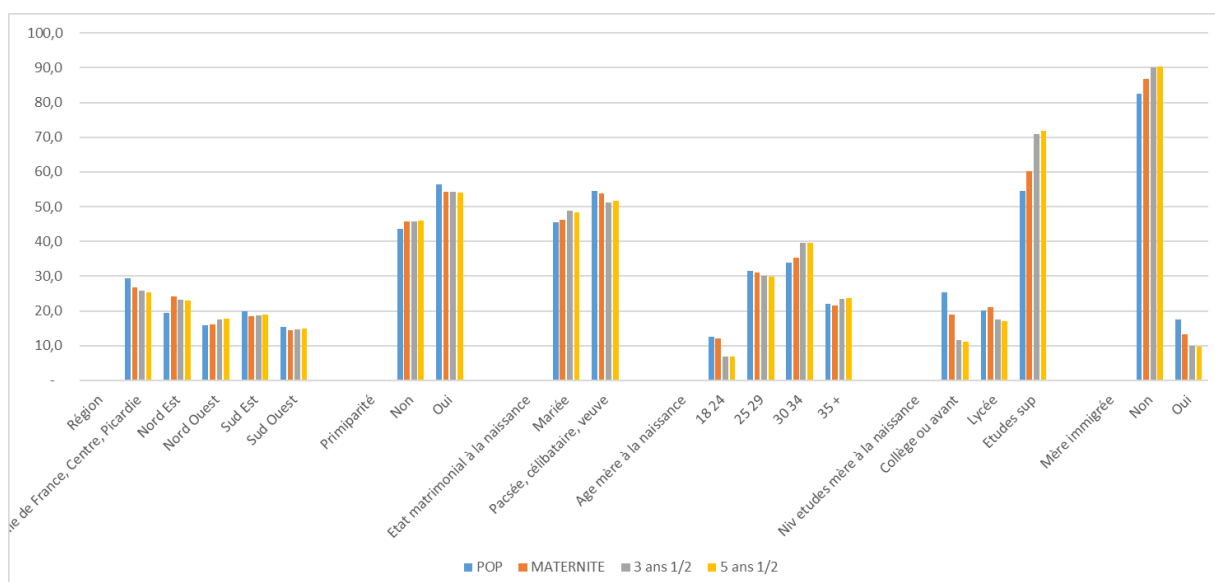
Qui sont les répondants à cette phase d'enquête

2 éléments peuvent être analysés : qui sont les répondants à cette enquête et comment les non-répondants évoluent au cours du temps. Pour ce faire, on peut comparer, pour l'ensemble des variables de calage, la répartition sur la population cible, sur les répondants en maternité, les répondants aux 3 ans ½ de l'enfant (afin d'estimer l'attrition avant cet âge) et les répondants aux 5 ans ½ de celui-ci (afin d'estimer l'attrition depuis la dernière enquête).

