

PRESENTATION D'UN PROCÉDE  
D'ENREGISTREMENT SYSTEMATIQUE  
DES CARACTERISTIQUES BIO-DEMOGRAPHIQUES  
DE LA POPULATION (\*)

par

Yvan LEPAGE

Laboratoire d'Anthropologie, Université Libre de Bruxelles.  
Prof. Fr. Twiesselmann.

**Introduction.**

Voici les modalités d'élaboration d'un procédé d'enregistrement systématique et exhaustif des populations humaines. Avant toute chose, il convient de démontrer le bien-fondé d'une telle démarche.

L'élaboration des statistiques de population en Belgique sont l'objet de nombreuses transcriptions manuelles des données. Ainsi, l'officier d'état civil chargé d'enregistrer une naissance opère en deux étapes. La première étape consiste en la rédaction de l'acte au registre des naissances. Ces registres restent aux archives de la maison communale. Ils constituent les actes originaux de référence. La seconde étape consiste en la communication à l'Institut National de Statistique des données démographiques. Les caractéristiques à fournir à l'Institut ne concordent pas avec celles des archives. L'officier d'état civil est obligé de recourir aux registres de la population pour remplir correctement les bulletins. On trouve, outre les données démographiques inscrites dans les actes, des questions relatives au mode de constitution de la famille, à la situation sociale du couple, etc.

On s'accorde en général, à reconnaître que les registres d'état civil sont la meilleure source pour l'étude de la démographie.

---

(\*) Communication présentée le 27 mars 1972.

Que les registres d'état civil constituent la meilleure source, ne signifie pas qu'ils soient parfaits, ni même que leur qualité soit équivalente d'une commune à l'autre et d'une époque à l'autre. La qualité des registres n'est pas seule en cause lorsqu'il s'agit d'entreprendre l'étude de la population. La principale difficulté réside dans l'exiguïté du champ d'observation. Chaque administration communale forme une entité autonome. Les renseignements sont saisis au lieu où un des actes a été enregistré. Pour un sujet, qui naît, vit et meurt dans la même commune, l'enregistrement du mariage pose déjà un problème. Si le jeune homme va chercher une femme dans une localité voisine, et si la mariée est célébrée, comme c'est la coutume, au lieu de résidence des parents de l'épouse, alors les données manquent au lieu principal de l'observation : le village de l'époux. Des informations sont perdues, en l'occurrence tout ce qui concerne la famille de l'épouse. Or, comme le mariage est couramment contracté dans le village de résidence des parents de la femme et que la résidence du couple est généralement celle du mari, on enregistre un nombre important d'unions dont la vie se déroule en dehors du champ d'observation.

Les registres d'une seule commune ne permettent pas de suivre avec rigueur la vie de ses habitants et l'émigration temporaire ou définitive ajoute maintes difficultés. Ces raisons font qu'il importe d'étendre le champ d'observation et de le porter soit au niveau de l'arrondissement ou de la province, soit au niveau national.

Il se trouve que l'on dispose, maintenant, grâce à la technologie des ordinateurs, d'un instrument qui permet de changer radicalement les conditions d'observation et d'enregistrement des données d'état civil. Le traitement des archives d'état civil est désormais possible. Il exige certes la consultation et la transcription de masses énormes de données. On sait que pour reproduire ces documents écrits, les faire entrer en ordinateur, la difficulté réside dans la normalisation et l'épuration des données. Le moment vient d'aborder un dispositif d'ensemble pour le maniement des données, de faire en sorte qu'il serve non seulement les intérêts de disciplines connexes comme la démographie, la génétique des populations, la sociologie, la statistique d'état civil ; mais encore des buts

théoriques conjoints comme l'interprétation des régularités, la construction de modèles, l'élaboration de concepts, la manifestation des structures.

L'utilisation des ordinateurs et diverses techniques apparentées ouvrent des possibilités considérables, tant pour la collecte des renseignements que pour leur exploitation, dont le corollaire à ces nouvelles possibilités est l'exigence et la précision des sources. Dans cette perspective, il est fondamental pour les futurs utilisateurs qu'il soit mis sur pied une uniformisation des données.

La nécessité de connaître avec précision les faits de population se fait sentir d'une manière aiguë lorsqu'on désire répondre aux questions que posent les autorités gouvernementales ou internationales aux chercheurs de diverses disciplines : statisticiens de l'état civil, biologistes, économistes, médecins, etc.

Il nous a paru possible de confectionner un dispositif d'ensemble répondant aux exigences qu'impose un état doté de techniques modernes d'enregistrement.

