

## NOUVELLES RECHERCHES SUR LA DISTRIBUTION DU SEXE A LA NAISSANCE

par Marcel Paul SCHUTZENBERGER,  
*Attaché de recherche à l'Institut National d'Hygiène*

DANS un article publié l'année dernière sous le titre de « Résultats d'une enquête sur la distribution du sexe dans les familles nombreuses », nous exposons les conclusions auxquelles nous avait conduit l'étude statistique de 14.230 familles françaises contemporaines de 5 enfants et plus, poursuivie au centre de génétique médicale de l'Hôpital Saint-Louis sous la direction de notre maître M. le Professeur R. Turpin et grâce à l'appui du fonds d'études de la Société Médicale des Hôpitaux de Paris. Par la considération des séquences d'enfants de même sexe nous montrions qu'il existe une liaison entre le sexe des enfants consécutifs dans une famille. Cette liaison, cette solidarité entre sexes successifs, nous pouvions démontrer qu'elle était indépendante dans une très large mesure de facteurs héréditaires éventuels favorisant dans un couple de progéniteurs telle ou telle déviation de la *sex ratio* par rapport à sa valeur moyenne dans l'ensemble de la population.

En même temps que nous mettions en valeur ce phénomène, nous établissions que dans les limites de précision que permettait notre échantillon aucune variation systématique de la fréquence des naissances masculines ne se manifestait en fonction du rang de naissance ou de l'importance de la famille.

Ces résultats avaient amené notre maître M. le Professeur Turpin et nous-même à conclure à l'existence de périodes gynophiles et androphiles, s'étendant sur l'inter-

valle de plusieurs gestations et favorisant par un mécanisme encore à préciser la naissance de filles ou de garçons.

Pour pousser plus loin cette étude attachante, nous indiquions qu'il était nécessaire non pas tant d'augmenter le volume de l'échantillon examiné, chose assez malaisée d'ailleurs, que surtout d'obtenir une information plus riche sur ces naissances. C'est ce travail que M. le Professeur Turpin a confié à deux de ses élèves : Mme Denise Derroche et Mme Jacqueline Leviné, dont il a bien voulu nous charger d'aider la préparation des thèses de doctorat en médecine et dont nous exposerons ici les conclusions statistiques les plus frappantes. Conclusions parfois provisoires d'ailleurs car, comme nous l'expliquerons plus loin, divers problèmes se sont posés en cours de dépouillement et ont nécessité la collecte d'un matériel plus abondant qui se poursuit encore actuellement.

On ne sera donc pas surpris que ce compte rendu soit volontairement limité à l'exposé de nos seuls résultats et que nous ayons réservé pour une publication ultérieure toute la discussion des nombreux travaux français et étrangers consacrés aux problèmes de la détermination du sexe chez l'homme.

### Les documents.

De même que dans notre précédente enquête les matériaux proviennent du service de la « Médaille d'honneur

aux mères » dépendant du Ministère de la Santé Publique et de la Population. Il s'agit donc encore de familles nombreuses françaises contemporaines représentant un échantillon réparti assez régulièrement sur l'ensemble du territoire français. Nous ne reprendrons pas l'exposé des raisons qui nous incitent à une très grande confiance dans l'exactitude et dans le caractère représentatif de cet échantillon et nous nous bornerons à rappeler toutefois que des causes d'ordre administratif (médailles d'or, d'argent, de bronze, attribuées en fonction du nombre des enfants) privent de signification démographique précise les fréquences relatives des familles de  $n$ ,  $n + 1$ ,  $n + 2$ ... enfants pour  $n$  plus petit que 10.

Nous signalerons en outre que pour une fraction difficile à préciser il s'agit de familles « en cours » dont l'accroissement n'est peut-être pas encore terminée et qu'ainsi « une famille de 8 enfants » par exemple, signifie une famille dans laquelle 8 naissances avaient déjà eu lieu en 1948, année de la constitution des dossiers que nous avons utilisés. La distinction entre famille de 7, 8... enfants est donc un peu formelle quoiqu'à mesure qu'augmente la taille de la famille elle correspond de plus en plus à une différence réelle.

Deux échantillons ont été collectés initialement l'un de 502 familles de 5 enfants et plus comprenant au moins une paire gemellaire, l'autre de 826 familles de 7 enfants et plus, mais sans naissance multiple. L'information recueillie est constituée par le sexe des enfants successifs et l'âge de la mère à chaque naissance.

A la suite de ce premier travail que nous discuterons ici, il est apparu nécessaire d'augmenter le volume de nos échantillons, afin de porter à environ un millier le nombre des familles à naissances gémellaires et à 1.500 le nombre des familles de plus de 7 enfants.

Cette extension de l'enquête étant actuellement en cours, avec la participation de Mme Pouzzole nous nous baserons le plus fréquemment sur les chiffres obtenus à partir des seuls échantillons initiaux.

