

Annexe IV

Analyse des résultats d'Inter Aide sur les populations

I. PROBLÉMATIQUE

Dans l'analyse des enquêtes Inter Aide, la première donnée à valider concerne la population recensée. Il existe en fait trois données sur la population: la population globale du village, celle des mères d'enfants de moins de 5 ans, et celle des enfants de moins de 5 ans. Dans ces données, recherchées chaque année, seule la dernière sert de base au calcul de TMM5, qui est basé sur l'extrapolation du nombre de décès annuels rapporté au nombre total d'enfants de moins de 5 ans vivant dans le village.

En théorie, le calcul du TMM5 doit se faire sur la population moyenne d'enfants de moins de 5 ans, calculé ainsi:

$$\frac{\text{nombre d'enfants} < 5 \text{ ans année N} + \text{nombre d'enfants} < 5 \text{ ans année N-1}}{2}$$

Lorsque l'on regarde les données d'une année sur l'autre, ces populations, qui devraient être proches, varient dans des proportions souvent étonnantes.

Cette constatation nous a amené à étudier les déterminants des variations afin d'évaluer les variations qui relèvent de variations normales de celles qui relèvent de biais secondaires à des enquêtes mal conduites.

II. DÉTERMINANTS DES VARIATIONS DE LA POPULATION DES ENFANTS DE MOINS DE 5 ANS.

Supposons que l'on appelle "cohorte" la population des enfants de moins de 5 ans.

Cette cohorte varie ainsi d'une année sur l'autre:

$$\text{Cohorte}_{\text{année N}} = (\text{Cohorte} + \text{Naissances} - \text{Décès} - \text{sortie de cohorte})_{\text{année N-1}}$$

La cohorte_{année N-1} est la population recensée l'année précédente

Les Naissances sont comptabilisées dans l'enquête de l'année N, de manière rétrospective

Les décès sont comptabilisées dans l'enquête de l'année N, de manière rétrospective

Les sorties de cohorte correspondent aux enfants qui avaient entre 48 et 59 mois en année N-1 et qui ont 5 ans révolus en année N

Cette dernière donnée n'est pas connue. Dans les zones où des enquêtes ont été réalisées tous les ans, depuis 5 ans ou plus, dans lesquelles toutes les naissances et tous les décès ont été recensés, il serait possible de le déterminer. Mais cette approche est lourde et n'est pas possible dans la plupart des cas.

III. DÉTERMINATION DE LA VALIDITÉ DE LA POPULATION DES ENFANTS DE MOINS DE 5 ANS.

Dès lors, l'approximation qui est faite consiste à estimer que, à l'échelle des communes ou des zones (qui comprennent 2 ou 3 communes), cette population est stable et que l'on peut calculer le taux de

mortalité annuel μ en prenant la population de l'année N comme population de référence.

C'est ainsi que le TMM5 est calculé par Inter Aide actuellement.

Nous pensons que cette approximation est acceptable si les enquêtes sont réellement exhaustives.

IV. CONTRÔLE DE L'EXHAUSTIVITÉ DES RECENSEMENTS EFFECTUÉS LORS DES ENQUÊTES.

Pour évaluer le caractère exhaustif des enquêtes, nous avons repris une soixantaine de cahiers d'enquêtes réalisées entre 2005 et 2007 et nous avons comparé comment évoluaient les populations d'enfants de moins de 5 ans, en notant les nombre de maisons et de familles visitées. Dans trois communes, nous avons en outre relevé le nombre de familles sans enfant, avec 1 enfant, avec 2, avec 3, 4 et 5 enfants de moins de 5 ans.

Les tableaux qui suivent présentent les 7 communes, prises au hasard dans lesquelles une analyse comparative des populations d'enfants de moins de 5 ans a été faite.

Initialement, nous n'avons étudié que les résultats discordants, c'est à dire lorsque la cohorte variait beaucoup d'une année à l'autre (rapport année N/année N-1 $> 125\%$ ou $< 75\%$).