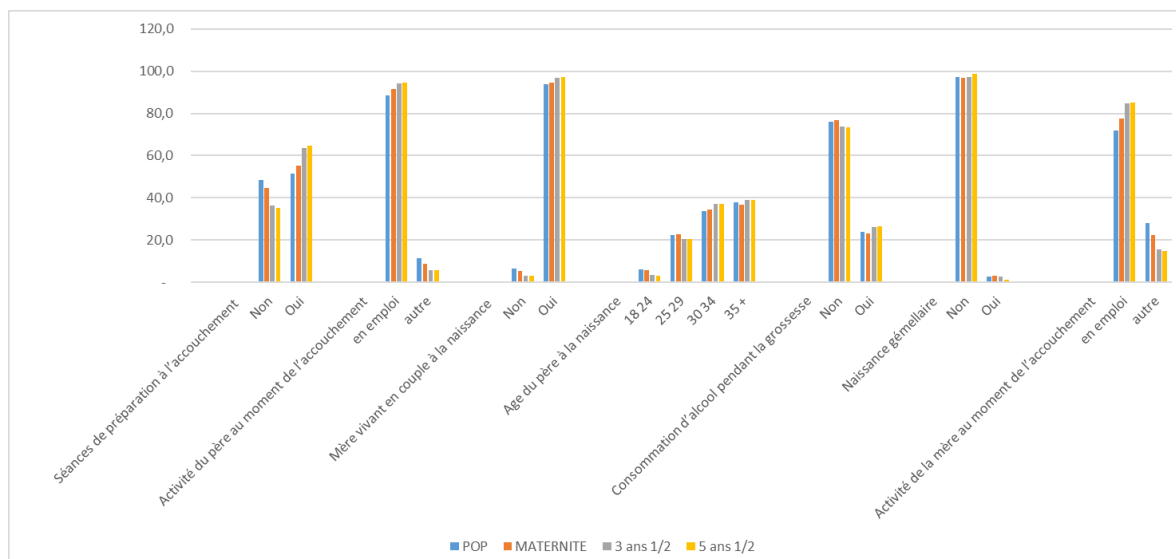
	POP	MATERNITE	3 ans 1/2	5 ans 1/2		POP	MATERNITE	3 ans 1/2	5 ans 1/2
Région					Séances de préparation à l'accouchement				
Ile de France, Centre, Picardie	29,3	26,8	25,9	25,4	Non	48,4	44,7	36,4	35,3
Nord Est	19,5	24,2	23,2	23,0	Oui	51,6	55,3	63,6	64,7
Nord Ouest	15,8	16,1	17,6	17,8	Activité du père au moment de l'accouchement				
Sud Est	20,0	18,5	18,6	18,9	en emploi	88,7	91,4	94,1	94,4
Sud Ouest	15,5	14,5	14,7	14,9	autre	11,3	8,6	5,9	5,6
Primiparité					Mère vivant en couple à la naissance				
Non	43,6	45,6	45,8	45,9	Non	6,3	5,5	3,0	3,0
Oui	56,4	54,4	54,2	54,1	Oui	93,7	94,5	97,0	97,0
Etat matrimonial à la naissance					Age du père à la naissance				
Mariée	45,5	46,2	48,8	48,4	18 24	6,1	5,8	3,3	3,2
Pacsée, célibataire, veuve	54,5	53,8	51,2	51,6	25 29	22,4	22,9	20,5	20,6
Age mère à la naissance					30 34	33,5	34,5	37,1	37,0
18 24	12,6	12,1	7,0	6,9	35 +	37,9	36,8	39,1	39,2
25 29	31,4	31,1	30,0	29,9	Consommation d'alcool pendant la grossesse				
30 34	33,9	35,3	39,6	39,7	Non	76,2	76,8	73,7	73,5
35 +	22,0	21,6	23,5	23,6	Oui	23,8	23,3	26,3	26,5
Niv études mère à la naissance					Naissance gémellaire				
Collège ou avant	25,2	18,8	11,5	11,1	Non	97,4	96,9	97,1	98,7
Lycée	20,1	21,0	17,5	17,1	Oui	2,6	3,1	2,9	1,4
Etudes sup	54,6	60,2	71,0	71,8	Activité de la mère au moment de l'accouchement				
Mère immigrée					en emploi	72,1	77,7	84,6	85,2
Non	82,6	86,7	90,1	90,4	autre	27,9	22,3	15,4	14,8
Oui	17,5	13,3	9,9	9,7					

On peut constater avec les graphiques suivants quelques phénomènes :

- Une certaine population est de toute façon non-répondante (lorsque la barre orange est inférieure à la barre bleue, cette catégorie est non-répondante dès l'enquête en maternité) – les jeunes parents, immigrés, peu diplômés et sans emploi.
- Puis une certaine population a abandonné l'enquête au cours du temps (lorsque la barre grise est inférieure à la barre orange) – globalement la même population que la précédente pour laquelle l'attrition s'est accélérée.
- Puis enfin une certaine population a quitté l'enquête lors de cette seule dernière phase aux 5 ans ½ (lorsque la barre jaune est inférieure à la barre grise) – c'est heureusement très peu le cas.