## II. — LES INTERVALLES ENTRE NAISSANCES SUCCESSIVES DANS LES FAMILLES DE PLUS DE 7 ENFANTS.

1° Il semble bien qu'à ce niveau, des facteurs psychologiques si souvent évoqués, tels par exemple que le désir des parents d'avoir des garçons (ou des filles) ne jouent qu'un rôle des plus minimes et que l'on se trouve en présence de phénomènes biologiques relativement peu perturbés. C'est cette considération qui nous a conduit à n'utiliser ici que ces familles malgré leur rareté relative. Le total de nos 826 familles se décompose comme suit :

7 enfants	: 180 familles
8 »	: 127 »
9 »	: 166 »
10 »	: 159 »
11 »	: 117 »
12 »	: 46 »
13 »	: 31 »

soit au total 7.602 naissances dont 3.875 garçons. La proportion des naissances masculines est donc de  $0,5097 \pm 0,0057$  ; cette valeur ne diffère pas significativement

de celle que nous avons observée dans notre précédente enquête ( $0,5066 \pm 0,0016$  sur 95.725 naissances).

Bien qu'évidemment les chiffres de notre échantillon soient quelque 5 fois trop faibles pour que le phénomène apparaisse comme statistiquement prouvé sur ce seul matériel nous retrouvons cependant ici encore une liaison entre le sexe des enfants consécutifs, qui confirme nos observations antérieures.

Appelons (gg) les intervalles entre deux garçons consécutifs.  
» (gf) » » un garçon et une fille.  
» (fg) » » une fille et un garçon.  
» (ff) » » deux filles.

Nos 6.776 intervalles se répartissent ainsi :

Intervalle	Nombre observé effectivement	Nombre attendu dans l'hypothèse où les naissances seraient indépendantes.	Différence
(gg)	1.778	1.759,7	+ 18,3
(gf) ou (fg)	3.362	3.386,7	- 24,7
(ff)	1.636	1.629,6	+ 6,4
Total . . . . .	6.776	6.776,0	0,0

ce qui montre bien l'excès des paires d'enfants consécutifs de même sexe par rapport aux valeurs qu'aurait prévues une théorie ne tenant pas compte de la liaison, même si un test par chi-carré ne parvient pas à établir le degré de signification statistique de ces écarts.

Nous avons calculé l'intervalle séparant chaque naissance de la naissance qui la suit pour toutes nos familles. Voici par exemple, pour un échantillon de 337 familles de 7 enfants, et un autre de 258 familles de 8 enfants la répartition de la durée des intervalles, l'unité employée

8, 9, 10 mois	11, 13 mois	12, 16 mois	14, 18 mois	15, 19 mois	17, 22 mois	18, 25 mois	20, 28 mois	21, 25 mois	23, 28 mois	24, 26 mois	26, 27 mois
14 familles de 7 enfants . . .	157	263	236	226	201	167					
5 familles de 8 enfants . . .	137	263	257	234	197	159					
29, 30, 31 mois	32, 34 mois	33, 37 mois	35, 40 mois	36, 43 mois	38, 46 mois	39, 43 mois	41, 46 mois	42, 46 mois	44, 46 mois	45, 49 mois	47, 48 mois
141 familles de 7 enfants . . .	118	89	76	47	45	39					
96 familles de 8 enfants . . .	74	56	64	53	29	33					
50, 51, 52 mois	53, 55 mois	54, 58 mois	56, 61 mois	57, 61 mois	59, 61 mois	60, 61 mois	62 et plus				
26 familles de 7 enfants . . .	22	21	23	111							
15 familles de 8 enfants . . .	27	19	11	77							

TABLEAU I

	Familles de 7 enfants	Familles de 8 enfants	Familles de 9 enfants	Familles de 10 enfants	Familles de 11 enfants
1 <sup>o</sup> intervalle ....	23,67 m. $\pm$ 0,74	20,56 m. $\pm$ 0,82	22,03 m. $\pm$ 0,95	20,28 m. $\pm$ 1,17	20,41 m. $\pm$ 0,85
2 <sup>o</sup> » .....	23,88 »	23,40 »	22,17 »	21,15 »	19,62 »
3 <sup>o</sup> » .....	26,47 »	25,98 »	25,30 »	21,28 »	19,56 »
4 <sup>o</sup> » .....	28,61 »	25,41 »	24,04 »	19,92 »	20,69 »
5 <sup>o</sup> » .....	30,83 »	27,40 »	26,06 »	22,32 »	21,43 »
6 <sup>o</sup> » .....	34,28 »	30,18 »	26,86 »	23,00 »	20,28 »
7 <sup>o</sup> » .....		33,10 »	27,87 »	23,49 »	21,67 »
8 <sup>o</sup> » .....			32,78 »	24,49 »	23,72 »
9 <sup>o</sup> » .....				27,40 »	26,23 »
10 <sup>o</sup> » .....					27,00 »

étant le trimestre (les calculs initiaux effectués en mois ont été regroupés).