Puis, constatant dans certains cas que des biais multiples se compensaient à l'échelle d'un village, nous avons vérifié l'intégralité des cahiers pour les dernières communes.

Nous ne commenterons dans ce document qu'une partie des tableaux car certaines observations se retrouvent d'une commune ou d'un village à l'autre.

V. Données brutes

		N enfants<5 ans		Rapport 07/05
		2005	2007	
Lokomby	Ambodimanga	239	244	102%
	Fandova	341	187	55%
	Lokomby	273	250	92%
	Mahavelo sirana	99	106	107%
	Mitanty est	239		
	Vohibazimba	239		
sous Total		1430	787	55%
Ambahive	AMBALATENINA	245	149	61%
	AMBATOMAINTY	190	216	114%
	IABAZAZA	91	102	112%
	MISIRINTONA	171	238	139%
	MITANTY OUEST	221	166	75%
	NANARENA	235	224	95%
	TAZOKONDRO BE	180	255	142%
	VOHIBOATAVO	248	402	162%
sous Total		1581	1387	88%
Total		3011	2174	72%

	N enfants<5 ans		
	2006	2007	2007/06
Vohimasy			
AMBODIVOANGY NORD	75	78	104%
AMBODIVOANGY SUD	320	267	83%
AMBOHIMANARIVO	157		0%
AMBOHIMIARY	72	80	111%
ANIVOTSARA	131	85	65%
ANKARIMBARY	145	129	89%
MAROFINARITRA	174	196	113%
VOHIMARY	126	145	115%
VOHIMASY NORD	91	71	78%
VOHIMASY SUD	219	200	91%
TOTAL	1 510	1251	83%
TOTAL CORRIGE	1353	1251	92%

	N enfants<5 ans		Rapport 07/06
	2006	2007	
Ambodiara	293	331	113%
Ambohitsara	166	158	95%
Marolanana	195	221	113%
Vinanitelo	366	350	96%
Sous-total	1020	1060	104%
Ambalafary	190	174	92%
Bekatra	290	293	101%
Marokoriaka	205	236	115%
Niahonana	110	94	85%
Sous-total	795	797	100%
Total	1815	1857	102%

	N enfants<5 ans		Rapport 07/06
	2006	2007	
Vatana			
Andranomavo	165	149	90%
Fotatra	67	93	139%
Nato	96	67	70%
Vatana be	231	209	90%
Vatana kely	75	109	145%
Vatana anamba	42	51	121%
Vohimary	70	114	163%
total	746	792	106%

	N enfants<5 ans		
	2006	2007	2007/06
Sahasinaka			
AMBODIBONARA	402	337	84%
AMBODILAZA	71	116	163%
AMPIRINGALAVA	111	114	103%
ANALABE	192	179	93%
MAROANDRONANA	129	150	116%
Shk AMBODIRAMIAVONA	379	418	110%
Shk VILLE BASSE	117	125	107%
TSIMANIMANINA	91	116	127%
VANGAINDRANO KELY	102	90	88%
VATOLAMBO	169	178	105%
Total	1 763	1823	103%

VI COMMENTAIRES

VI.1. Villages enquêtés en année N-1 et non enquêtés en année N

Cette situation est arrivée dans plusieurs communes puisqu'en 2007, environ 15% des enquêtes prévues n'ont pas été réalisées, essentiellement dans les premières zones.

Dans la commune de Lokomby, cette situation est survenue dans 2 villages sur 6. La cohorte 2007 correspond à 55% de la cohorte 2005.

		N enfants < 5 ans		
		2005	2007	Rapport 07/05
Lokomby	Ambodimanga	239	244	102%
	Fandova	341	187	55%
	Lokomby	273	250	92%
	Mahavelo sirana	99	106	107%
	Mitanty est	239		
	Vohibazimba	239		
sous Total		1430	787	55%

Si on apparie les villages, en excluant Mitanty Est et Vohibazimba, ce rapport remonte à 83%

		N enfants < 5 ans		
		2005	2007	Rapport 07/05
Lokomby	Ambodimanga	239	244	102%
	Fandova	341	187	55%
	Lokomby	273	250	92%
	Mahavelo sirana	99	106	107%
	Mitanty est			
	Vohibazimba			
sous Total		952	787	83%

ce qui reste bas, en raison d'un résultat discordant dans le village de Fandova (voir plus loin).