Le procédé préconisé ici vise à enregistrer tous les individus nés en Belgique, à partir d'une certaine date donnée, à l'aide d'un code chiffré.

Le procédé adopté s'inspire directement du projet de Fr. TWIESSLMANN (1962) conçu pour l'enregistrement systématique des maladies et tares de tous les sujets d'une population humaine. Il en appelle également à celui de l'expérience canadienne (NEWCOMBE, H. B., 1959, 1962, 1968, 1969), à l'utilisation des machines à calculer électroniques mais il en diffère par la nature des renseignements récoltés et par les modalités d'application et d'exécution.

La conception du code chiffré d'identification du sujet diffère quelque peu de celui du projet de Fr. Twiesselmann, car nous avons voulu qu'il serve simultanément les intérêts des médecins, des généticiens, des gouvernants, des démographes, des économistes.

Nous envisagerons successivement les points suivants :

1. La confection du code d'identification du sujet, ainsi que les avantages qu'il présente.
2. La confection des cartes reproduisant tous les actes d'état civil.

3. Le commentaire du choix des systèmes de codification.
4. L'application de ce système d'enregistrement pour l'élaboration des statistiques de population.
5. Quelques possibilités d'utilisation d'un tel procédé pour des disciplines comme la médecine, la génétique, etc.

#### **1. Procédé de confection du code d'identification d'un sujet.**

Chaque individu se voit attribuer un numéro d'identification qui le caractérisera dans quelque situation que ce soit, soit comme EGO, comme parent, comme défunt, comme conjoint. Cette opération de numérotation est la plus délicate. Il est fondamental de classer les individus d'après les éléments qui ont la plus grande puissance de discrimination et qui simultanément présentent le moins de risque de varier dans la suite des données relatives au même individu.

Or, malgré la pléthore de renseignements figurant dans les différents actes des registres de l'état civil belge, ceux-ci manquent d'uniformité quant aux données relatives à la définition d'un même individu. Seuls, le patronyme et les prénoms y figurent d'une manière constante. Encore que des variantes orthographiques dans les noms de famille puissent surgir dans la série d'actes concernant une même lignée.

Le procédé d'identification d'un individu est extrêmement simple. Le sujet est identifié par 18 chiffres.

Les cinq premiers de ces 18 chiffres forment le code de la commune de naissance, tel qu'il est couramment utilisé par l'Institut National de Statistique. Ainsi, le 8.5.035 représente la 35<sup>e</sup> commune (St-Mard), du 5<sup>e</sup> arrondissement (Virton), de la 8<sup>e</sup> province (Luxembourg).

Les 7 chiffres suivants indiquent la date de naissance du sujet. Ainsi, 17.09.970 représente le 17 septembre 1970. Le premier chiffre du millésime n'est pas noté pour économiser une colonne de la carte. Cette suppression permettra néanmoins l'utilisation de ces cartes jusqu'à la 999<sup>e</sup> année après la mise en route du fichier.

Les deux chiffres suivants sont les deux chiffres de droite du numéro d'ordre sous lequel le sujet est enregistré à la naissance dans les registres de l'état civil. Deux chiffres suffisent, car il naît en Belgique plus ou moins 150.000 enfants par an. Ce qui fait en moyenne par jour un peu plus de 400 naissances, et il est

hautement improbable qu'une commune enregistre 100 naissances la même journée.

Avec ces 14 chiffres, le sujet est défini et ne peut plus être confondu avec un autre.

Les trois chiffres suivants sont constitués par la nationalité.

Enfin le dernier chiffre indique le sexe du sujet ou son appartenance éventuelle à un couple de jumeaux. Le 1 indique le sexe masculin ; le 2, le sexe féminin ; le 3, l'appartenance à une paire de jumeaux masculins ; le 4, l'appartenance à une paire de jumeaux féminins ; le 5, un garçon et le 6, une fille appartenant à une paire de jumeaux de sexes différents.

En résumé, le sujet est identifié comme suit :

8.5.035.17.09.970.12.404.1

C'est un garçon, né le 17 septembre 1970 dans la commune de Saint-Mard et inscrit dans les registres d'état civil sous un numéro d'ordre se terminant par 12 ; il est de nationalité belge.

On voudra bien remarquer que le nom du sujet de même que son prénom n'auront plus aucun intérêt. De la sorte, une fois établie en code, l'identification du sujet est déterminée tout en restant anonyme. Ce code situe le sujet dans l'espace et dans le temps avec des données immédiatement utilisables en vue de statistique démographiques usuelles. Il évite la codification du nom et des prénoms soit par un système numérique ou phonétique.