Nous avons aussi regroupé en une seule classe les intervalles dépassant 62 mois (5 ans) : il semble difficile d'admettre qu'il s'agisse pour ces longs arrêts de la procréation d'un phénomène biologique intéressant de notre point de vue et l'introduction dans les calculs de valeurs parfois 5 ou 6 fois plus grandes que la moyenne (certains intervalles atteignent 12 ans) aurait risqué de fausser complètement le sens de ceux-ci. D'autre part l'utilisation d'autres caractéristiques de ces distributions, moins sensibles aux valeurs extrêmes, telles que le médian par exemple (durée de l'intervalle, telle que 50 % de ceux-ci soient plus courts qu'elle-même) nous aurait contraint de faire appel à des échantillons beaucoup plus nombreux. Par exemple : les valeurs médianes pour les familles de 7 et 8 enfants sont respectivement de 25,71 et 24,20 ; cette différence ne serait pas significative et pourtant la différence des durées moyennes (respectivement 27,96 et 26,56 mois) est indiscutable. Il en serait encore de même si l'on éliminait purement et simplement les intervalles supérieurs à 62 mois (valeurs moyennes : 25,92 et 24,93) ce qui justifie le groupement utilisé. Enfin comme il s'agira presque toujours ici de calculs portant sur les moyennes d'un grand nombre de valeurs individuelles, il a semblé préférable de ne pas effectuer de transformation de l'échelle des temps (telle que logarithmique ou par l'opération « racine carrée de ») faisant disparaître l'asymétrie marquée de la distribution mais faisant perdre en même temps tout caractère concret aux

opérations numériques ultérieures. D'ailleurs, et c'est la base d'une bonne partie de la statistique moderne, l'asymétrie et même la bimodalité disparaissent progressivement sous certaines conditions (ici remplies) avec le nombre des observations quand on effectue la moyenne de celle-ci.

2<sup>o</sup> Variations de la valeur moyenne des intervalles avec le nombre d'enfants dans la famille.

Nous avons calculé pour les familles de 7, 8, 9... 13 enfants la valeur moyenne de l'intervalle entre deux naissances consécutives : comme il était bien prévisible *a priori* ceux-ci décroissent assez régulièrement avec la taille de la famille :

Nombre d'enfants dans la famille	Nombre d'intervalles utilisés	Valeurs moyennes en mois
7	2022	27,96 $\pm$ 0,30
8	1806	26,56 $\pm$ 0,31
9	1328	25,91 $\pm$ 0,32
10	1431	22,59 $\pm$ 0,39
11	1170	22,06 $\pm$ 0,27
12	506	21,31 $\pm$ 0,41
13	372	19,08 $\pm$ 0,40

Remarquons que notre groupement de tous les intervalles supérieurs à 62 mois a pour résultat une réduction apparente probablement assez sensible de ces différences : dans les familles de 12 enfants par exemple, les intervalles supérieurs à 62 mois ne représentent que 1 % de l'ensemble des intervalles (il n'en existe pas pour les

TABLEAU II

7 enfants	1 <sup>er</sup> intervalle	2 <sup>e</sup> intervalle	3 <sup>e</sup> intervalle	4 <sup>e</sup> intervalle	5 <sup>e</sup> intervalle	6 <sup>e</sup> intervalle
Valeur calculée..	22,53 mois	24,70 mois	26,87 mois	29,04 mois	31,21 mois	33,38 mois
Valeur observée..	20,56 »	23,88 »	26,47 »	28,61 »	30,83 »	34,28 »

  

8 enfants	1 <sup>er</sup> interv.	2 <sup>e</sup> interv.	3 <sup>e</sup> interv.	4 <sup>e</sup> interv.	5 <sup>e</sup> interv.	6 <sup>e</sup> interv.	7 <sup>e</sup> interv.
Valeur observée..	19,94 mois	22,43 mois	24,92 mois	27,40 mois	29,84 mois	32,38 mois	34,87 mois
Valeur calculée..	19,41 »	23,19 »	26,64 »	25,88 »	28,53 »	32,23 »	36,14 »

familles de 13 enfants de cet échantillon) alors que leur fréquence atteint 5 % pour les familles de 7 enfants. Il est inutile d'insister sur la signification (au sens statistique) de ces différences bien apparentes d'après les valeurs de la déviation standard. Notons enfin l'intérêt de ce calcul qui légitime, en partie tout au moins, les divisions un peu formelles que nous avons adoptées entre familles de 7, 8, 9... enfants.

3° Variation avec le rang de naissance de la valeur moyenne des intervalles :

Ici aussi, ce facteur pouvait être prévu à l'avance : l'intervalle moyen séparant deux naissances est d'autant plus long que les rangs de celles-ci sont plus élevés, ou, ce qui revient à peu près au même statistiquement dans l'ensemble, que les progéniteurs sont plus âgés puisqu'il existe une corrélation élevée entre âge de la mère et rang de naissance et ceci d'autant plus que le nombre élevé des enfants limite l'échantillon aux femmes présentant une vie obstétricale plus régulière : voici les résultats de ces calculs dans tous les cas où le nombre des familles est suffisant pour obtenir une précision raisonnable : (voir Tableau I).

La comparaison de ces chiffres à leur déviation standard montre que les quelques irrégularités qui se présentent peuvent être légitimement attribuées à des fluctuations d'échantillonnage. Plus rigoureusement une analyse de variance de type classique ne révèle sur cet échantillon aucune discordance systématique réellement importante de ces chiffres d'avec un schéma de régression linéaire en fonction du rang de l'intervalle. Par exemple, pour 7 enfants et 8 enfants, respectivement on obtient les droites de régression :

$$Y_n = 20,36 + 2,17 n \text{ et } Y_n = 17,45 + 2,49 n$$

où  $Y_n$  est la valeur moyenne du nième intervalle, les équations conduisent aux valeurs suivantes : (voir Tableau II).