La conséquence sur le calcul du TMM5 n'est pas anodine.

En 2006, le TMM5 calculé = 239%

Si on retranche les deux villages non enquêtés en 2007, le TMM5 de 2006 devient = 263% soit une différence de 24%.

Dans certains cas, la différence peut être encore plus grande et varier dans le sens d'une amélioration du TMM5 de l'année N-1.

VI.2. Analyse des écarts supérieurs à 30%

VI.2.1. Lokomby

Une fois les villages appariés, on remarque que les résultats sont cohérents sauf pour le village de Fandova, dont la cohorte est divisée par deux entre 2005 et 2007.

En reprenant les cahiers d'enquête, on peut mettre en évidence plusieurs raisons. En 2007, un cahier

qui comportait 53 enfants de moins de 5 ans n'est pas apparu fiable à l'équipe terrain, qui ne l'a pas pris en compte. En ajoutant ce cahier, la cohorte 2007 correspond à 70% de celle de 2005.

Par ailleurs, en 2005, les enquêtes étaient réalisées par plusieurs enquêteurs pour un même village. A Fandova, 4 enquêteurs ont travaillé. L'un d'eux a probablement « gonflé » les chiffres puisqu'il a noté plusieurs familles de 5 enfants de moins de 5 ans. Enfin, puisque plusieurs enquêteurs ont travaillé sur le même village, on ne peut exclure des doublons. Ce biais potentiel a été prévenu par la suite, puisque désormais 1 seul enquêteur (très rarement 2) réalise toute l'enquête dans un village.

VI.2.2. Analyse des écarts supérieurs à 25%

Dans un premier temps, nous avons cherché à déterminer les causes des variations de cohorte lorsque celles-ci étaient manifestes. Dans la commune de Ambahive, nous avons repris les cahiers des villages présentant un rapport Cohorte 2007/cohorte 2006 $\leq 25\%$ ou $\geq 25\%$.

Les résultats sont présentés dans le tableau ci-dessous:

AMBAHIVE	2005				2007				07/05
	N E<5 ans	Erreur report	Familles enquêtées	E<5 ans / Famille	N E<5 ans	Erreur report	Familles enquêtées	E <5 ans / Famille	
Ambalatenina	245	0	138	1,78	149	0	146	1,02	61%
Misirintona	171	0	72	2,38	238	0	158	1,51	139%
Mitanty Ouest	222	1	179	1,24	166	0	141	1,18	75%
Tazokondro be	183	3	180	1,02	237	18	180	1,32	130%
Vohiboatavo	241	7	220	1,1	402	0	293	1,37	167%

La dernière colonne montre le rapport entre les cohortes de 2005 et celles de 2007. Dans tous ces villages, le nombre d'enfants paraît varier de façon trop importante.

Dans 2 cas, à la baisse.

Dans 3 cas, à la hausse.

A Mitanty Ouest, la baisse de la cohorte est proportionnelle à la diminution du nombre de familles visitées (179 en 2005, 141 en 2007). Si on effectue une péréquation, le rapport entre la cohorte 07/cohorte 05 est de 95%.

A Ambalatenina, le nombre de familles visitées est sensiblement le même. Par contre, dans les cahiers en 2005, on remarque que 100% des familles ont au moins 1 enfant de moins de 5 ans et qu'il existe plusieurs familles notées comme ayant 5 enfants de moins de 5 ans. Il semble donc l'enquêteur de 2005 ait volontairement ou pas *gonflé* les chiffres de la cohorte 2005. Les familles en 2007 ont en général entre 0 et 2 enfants dans la cohorte.