## **2. Confection des cartes des actes de l'état civil.**

Afin d'intégrer l'individu dans sa cellule familiale, il suffit d'identifier les parents par un code semblable à celui décrit pour la définition du sujet. L'ordre d'inscription est à la suite de la définition du sujet, celui du père d'abord, celui de la mère ensuite. Il n'est pas nécessaire de reproduire dans la définition du père et de la mère, l'indication du sexe. L'ordre d'inscription, le père d'abord, la mère ensuite, le remplace. Il en est de même pour les actes de mariage, pour les actes de divorce, et pour les actes de séparation de corps et biens : on trouvera d'abord la définition de l'homme ensuite celle de la femme.

Les données d'état civil sont retranscrites sur des cartes mécanographiques de type courant comportant 80 colonnes. Le

procédé de transcription est extrêmement simple. On réserve les deux premières cases à l'identification des différents types d'actes. Ainsi, le 01 identifie l'acte de naissance ; le 02, l'acte de mariage ; le 03, l'acte de décès ; le 04, l'acte d'émigration ; le 05, l'acte d'immigration ; le 06, l'acte de divorce ; le 07, l'acte de séparation de corps et biens. En prévoyant deux cases pour l'identification des différents types d'actes, cela nous permet d'envisager la création de nombreux fichiers avec des données correspondant à différentes étapes de la vie du sujet.

Ainsi constitué, le fichier national exhaustif permanent de la population ne devrait comporter que les données telles qu'elles sont présentées ici.

Les cartes relatives aux actes de naissance comprennent les renseignements suivants :

— lieu de naissance du sujet	5 cases
— date de naissance du sujet	7 »
— numéro d'enregistrement à la naissance du sujet	2 »
— nationalité	3 »
— sexe	1 »
— lieu de naissance du père	5 »
— date de naissance du père	7 »
— numéro d'enregistrement à la naissance du père	2 »
— nationalité du père	3 »
— niveau d'études	1 »
— catégorie socio-économique du père	2 »
— lieu de naissance de la mère	5 »
— date de naissance de la mère	7 »
— numéro d'enregistrement à la naissance de la mère	2 »
— nationalité de la mère	3 »
— niveau d'études	1 »
— catégorie socio-économique de la mère	2 »
— domicile des parents	5 »

Les cartes des actes de naissance totalisent 63 cases.

Les cartes relatives aux actes de mariage comprennent les données suivantes :

— lieu de naissance du mari	5 cases
— date de naissance du mari	7 »
— numéro d'enregistrement à la naissance	2 »
— nationalité du mari	3 »
— niveau d'études	1 »

— catégorie socio-économique	2 cases
— domicile du mari	5 »
— état civil du couple	1 »
— lieu de célébration du mariage	5 »
— date de la célébration du mariage	7 »
— lieu de naissance de la femme	5 »
— date de naissance de la femme	7 »
— numéro d'enregistrement à la naissance	2 »
— nationalité de la femme	3 »
— niveau d'études	1 »
— catégorie socio-économique de la femme	2 »
— domicile de la femme	5 »
— contrat de mariage	1 »

Les cartes des actes de mariage totalisent 64 cases.

Les cartes relatives aux actes de décès contiennent les caractéristiques suivantes :

— lieu de naissance du défunt	5 cases
— date de naissance du défunt	7 »
— numéro d'enregistrement à la naissance	2 »
— nationalité du défunt	3 »
— sexe	1 »
— domicile du défunt	5 »
— lieu du décès	5 »
— état civil du défunt	1 »
— niveau d'études	1 »
— catégorie socio-économique	2 »
— date du décès	7 »
— localisation du décès	1 »

Les cartes des actes de décès totalisent 40 cases.

Les cartes relatives aux actes d'émigration contiennent les données suivantes :

— lieu de naissance	5 cases
— date de naissance	7 »
— numéro d'enregistrement à la naissance	2 »
— nationalité	3 »
— sexe	1 »
— niveau d'études	1 »
— catégorie socio-économique	2 »
— état civil	1 »
— commune de sortie	5 »
— date de sortie	7 »
— commune de destination	5 »

Les cartes d'émigration totalisent 39 cases.