L'accord est pratiquement assez satisfaisant et le calcul donne (l'unité employée est le trimestre) :

	Somme des carrés		Nombre de degrés de liberté		Carré moyen	
	7 enfants	8 enfants	7 enfants	8 enfants	7 enfants	8 enfants
Total .....	43.309	34.592	2.021	1.085		
Déviati on intra-cl. ....	40.028	33.645	2.016	1.799	19,9	18,7
Déviati on inter-classe .....	3.281	2.947	5	6	656,2	491,2
Terme linéaire ..	3.095	2.797	1	1		2.797
» quadratique ..	97	12	1	1	97	12
» cubique ..	67	86	1	1	67	86
Résidu .....	22	42	2	3	11,0	14,0

Toutefois à l'examen des chiffres on est surpris de l'excès général pour presque tous les types de familles de l'accroissement du dernier intervalle par rapport aux autres. Ceci peut naturellement s'expliquer par une régression non linéaire, mais l'on pourrait aussi peut-être

penser à une période d'infécondité relative précédant le dernier né. Nos documents sont encore trop incomplets pour que nous puissions voir là autre chose qu'une indication pour une étude plus poussée.

4° Variation de la durée moyenne des intervalles avec le sexe des enfants :

Nous avons regroupé dans notre échantillon de 876 familles les intervalles en intervalles (gg), (gf), (fg), ou (ff) d'après le sexe des enfants qui les limitent et au total nous avons obtenu les valeurs moyennes suivantes :

Intervalles (gg) .....	24,38	mois
» (gf) .....	25,03	»
» (fg) .....	24,27	»
» (ff) .....	23,56	»

Pour mieux apprécier ces chiffres il semble préférable de faire les 3 comparaisons suivantes :

a) intervalles commençant par un garçon/intervalles commençant par une fille [(soit : (gg) + (gf) — (fg) — (ff)] cette comparaison montre un excès de 0,776 mois en faveur des intervalles suivant une naissance masculine. On peut ici évoquer, quoique nous le croyons pas en raison de la régularité de phénomène, l'influence de causes d'ordre psychosociologique (désir d'un héritier mâle ?).

b) intervalles se terminant par un garçon/intervalles se terminant par une fille [(soit : (gg) + (fg) — (gf) — (ff)] la différence est infime : 0,023 mois.

c) intervalles séparant 2 enfants de même sexe/intervalles séparant deux enfants de sexe opposé [(soit (gg) + (ff) — (fg) — (gf)] les intervalles entre enfants de même sexe sont plus courts que les autres de 0,658 mois, soit environ de 3 % de la durée moyenne.

Pour établir la réalité de ces différences, pour prouver qu'elles ne résultent pas seulement d'une fluctuation d'échantillonnage, nous avons eu recours à l'analyse de variance malgré que les 3 conditions classiques (normalité, indépendance, orthogonalité) ne soient pas rigoureusement remplies. Nous nous sommes déjà expliqué sur la normalité des distributions. En ce qui concerne le deuxième point (non indépendance des fluctuations aléatoires des intervalles successifs dans une même famille) son influence, difficile à estimer correctement, paraît toutefois s'exercer dans un sens qui rend le test plus rigoureux en réalité qu'il ne paraît. Enfin, si les 4 types d'intervalles ne sont pas exactement représentés par le même nombre de cas, il semble cependant que les écarts soient faibles par rapport aux valeurs idéales. Nous indiquerons d'ailleurs plus loin un autre test moins puissant mais indiscutable de ce résultat essentiel : l'allongement des intervalles entre enfants de sexes opposés.

Voici les résultats de cette analyse. L'unité ici est le trimestre et l'échantillon utilisé est l'échantillon initial de 826 familles (voir tableau ci-dessous).

Quelques explications sont nécessaires. La mention « ligne » se rapporte aux valeurs moyennes des intervalles de nième rang dans les familles de  $m$  enfants, et l'on a donné dans le bas du tableau la décomposition de cette partie de la variance ; la mention « colonne » se

Source de variation	Somme des carrés	Carrés moyens	Nombre de degrés de liberté
Total	103.397	15,26	6.775
Lignes	12.611		62
Colonnes	199	66,33	3
Interactions	9.595	51,59	186
Résidu	80.992	12,41	6.524
gg + gf — fg — ff	11,3	11,3	1
gg + fg — gf — ff	0	0	1
gg + ff — gf — fg	8,2	8,2	1
Familles	5.583	930,50	6
Lignes-familles	7.028	125,50	56

rapporte aux valeurs moyennes des quatre types d'intervalles : (gg), (gf), (fg) et (ff). L'interaction est relative aux différences possibles entre les effets des types d'intervalles aux différents rangs et dans les familles de dimension diverse.