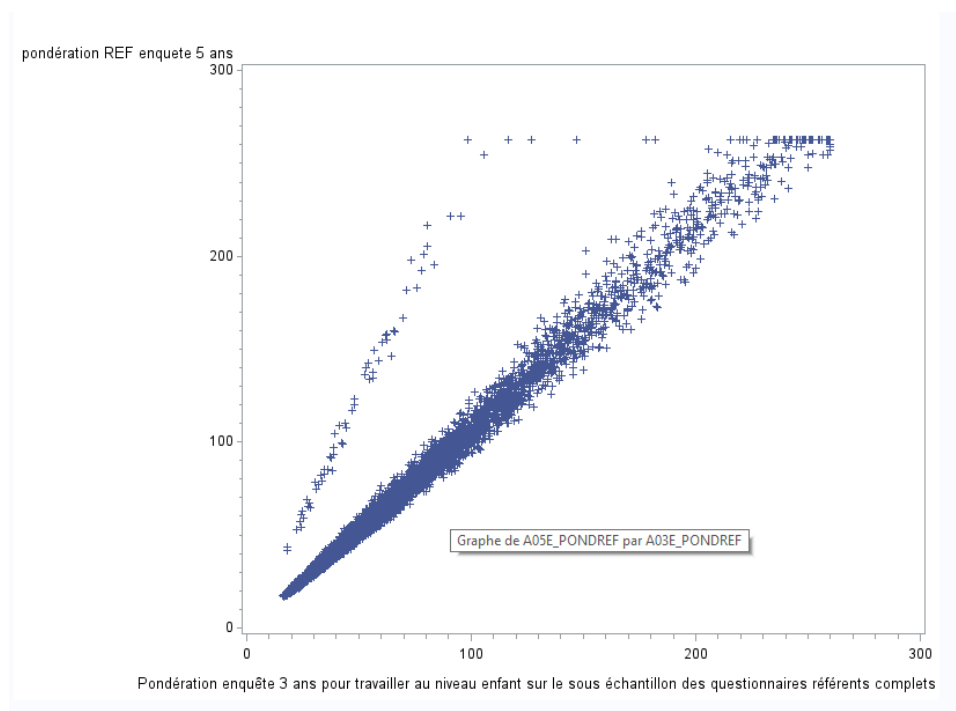


Distribution des variables de calage sur la population cible, répondants en maternité, répondants aux 3,5 ans, répondants aux 5,5 ans



Distribution des variables de calage sur la population cible, répondants en maternité, répondants aux 3,5 ans, répondants aux 5,5 ans

À noter également, lorsqu'on compare, pour les 10 441 enfants dont le parent référent a répondu aux enquêtes 5 ans ½ et 3 ans ½ la part de la population que chacun d'eux représente (le poids qu'on leur attribue), une corrélation très forte de ces 2 pondérations, ce qui permet d'assurer une certaine continuité des analyses (une variabilité d'un phénomène mesurée aux 3 ans ou au 5 ans des enfants résultera bien d'une variabilité du phénomène, et non d'une variabilité importante des pondérations).



Cependant, environ 2% des enfants (moins de 200 en représentent 26 000 de la population cible) subissent une évolution de leur poids supérieure à 20%. Ces 200 enfants, qui représentent ceux dont la probabilité d'être concernés par cette enquête diminue encore plus vite que le reste de la population, sont en réalité les jumeaux (3% de la population mais... 70% de ces individus). En effet, leur part diminue de 50% dans les répondants, ce qui n'était pas le cas jusque-là, et les obliger à représenter 3% de la population cible fait fortement augmenter leur poids.

3. Pondération longitudinale

La seconde pondération est une pondération **longitudinale** (A05E_LONGPONDREF). Elle concerne les enfants dont le parent référent a répondu à l'ensemble des questionnaires précédents **et** au questionnaire 5 ans ½, soit 9 798 enfants.

Les résultats de cette pondération sont les suivants (la limite a été fixée à 300 pour les poids maximum) :

Mesures statistiques de base				Quantiles (Définition 5)	
Location		Variabilité		Niveau	Quantile
Moyenne	77,9751	Ecart-type	68,29288	100Max 100%	322,93
Médiane	52,545	Variance	4664	99%	322,93
Mode	322,9253	Intervalle	306,55828	95%	246,34
		Ecart interquartile	53,91454	90%	168,74
				75% Q3	89,68
				50% Médiane	52,55
				25% Q1	35,76
				10%	27,57
				5%	24,09
				1%	19,88
				0% Min	16,37