Les cartes relatives aux actes d'immigration contiennent les rubriques suivantes :

— lieu de naissance	5 cases
— date de naissance	7 »
— numéro d'enregistrement à la naissance	2 »
— nationalité	3 »
— sexe	1 »
— niveau d'études	1 »
— catégorie socio-économique	2 »
— état civil	1 »
— commune d'entrée	5 »
— date d'entrée	7 »
— commune de provenance	5 »

Les cartes relatives aux actes d'immigration totalisent 39 cases.

Les cartes relatives aux actes de divorce et de séparation de corps et biens comportent les mêmes caractéristiques. La différence réside dans la détermination du fichier. Les données se trouvant dans ces types d'actes sont les suivantes ;

— lieu de naissance de l'homme	5 cases
— date de naissance de l'homme	7 »
— numéro d'enregistrement à la naissance	2 »
— nationalité	3 »
— niveau d'études	1 »
— catégorie socio-économique	2 »
— date de dissolution du mariage	7 »
— lieu de dissolution du mariage	5 »
— lieu de naissance de la femme	5 »
— date de naissance de la femme	7 »
— numéro d'enregistrement à la naissance	2 »
— nationalité	3 »
— niveau d'études	1 »
— catégorie socio-économique	2 »

Les cartes de divorce ou de séparation de corps et biens totalisent 52 cases.

On dispose, pour les Belges, de toutes les données à porter sur les cartes mécanographiques. Si les parents sont nés à l'étranger, seront codées les caractéristiques dont on disposera. L'absence de renseignements sera uniformément transcrite par des O.

Si le lieu de naissance ou les domiciles précédents sont connus pour les étrangers, le domaine de variation du code des communes belges nous laisse la possibilité d'exploiter l'intervalle

compris entre 00001 et 09999. Il suffit de convenir une fois pour toutes, que les codes formés par les chiffres allant de 00001 à 00095 identifieront les départements français ; de 00101 à 00192, les provinces italiennes ; de 00201 à 00233, les comtés écossais, par exemple. Tous les pays européens peuvent être aisément codifiés de cette manière. Pour les pays extra-européens, on se contentera de la codification du pays.

### 3. **Choix des systèmes de codification.**

Avant d'aborder les possibilités d'exploitation d'un tel enregistrement de l'état civil, voyons le détail du choix de certains systèmes de codification.

#### A.— CATÉGORIE SOCIO-ÉCONOMIQUE.

Le classement de la population suivant la profession peut être fait de divers point de vue. On peut s'intéresser au métier : cultivateur, maçon, boulanger, avocat, ... ; on peut, d'un point de vue plus économique, mettre l'accent sur la branche d'activité où s'exerce la profession et classer dans la même catégorie, la métallurgie par exemple, aussi bien les ouvriers métallurgistes que les comptables et les secrétaires qui travaillent dans les entreprises de cette branche ; on peut enfin essayer d'atteindre, à travers l'activité, le milieu social ; on classe alors les personnes suivant leur catégorie socio-professionnelle ; celle-ci dépend du métier exercé, de la branche d'activité dans laquelle il s'exerce, de la qualification, du niveau hiérarchique et du statut.

Les métiers individuels, les activités collectives, la hiérarchie sociale se modifient au cours de l'histoire et l'on ne saurait donner un classement convenant à toutes les époques. D'autre part, les activités humaines sont de plus en plus nombreuses et l'on est obligé de simplifier beaucoup si l'on veut en prendre une vue d'ensemble ; et cette simplification est rendue nécessaire par le manque fréquent de précision des indications que fournissent les actes d'état civil.

Voici à titre d'exemple, les grands groupes socio-professionnels utilisés lors de la publication des recensements de 1954 et de 1962 en France :

1. professions libérales et cadres supérieurs ;

2. patrons de l'industrie et du commerce (y compris les artisans) ;
3. cadres moyens ;
4. employés ;
5. ouvriers ;
6. agriculteurs exploitants ;
7. salariés agricoles ;
8. personnel de service ;
9. autres catégories (clergé, armée, police, arts et spectacle).

La nomenclature des professions, publiée par l'Institut National de Statistique, répertorie les professions par un code formé de cinq chiffres. Les deux premiers chiffres de ce code identifient le corps d'activité. Ces corps d'activités sont catalogués sous 46 rubriques, numérotées de 1 à 45 et un rebut reconnu par le chiffre 99. Les trois chiffres suivants indiquent le rang de la profession à l'intérieur de la rubrique. Ainsi, le 01.046 constitue la 46<sup>e</sup> profession (planteur de tabac) du premier groupe d'activité (agriculture et élevage).