La comparaison de la variance résiduelle aux autres variances nous permet alors d'aboutir aux conclusions suivantes :

1° l'accroissement de la durée moyenne des intervalles commençant par un garçon ne peut être due au hasard (F. de Snedecor, 9,03) probabilité inférieure à 1 % ;

2° l'accroissement de la durée moyenne des intervalles entre enfants de sexe opposé ne peut être due au hasard (F. de Snedecor : 6,03 ; probabilité : 2 %) ;

3° il existe vraisemblablement une interaction entre ces causes de variation et le rang et la taille des familles (F. de Snedecor : 4,16 ; probabilité : 1 0/00). Observons cependant que cette conclusion est sans doute assez discutable car c'est pour cette rubrique que jouent de la manière la plus grave les défauts d'orthogonalité, c'est-à-dire de non proportionnalité parfaite du nombre d'intervalles de chaque type à chaque rang de naissance ;

4° on retrouve évidemment ce résultat déjà mentionné que les intervalles varient significativement avec leur rang et l'importance des familles.

Nous évoquons plus haut les critiques qui pouvaient être adressées à notre application de l'analyse de variance, aussi il est peut-être plus convainquant de traiter directement le problème qui nous intéresse par une méthode plus grossière :

	Durée moyenne de l'intervalle		Différence
	entre 2 enfants de même sexe	entre 2 enfants de sexes opposés	
Familles de 7 enfants	27,73	29,64	+ 0,91
» 8 »	26,47	27,31	+ 0,81
» 9 »	25,58	26,24	+ 0,66
» 10 »	22,15	23,03	+ 0,88
» 11 »	21,74	22,43	+ 0,69
» 12 »	21,81	20,88	— 0,93
» 13 »	18,54	19,63	+ 1,09

Valeur moyenne des différences	0,597
Variance des différences	0,472
t de Student	2,30
Probabilité : $\leq 3\%$	

Pour mieux apprécier la différence entre les intervalles commençant par un garçon et ceux commençant par une fille nous donnerons le tableau suivant :

	Durée moyenne des intervalles		Différence
	commençant par un garçon	commençant par une fille	
	mois	mois	
Familles de 7 enfants	28,58	27,76	+ 0,82
» 8 »	27,46	26,16	+ 1,30
» 9 »	26,52	25,36	+ 1,16
» 10 »	22,50	22,69	— 0,19
» 11 »	22,47	21,61	+ 0,86
» 12 »	22,02	20,54	+ 1,48
» 13 »	19,22	18,95	+ 0,27

Valeur moyenne des différences	0,814
Variance des différences	0,351
t de Student	3,62
Probabilité : 1,5 %	

De même que pour le tableau précédent, la comparaison de ces différences entre elles et à leur déviation standard (calculée sur l'ensemble des intervalles dans les familles de même nombre d'enfants) ne révèle aucune discordance intéressante ; aucun des deux phénomènes ne semble s'intensifier régulièrement dans les familles les plus ou les moins nombreuses.

En ce qui concerne la comparaison [(gg) + (ff) — (gf) — (fg)] en fonction du rang de naissance, son étude sur un total de 337 familles de 7 enfants et 258 familles de 8 enfants, conduit aux chiffres suivants :

	Excès moyen des intervalles séparant les enfants de sexes opposés sur les intervalles séparant les enfants de même sexe.	
	Familles de 7 enfants	Familles de 8 enfants
1 <sup>er</sup> intervalle	+ 5,86 mois $\pm$ 0,74	— 0,20 mois $\pm$ 0,82
2 <sup>e</sup> »	— 0,64 »	+ 0,70 »
3 <sup>e</sup> »	+ 0,90 »	+ 0,70 »
4 <sup>e</sup> »	+ 1,25 »	+ 1,58 »
5 <sup>e</sup> »	+ 1,01 »	+ 0,37 »
6 <sup>e</sup> »	+ 0,50 »	— 1,93 »
7 <sup>e</sup> »	...	+ 5,33 »

Nous n'avons encore aucune interprétation satisfaisante de l'irrégularité extrême de ces chiffres : l'excès des variances des différences ainsi calculées (celles-ci sont de 5,05 et 4,99 pour les familles de 7 et 8 enfants respectivement) pose un problème que nous espérons pouvoir résoudre grâce à l'emploi d'un échantillon plus étendu. Notons toutefois qu'il ne semble y avoir aucune variation systématique régulière avec le rang de naissance, de l'allongement des intervalles entre enfants de sexes opposés.

III. — CAS DE NAISSANCES GÉMELLAIRES.

1° Notre échantillon initial comportait 502 familles se répartissant de la manière suivante :

72 fam. de 5 enf.	57 fam. de 10 enf.	2 fam. de 15 enf.
89 » 6 » 41 »	11 » 1 »	16 »
61 » 7 » 26 »	12 » 1 »	18 »
81 » 8 » 18 »	13 »	
49 » 9 » 4 »	14 »	
455 familles comprennent une seule paire gémellaire.		
40 » » 2 » »		
5 » » 3 » »		

Enfin dans une famille des jumeaux sont nés 4 fois et 5 fois dans une autre. Au moment où nous rédigeons cet article, la collecte des matériaux se poursuit et nous ne pouvons encore utiliser les nouveaux documents à notre disposition que pour préciser quelques points très particuliers.

2° Distribution du sexe dans les paires gémellaires :

- 268 paires (garçons, garçons).
- 282 paires (garçon, fille).
- 282 paires (fille, fille).