A Tazonkondro, il existait une erreur de report: le résultat noté était de 255 enfants. Nous n'en avons compté que 237. Malgré cela, l'augmentation reste importante, alors que le nombre de familles visitées est homogène. L'analyse des cahiers ne donne pas grand chose. Il n'est pas possible de croiser avec les données des naissances et des décès car il n'y a pas eu d'enquête dans cette zone en 2006. Concernant l'erreur de report, nous n'avons rencontré qu'exceptionnellement des erreurs de cette importance ($> 5\%$), la plupart étant inférieures à 3%. De plus, comme elles se compensent en général, nous avons trouvé, à l'échelle des communes, des variations inférieures à 1%.

Le village de Misirintona est intéressant car il cumule deux mécanismes qui se compensent partiellement. Entre 2005 et 2007, le nombre de familles visitées est multiplié par plus de deux, mais le nombre d'enfants de moins de 5 ans par famille chute de 2,38 (résultat très inhabituel) à 1,51 (résultat plus conforme aux standards). Ici, il semble que les données 2007 soient plus fiables que

les données 2005.

A Vohoboato, 2 mécanismes sont observés, qui s'additionnent. Le nombre de familles visitées progresse de 33% et le nombre d'enfants de moins de 5 ans par famille progresse lui aussi. Ici aussi, l'absence de données en 2006 empêche de poursuivre l'analyse.

VI.2.3. Analyse des écarts inférieurs à 15%

Nous avons repris les cahiers des zones III et VII, qui permettaient de comparer les données de 2006 et de 2007.

Pour les écarts importants, les mécanismes trouvés sont identiques à ceux décrits au paragraphe précédent. Même avec le nombre de naissances et de décès, les variations des cohortes sont parfois totalement inexplicables.

Nous nous sommes aussi intéressés aux résultats qui paraissent, en première lecture, fiables. En général, lorsque le rapport entre la cohorte 07/cohorte 06 est $> 85\%$ et $< 115\%$, les données ne présentent pas d'incohérence évidente (ce qui est logique).

VOHIMASY	2006				2007				07/06
	N E<5 ans	Erreur report	Familles enquêtées	E<5 ans / Famille	N E<5 ans	Erreur report	Familles enquêtées / Famille	E <5 ans / Famille	
Marofinaritra	174	1	143	1,22	196	0	149	1,32	1,13
	175	0	143	1,22	196	0	149	1,32	1,12

Cet exemple montre une augmentation de 12% (une fois corrigée une petite erreur de report) entre la cohorte 2006 et 2007. Le nombre de familles enquêtées est homogène. Le rapport enfants < 5 ans/ nombre de familles augmente mais il y a eu dans l'année 52 naissances et 9 décès, ce qui peut expliquer les augmentations constatées.

Dans certains cas cependant, la cohérence n'est qu'apparente. L'exemple suivant a aussi été observé dans la commune de Vohimasy.

VOHIMASY	2006				2007				07/06
	N E<5 ans	Erreur report	Familles enquêtées	E<5 ans / Famille	N E<5 ans	Erreur report	Familles enquêtées / Famille	E <5 ans / Famille	
Ankarimbary	145	15	129	1,12	129	3	99	1,3	0,89
Ankarimbary corr	160	0	129	1,24	126	0	99	1,27	0,79

Dans ce village, la différence observée n'est que de 11% à la baisse entre 2006 et 2007. En reprenant les données, nous avons trouvé une erreur de report importante ($\approx 10\%$) en 2006 et une autre de 3% en 2007. La baisse réelle était donc de plus de 20%. Une fois corrigés, les résultats paraissent d'ailleurs plus homogènes, mis à part que le nombre de familles enquêtées a diminué de plus de 23%. Dans cet exemple, on voit que si seule l'évolution de la cohorte est suivie, on peut considérer à tort que le résultat 2007 est cohérent avec le résultat 2006. Le suivi du nombre de familles enquêtées reste important.