Ce système de codification est peu efficient. En effet, devant la difficulté que représente la détermination des professions, la lourdeur de manipulation et le peu d'utilité pratique d'un tel code, il serait inopportun d'enregistrer des caractéristiques qui lors des phases d'exploitation seraient immédiatement simplifiées ou inemployées. C'est pourquoi, par un souci de simplification et d'utilisation pratique, il vaut mieux s'en tenir à des classifications couramment utilisées dans les enquêtes, dans les études et même lors de la publication de certains recensements (cf. en France). La sériation en catégories socio-professionnelles pose des problèmes d'interprétation. Il faut en combinant des notions différentes, telles la profession proprement dite, la situation matérielle et la catégorie sociale, tenter de classifier en se rapprochant le plus possible de la réalité socio-économique. Devant les différentes solutions adoptées, il nous paraît opportun de s'en tenir à une solution combinant les sériations par branche d'activité économique suivant la classification-type préconisée par les Nations-Unies et le statut social.

Ces activités économiques sont :

1. agriculture, sylviculture, pêche et chasse ;
2. industrie extractive (extraction du charbon, minerais et

- minerais divers, extraction de pierres à bâtir et autres matériaux utilisés pour la construction) ;
3. industries manufacturières ;
  4. industries du bâtiment et de la construction ;
  5. électricité, gaz, eau, vapeur et chaleur ;
  6. commerce, banque, assurances, affaires immobilières ;
  7. transports, communications et connexes ;
  8. services etc. ;
  9. activités inconnues.

Les catégories de statuts sociaux sont les suivantes :

1. chefs d'entreprises, indépendants, patrons non liés par un contrat d'emploi, professions libérales ;
2. chefs d'entreprises et d'établissements liés par un contrat d'emploi ;
3. employés des cadres supérieurs ;
4. autres employés ;
5. ouvriers ;
6. ouvriers à domicile ;
7. aidants ;
8. état social inconnu.

La classification des activités socio-économiques combinée avec le statut social nous donne un code à deux chiffres. Le tableau suivant illustre de manière concrète la construction de ce code.

État social		1	2	3	4	5	6	7	8
Branches d'activités économiques	1	11	12	13	14	15	16	17	18
	2				24				
	3				34				
	4				44				
	5				54				
	6				64				
	7				74				
	8				84				
	9				94				

Ainsi, si l'on désire isoler la branche d'activité «agriculture» on trie sur le 1 des dizaines. Par contre si on travaille sur l'état social, on trie sur les unités.

#### B.— NATIONALITÉ.

La composition actuelle des territoires risque à chaque instant de subir des modifications de frontières. Il nous est apparu opportun d'adopter un code ouvert. Le premier des trois chiffres indique le continent. Ainsi, l'Afrique sera reconnue par le 1 ; l'Amérique, par le 2 ; l'Asie, par le 3 ; l'Europe, par le 4 ; l'Océanie, par le 5 et l'U.R.S.S., par le 6. Dans les publications démographiques des Nations-Unies, l'U.R.S.S. forme une rubrique à part. Les deux derniers chiffres identifient le pays de ce continent. Les pays sont classés par ordre alphabétique dans l'orthographe anglo-saxonne. C'est ainsi que la nationalité belge sera formée par le chiffre 4.04. A titre indicatif, la nationalité française sera identifiée par le 4.10 ; l'allemande (ouest), par le 4.11 ; la luxembourgeoise, par le 4.21 ; la néerlandaise, par le 4.24.

#### C.— ÉTAT CIVIL DU COUPLE DANS L'ACTE DE MARIAGE.

Dans l'acte de mariage, l'état civil du couple n'occupe qu'une seule case. En effet, il n'est pas nécessaire de reproduire l'état civil au moment du mariage de chacun des conjoints. Il n'existe que 9 possibilités. Le système de codification est façonné à l'aide d'une grille. Cette grille est construite de la manière suivante ; les possibilités au moment du mariage pour l'époux sont placées en abscisses, pour la femme en ordonnées. La lecture se fait verticalement pour le mari et horizontalement pour la femme.

Homme Femme	Célib.	Veuf	Divorcé
Célib.	1	2	3
Veuve	4	5	6
Divorcée	7	8	9

Ainsi, le 8 sera le code d'un couple constitué par un homme veuf se remariant avec une divorcée.