L'on sait que si le phénomène de monozygotie n'intervenait pas pour élever la proportion des paires gémellaires constituées par deux enfants du même sexe les fréquences des trois types devraient être proportionnelles à  $p^2$ ,  $2pq$ , et  $q^2$  respectivement (avec  $p = 1 - q =$  proportion des naissances masculines). De fait les proportions observés sont beaucoup plus voisines de  $1/3$ ,  $1/3$ ,  $1/3$  comme l'ont déjà signalé tous les auteurs. Il en est à peu près de même pour les triplets : dans un échantillon de triplets né en 1947, 1948, 1949 dont nous devons la plus grande partie à l'obligeance du service médical du lait Gloria, on relève :

3 garçons .....	12 cas
2 » 1 fille .....	17 »
1 » 2 » .....	18 »
0 » 3 » .....	13 »
Total .....	60 cas

Et de même G.W.D. Hamlett ayant recherché toutes les naissances quadruples survenues aux Etats-Unis entre 1915 et 1930 donne les chiffres suivants :

4 garçons .....	13 naissances
3 » 1 fille .....	6 »
2 » 2 » .....	12 »
1 » 3 » .....	7 »
0 » 4 » .....	10 »
Total .....	48 naissances

Dans aucun de ces cas il n'est possible d'écarter par un test statistique l'hypothèse d'une équiprobabilité des diverses compositions, et il est assez amusant que les deux mécanismes de la polyembryonie et de la polyovulation se balancent assez exactement pour remplacer apparemment une répartition du type Gibbs par une répartition du type Bose-Einstein !

Il faut remarquer toutefois le léger excès des naissances féminines (non significatif statistiquement sur ces faibles échantillons) dans les naissances doubles et triples, ce qui confirme un travail antérieur de notre maître M. le Professeur R. Turpin sur l'ensemble des naissances gémellaires en France.

Cette opposition entre la sex-ratio générale et la sex-ratio dans les naissances gémellaires est encore plus nette si l'on compare les trois cas suivants :

1° Proportion des garçons dans les familles françaises contemporaines de plus de 5 enfants (1) sans naissances gémellaires :  $48.495/95.725 = 0,5066 \pm 0,0016$ .

2° Proportion des garçons parmi les naissances simples dans les familles ayant présenté au moins une paire gémellaire :  $2.570/4.887 = 0,5259 \pm 0,0071$ .

3° Proportion des garçons à l'intérieur des paires gémellaires :  $818/1.664 = 0,4916 \pm 0,0123$ .

Un test par chi-carré conduit à une valeur de 8,58 pour deux degrés de liberté (soit une probabilité inférieure à 2 %) et permet de conclure à l'hétérogénéité de ces trois proportions, il semble qu'il y ait association entre la gémellité et l'excès des naissances masculines dans les familles.

Nous avons recherché si ce phénomène nouveau était plus marqué dans tel ou tel type de gémellité. En nous limitant à 776 familles n'ayant présenté qu'une seule naissance gémellaire nous avons obtenu :

- 236 familles avec une paire (garçon, garçon) : 1.500 enfants, dont 807 garçons.
  - 258 familles avec une paire (garçon, fille) : 1.505 enfants, dont 783 garçons.
  - 262 familles avec une paire (fille, fille) : 1.490 enfants, dont 775 garçons.
- soient des proportions de 0,5380 ; 0,5380 ; 0,5201 respectivement

Aucune différence significative entre ces trois chiffres n'apparaît donc nettement dans notre échantillon.

Il peut être intéressant enfin d'étudier plus spécialement les familles où se sont produites plusieurs fois des naissances gémellaires : nous disposons aujourd'hui de 66 familles avec 2 paires de jumeaux qui se répartissent ainsi :

2 paires (garçon, garçon) .....	3 familles
1 » » 1 paire (garçon, fille) ..	13 »
1 » » 1 » (fille, fille) ....	18 »
2 » (garçon, fille) .....	9 »
1 » » 1 paire (fille, fille) ....	16 »
2 » (fille, fille) .....	7 »

En prenant en première approximation comme valeurs théoriques  $22/3$ ,  $44/3$ ,  $44/3$ ,  $22/3$ ,  $44/3$  et  $22/3$ , on obtient un chi-carré de 4,14 (pour 5 degrés de liberté). Notre échantillon ne nous permet donc pas de mettre en évidence une éventuelle prédisposition de certaines familles à produire plus fréquemment tel ou tel type de jumeaux. Notons enfin que cette proportion élevée de familles présentant plusieurs naissances multiples (75 sur 851 au moment où nous écrivons cet article) est en réalité moins forte qu'elle ne paraît en raison de la sélection initiale de nos familles limitées à celles qui comprennent plus de 5 enfants.

3° Influence du rang de naissance et de l'âge de la mère sur la fréquence de la gémellité.

C'est un fait connu depuis longtemps que la fréquence de la gémellité (et tout spécialement de la gémellité dizygote) s'accroît avec l'âge de la mère. Etant donné la

(1) Ces chiffres proviennent de notre travail antérieur qui avait porté sur un matériel tout à fait comparable. On notera cependant la régularité de la faible valeur de la sex-ratio dans les familles sans naissances gémellaires.

corrélation très élevée qui existe entre ce paramètre et le nombre des grossesses antérieures il a été encore impossible de distinguer les effets relatifs de ces deux facteurs.