VI.2.4. Analyse complémentaire: la répartition des familles selon le nombre d'enfants

Puisque lorsqu'il y a des erreurs de report, des résultats d'apparence fiables peuvent masquer des enquêtes incomplètes, nous avons cherché une méthode pratique pour vérifier que les données des cahiers avaient été reportées sans erreur dans les feuilles de calcul informatisées.

Si nous avons bien compris la méthode utilisée par l'équipe terrain pour synthétiser les données,

celle-ci nécessite une concentration importante.

Les cahiers actuels se présentent ainsi:

	famille 1	famille 2	famille 3	famille 4	famille 5
Nom père	X1	X2	X3	X4	X5
Nom Mère	Y1	Y2	Y3	Y4	Y5
N personnes dans la famille	5	4	7	8	3
N naissances année précédente	0	0	0	1	1
N enfants <5 ans	1	0	2	3	1
etc...					

Dans un premier temps, les résultats sont additionnés par ligne, puis comptés pour l'ensemble du cahier, en tournant les pages.

Cette méthode est rapide. Elle est simple lorsque les cahiers sont courts (moins de 100 familles), plus pénible lorsque les cahiers contiennent aux alentours de 200 familles.

Lors des vérifications, nous avons pu constater combien cette méthode exposait au risque d'erreur même des personnes expérimentées.

Pour assurer nos propres vérifications, nous avons pris le nombre d'enfants par famille.

	Composition famille																	
	2006								2007									
								Total									Total	
	0	1	2	3	4	5	Familles	E<5A	0	1	2	3	4	5	E<5A	Familles		
Ambodiara	0	109	72	19			200	310	5	85	90	17	2	1	329	200		
Ambohitsara	40	68	36	8			152	164	1	76	40			1	161	118		
Marolanana	49	81	52	1			183	188	0	66	65	7	1		221	139		
Vinanitelo	0	87	104	19	1		211	356	3	100	87	22	3		352	215		
Ambalafary	16	52	52	10	1		131	190	26	58	53	2	1		174	140		
Bekatra	31	101	68	15	2		217	290	22	118	70	12			294	222		
Marokoriaka	2	81	55	5			143	206	55	68	66	9	1	1	236	200		
Niahonana	20	33	34	3			90	110	18	40	27				94	85		
totaux	158	612	473	80	4	0	1327	1814	130	611	498	69	8	3	1861	1319		

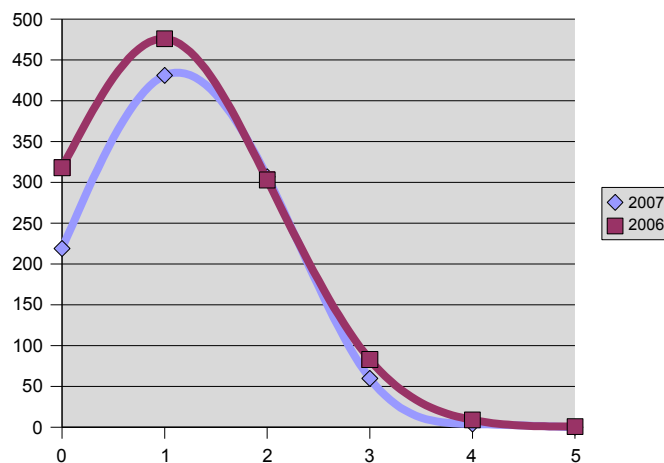
Cette méthode permet de croiser facilement les résultats du nombre de familles enquêtées (directement disponible dans les cahiers) et le nombre d'enfants de moins de 5 ans trouvé par les additions.

Ces tableaux permettent en outre de comparer la répartition des familles de moins de 5 enfants qui, sur des grands effectifs, sont stables.

Nous avons réalisé ce travail pour trois communes. Les résultats correspondant aux tableaux ci-dessus ont été présentés dans le rapport principal.

La dernière commune observée est celle de Vohimasy. Les résultats sont un peu moins superposables que pour Vinanitelo et Bekatra.