Nous avons retrouvé ce phénomène dans notre échantillon comme le montrent les chiffres ci-dessous calculés sur des familles n'ayant présenté qu'une seule naissance gémellaire :

	136 familles de 5 enfants	145 familles de 6 enfants	110 familles de 7 enfants
Age moyen de la mère au moment de la naissance de jumeaux .....	27,7	28,6 ans	32,7
Age moyen de la mère aux autres naissances .....	29,88	35,5 ans	29,6
Rang moyen des jumeaux ....	3,07	3,78	4,16
Rang moyen des autres enfants	2,31	2,81	3,57

De fait la discussion de la signification exacte de ces chiffres est assez difficile pour un ensemble de raisons qui tiennent tant à la nature de notre échantillonnage qu'à des causes psychosociologiques. Même lorsque nous savons qu'il s'agit d'une famille « achevée » (et ceci est rarement le cas pour les familles de 5 et 6 enfants, qui sont les plus fréquentes), nous ignorons encore si les naissances gémellaires survenant après 3 ou 4 autres grossesses n'ont pas arrêté l'activité reproductrice du couple : dans les familles de 5 et 6 enfants la proportion des naissances gémellaires survenant au dernier rang atteint 46 % et 41 % respectivement contre 28 % (familles de 7 enfants), 27 % (familles de 8 enfants), 11 % (familles de 9 enfants).

Quoiqu'il en soit de ce facteur, il n'a que vraisemblablement bien peu de raisons de jouer de manière différente pour les divers types de grossesse et nous pouvons considérer que la constatation régulière d'un rang moyen plus élevé pour les naissances de paires gémellaires formées de 2 enfants de sexe opposé (donc nécessairement dizygotes) traduit fidèlement un phénomène biologique constant.

#### 4° Les intervalles séparant les naissances gémellaires des naissances voisines.

Pour notre échantillon initial de 502 familles, nous avons calculé les intervalles séparant les naissances gémellaires des naissances simples les précédant ou les suivant immédiatement, et nous avons obtenu les chiffres du tableau suivant.

Nous trouvons encore ici le phénomène déjà mis en valeur pour les intervalles entre naissances simples : l'allongement de l'intervalle entre naissance de sexes opposés, qui est ici de 1 mois 1/2 à peu près pour les intervalles pré-gémellaires et de 2 pour les intervalles post-gémellaires. Remarquons qu'il est par contre difficile de voir dans la durée anormalement longue des intervalles

où g désigne un garçon.  
f » une fille.  
x » un enfant de sexe quelconque.

Type de l'intervalle	Nombre de cas	Durée moyenne
g. (xx) .....	254	31,45 mois
f. (xx) .....	224	32,40
x. (gg) .....	150	32,07
x. (gf) .....	158	32,53
x. (ff) .....	170	31,09
g. (gg) ou f. (ff) .....	163	30,84
x. (fg) .....	158	32,53
f. (gg) ou g. (ff) .....	157	32,35
Moyenne : x. (xx) .....	478	31,89 ± 0,92 mois
(gg) x. ....	108	30,27
(gf) x. ....	123	33,80
(ff) x. ....	135	29,81
(xx) g. ....	191	31,48
(xx) f. ....	175	31,08
(gg) g. ou (ff) f. ....	124	29,09
(gf) x. ....	123	33,80
(ff) g. ou (gg) f. ....	119	31,08
Moyenne : (xx) x. ....	366	31,32 ± 1,07

du type x. (gf) autre chose que l'influence de cet accroissement considérable de l'intervalle précédant la dernière grossesse. C'est ceci qui explique en outre, nous semble-t-il, que les intervalles post-gémellaires (pourtant situés à des rangs plus élevés en moyenne) n'apparaissent pas plus longs que les intervalles pré-gémellaires : une fraction importante de ces derniers est en effet constituée par les intervalles finals dont nous avons déjà mis en valeur la durée particulièrement importante dans les familles ne présentant que des naissances simples.

Par contre, nous ne trouvons ici aucun signe bien net confirmant le fait que les intervalles débutant par un garçon soient plus longs que ceux débutant par une fille, puisque, si la différence (gg) x — (ff) x est positive, la différence g (xx) — f (xx) est négative. Ce point appelle de nouvelles recherches. Nous ne pouvons encore savoir s'il ne s'agit pas simplement d'une fluctuation d'échantillonnage : le nombre des intervalles considérés laisse encore au hasard une marge de variation importante, et c'est pourquoi nous nous sommes bornés au calcul des valeurs moyennes sans en faire encore un test statistique rigoureux.

#### CONCLUSION

L'étude de ce matériel relativement considérable n'est pas encore entièrement terminée, et même, dans certains cas, la collecte des documents doit encore se poursuivre. Cependant, nous avons pu arriver à des conclusions certaines en ce qui concerne les variations de la durée des intervalles en fonction :

- du nombre d'enfants dans la famille,
- du nombre de grossesses antérieures,
- du sexe des enfants limitant l'intervalle.

En ce qui concerne les naissances gémellaires, nous avons pu montrer :

- l'accroissement de leur fréquence avec l'âge de la mère,
- avec le nombre des grossesses antérieures,
- l'élévation remarquable de la sex-ratio dans les familles de jumeaux.