N enfant<5 par famille - Vohimasy



Comme on peut le voir, le nombre de famille sans enfant ou avec un seul enfant est plus bas en 2007 qu'en 2006. Les effectifs (environ 1000) sont plus faibles que dans l'exemple précédent (environ 1800) et le couplage avec les données d'Ambandrika, l'autre commune de cette zone compenserait peut être cette différence. Mais cette différence peut aussi être liée aux consignes qui sont données aux enquêteurs.

Nous profitons donc de cet exemple pour attirer à nouveau l'attention sur une consigne qui est donnée aux enquêteurs et qui nous paraît plus nocive qu'utile. Lors de l'enquête initiale, un recensement exhaustif doit être fait. Lors des enquêtes de suivi il est demandé aux enquêteurs de passer dans toutes les maisons pour vérifier s'il y a un/des enfants de moins de 5 ans. Puis d'enregistrer dans les cahiers uniquement les données des familles ayant des enfants de moins de 5 ans.

Comme nous le voyons, en 2006, plus de 300 familles ont été notées dans l'enquête initiale. En 2007, 219 familles déclarent ne pas avoir d'enfants de moins de 5 ans. En raison de la consigne passée aux enquêteurs, il est difficile de savoir si la consigne a été partiellement suivie ou si cette diminution reflète autre chose.

Nous pensons que la réalisation systématique du recensement de tout le village permettra de sécuriser les données dans la mesure où tous les résultats pourront être comparés à la même population de référence.

Les indicateurs pour lesquels il existe des normes nationales (nombre de grossesses attendues, % d'enfants de moins de 5 ans dans la population) pourront être suivis d'année en année.

SYNTHÈSE

La population d'enfants de moins de 5 ans, telle qu'elle est recensée par Inter Aide dans ses enquêtes, est probablement représentative de la population réelle de cette tranche d'âge au-delà d'un effectif de 1800 à 2000 enfants. L'augmentation des effectifs permet de compenser des variations liées au principaux biais mis en évidence.

Néanmoins ces biais sont en partie liés à des erreurs de recueil de données qu'il est souvent possible d'améliorer. Nous avons décrit dans le document central des propositions concernant la participation communautaire (cartographie, validation des données) et le suivi des indicateurs démographiques.

L'harmonisation du recensement d'une année sur l'autre et la détermination précise, dès l'enquête initiale, du nombre de maisons et/ou de familles à visiter dans un village nous paraissent être 2 préalables indispensables.

Nous voudrions, dans ce rapport plus technique, faire une proposition de modification du cahier de recueil des données, dont vous jugerez le rapport contraintes/bénéfices.

En ligne 5, il est demandé aux parents combien ils ont d'enfants de moins de 5 ans. Peut être pourrait-on ajouter à côté l'âge de ces enfants:

	famille 1		famille 2		famille 3	
Nom père	X1		X2		X3	
Nom Mère	Y1		Y2		Y3	
N personnes dans la famille	5		4		7	
N naissances année précédente	0		0		0	
N enfants <5 ans	1		0		2	
etc...						

ou une autre formule permettant un contrôle d'une année sur l'autre.

Tous ces efforts permettront probablement d'affiner et de fiabiliser les résultats des enquêtes annuelles concernant les populations (générale et enfants de moins de 5 ans).

La précision des mesures en sera renforcée, en particulier le TMM5 dont les variations liées aux lois statistiques sont importantes sur les petits effectifs, et pour lequel tout biais évitable doit être supprimé.

Si ce n'est pas possible et que des résultats trop discordants sur la population conduisent à supprimer les données d'un village, il nous semble que les TMM5 antérieurs devraient être recalculés.

En effet, lors des comparaisons successives des TMM5, les villages pris en compte dans les enquêtes et les calculs devraient être appariés, en retirant de la comparaison les villages non vérifiés à l'année N. Si le TMM5 de l'année N-1 est modifié (ce qui est presque mathématique), c'est ce nouveau résultat TMM5 qui devra servir de référence dans la comparaison avec le TMM5 de l'année N, et non le résultat publié l'année précédente.