D'autres points n'ont encore pu être complètement élucidés :

- les interactions entre le rang de naissance et le sexe des enfants sur les variations des intervalles,
- l'accroissement excessif de l'intervalle précédant le dernier enfant,
- la sex-ratio des naissances multiples,
- la distribution des différents types de paires gémeillaires dans les familles en comportant plusieurs,
- les rôles respectifs de l'âge et du passé obstétrical de la mère dans les naissances gémeillaires.

Du point de vue biologique général, trois problèmes nous semblent particulièrement importants :

- la valeur anormalement haute de la sex-ratio dans les familles présentant des naissances gémeillaires ; il s'agit là d'un fait qui ne semble pas encore avoir été signalé et dont nous avouons ne pouvoir encore offrir aucune interprétation réellement satisfaisante,
- l'allongement de l'intervalle séparant deux naissances de sexes opposés,
- l'allongement des intervalles débutant par un garçon.

Ces faits apparaissent d'ores et déjà comme solidement établis avec notre premier échantillon et, s'il est possible, à la rigueur d'invoquer une cause psychologique pour le dernier, il n'en est évidemment pas de même pour le précédent qui traduit certainement un phénomène purement biologique. Ces faits apportent un complément et une précision nouvelle aux résultats que nous avons obtenus avec notre Maître, M. le Professeur R. Turpin : nous avons observé que, même après élimination de tout facteur propre au couple, et le prédisposant à donner naissance à des enfants de l'un ou l'autre sexe, il restait une corrélation entre le sexe des enfants immédiatement consécutifs. Nous proposons d'interpréter ce fait :

- soit par l'existence de périodes gynophiles ou androphiles distribuées de manière aléatoire dans la population et s'étendant sur la durée de plusieurs grossesses et imputables à des variations physiologiques de l'organisme maternel sans liaison avec le sexe des produits antérieurs à la conception,
- soit par une influence du sexe du dernier enfant inhibant partiellement la naissance ultérieure d'un enfant de sexe opposé et se traduisant également par une alternance de périodes gynophiles et androphiles.

Dans ces deux hypothèses, il restait bien entendu que cette action pouvait se faire soit au moment de la con-

ception par une action favorisant tel ou tel type de spermatozoïdes, soit dans les tout premiers temps de la grossesse par une augmentation de la léthalité pour l'un des sexes.

A la lumière de nos derniers résultats, il nous semble que ce soit à la deuxième hypothèse qu'il nous faille plutôt nous rallier. Actuellement, si l'on admet que la présence *in utero* d'un embryon masculin provoque en quelque sorte une réaction adaptative de l'organisme maternel, différente de celle que provoque un embryon féminin, l'on s'explique que puissent être allongés et les intervalles consécutifs à un garçon, et les intervalles séparant des enfants de sexe opposé. On comprend, en outre, que cette action joue de moins en moins à mesure que s'allonge l'intervalle moyen séparant les naissances, ce qui s'accorde avec le fait avancé par C. Gini (et que nos documents antérieurs ne contredisent nullement bien qu'ils soient sans doute trop peu nombreux pour pouvoir être utilisés dans ce but) que les séquences d'enfants de même sexe soient surtout fréquentes au début des familles.

Il ne s'agit là pour nous encore que d'une hypothèse, que l'on considère peut-être comme trop audacieuse. Sa valeur, en tout cas, ne saurait mettre en cause celle du phénomène statistique qu'elle nous a amené à découvrir. Nous espérons pouvoir en apporter des preuves nouvelles dans les travaux que nous poursuivons actuellement.

*Centre de Génétique de l'Hôpital Saint-Louis. (Service du Prof. R. TURPIN).*

*(Travail subventionné par le Fonds d'Etudes de la Société Médicale des Hôpitaux de Paris et de l'Institut National d'Hygiène).*

#### BIBLIOGRAPHIE

- CORRADO GINI. — Il Sesso dal Punto di vista Statistico (1908).  
G. W. D. HAMLETT. — Human twinning in the U. S. A. (1935). *Genetics*, 20, pp. 250-258.  
R. TURPIN et M. A. CARATZALI. — De l'influence de la gémeillité et de l'âge maternel sur la proportion des sexes. *Académie des Sciences* (1937), 104, pp. 151-153.  
R. TURPIN et M. P. SCHÜTZENBERGER. — Recherches statistiques sur la distribution des sexes à la naissance. *Académie des Sciences*, 1948, 226, pp. 1845-1846.  
R. TURPIN et M. P. SCHÜTZENBERGER. — Sur la détermination du sexe chez l'homme. *La Semaine des Hôpitaux de Paris*, 25<sup>e</sup> année, n<sup>o</sup> 60, pp. 2544-2545.  
R. TURPIN, Mme D. DEROCHE et M. P. SCHÜTZENBERGER. — Rapports entre le sexe des nouveau-nés et l'intervalle séparant leur naissance. *Académie des Sciences*, 230, pp. 335-336, 1950.  
M. P. SCHÜTZENBERGER. — Résultats d'une enquête sur la distribution du sexe dans les familles nombreuses. *Semaine des Hôpitaux*, 25<sup>e</sup> année, n<sup>o</sup> 60, pp. 2579-2582.