#### D.— NIVEAU D'ÉTUDES.

Ce type de renseignement ne figure sur aucun acte d'état civil. Il n'existe, a priori, aucun obstacle à l'y introduire. On gagnerait par là une information précieuse pour l'analyse de la population. On pourra relever les niveaux suivants :

- absence de scolarité ;
- niveau primaire ;
- niveau secondaire inférieur ;
- niveau secondaire supérieur ;
- niveau supérieur non-universitaire ;
- niveau supérieur non-universitaire technique ;
- niveau universitaire.

La classification du niveau des études vient raffiner la sériation des catégories socio-économiques. Introduire le niveau d'études n'est pas sans intérêt. En effet, le fait d'être chef d'entreprise n'implique pas que l'on soit porteur d'un diplôme de niveau supérieur. De même que l'on pourrait rencontrer un chauffeur de taxi porteur du titre de docteur en droit. La possession de ces renseignements pourra, par exemple, nous indiquer si le fait d'avoir poursuivi des études supérieures a une influence significative sur le comportement de ces familles en matière de procréation.

E.— L'ENREGISTREMENT DES ENFANTS MORT-NÉS ET DES FAUX MORT-NÉS suscite à l'heure actuelle des difficultés. Ce problème pourrait être assez aisément résolu. Que ce soit un faux mort-né ou un mort-né, on relève toutes les caractéristiques d'une naissance normale.

#### 1.— *Les faux mort-nés.*

Les faux mort-nés sont des enfants ayant vécu, mais morts endéans les trois jours suivant leur naissance. En effet, la juridiction belge offre un délai de trois jours pour déclarer une naissance. Au delà de 72 heures de vie, un sujet aura deux actes, celui de sa naissance et celui de son décès. Or, comme cette personne a vécu, on peut adopter le principe de la confection de

deux cartes ; une carte de naissance et une carte de décès. Comme toutes les caractéristiques enregistrées sur les cartes pour les statistiques d'état civil n'occupent pas les 80 colonnes des cartes mécanographiques, il nous est loisible de transcrire des données supplémentaires sur les cases restantes.

2.— *Les mort-nés.*

Étant donné que les mort-nés ne sont affectés d'aucun souffle de vie, ils ne seront dotés que d'une seule carte. Afin de les distinguer des autres, on adopte les chiffres 7 pour la détermination du sexe masculin et le 8 pour le sexe féminin. De cette manière, en triant sur la colonne réservée à l'indication du sexe, on sait immédiatement les isoler des autres naissances, et simultanément, en les enregistrant comme des naissances, on facilite la tâche dans les opérations ultérieurs de reconstitution des familles.

A la suite des renseignements d'état civil qui occupent 63 cases, il nous est loisible d'inscrire sur les cases restantes des données d'ordre médical. Par exemple, on peut réserver des cases à l'inscription d'un accident, d'une cause de maladie, d'une affection héréditaire. L'identification soit d'une maladie, soit d'un accident, soit d'une maladie héréditaire, soit d'un traumatisme sont obtenues en adoptant la codification selon le système décimalisé de l'Organisation Mondiale de la Santé (O.M.S.) ; d'autre part ce système codificateur peut être immédiatement utilisé. Rien non plus n'empêche de créer une seconde carte comportant des renseignements précis sur les causes de décès.

**4. Application de ce procédé pour l'élaboration des statistiques de population.**

Les statistiques de population s'élaborent à partir de trois agencements différents des cartes perforées.

a.— A partir des données inscrites sur les cartes on tire immédiatement certains résultats, tels que les naissances, les conceptions, les mariages, les décès par année civile, les mouvements saisonniers des différents événements, etc.

Pour des raisons d'économie d'exploitation d'un tel fichier, les programmes doivent être conçus de telle manière qu'ils utilisent au maximum les renseignements insérés dans le type d'enregistrement que l'on exploite et ne provoquent qu'un seul déroulement de la bande magnétique.

Dans l'exemple que nous présentons, à partir de l'enregistrement des actes de naissances, un seul programme a été conçu pour obtenir les résultats qui suivent :

- nombre de naissances vivantes, par année civile, par localité, par groupe de localités, selon le sexe et selon l'appartenance à une paire de jumeaux ;
- proportion de naissances multiples par rapport au nombre de naissance et selon le sexe ;
- indice de masculinité ;
- taux de masculinité ;
- mouvement mensuel des naissances ;
- nombres de naissances vivantes rapportés aux effectifs de la population, préalablement calculés, nous donne les taux de natalité « effectifs » ;
- addition du nombre de naissances vivantes et des faux mort-nés rapportés aux effectifs de la population nous donne le taux brut de natalité.

b.— Avant d'aborder la reconstitution des familles, il est d'une grande utilité de reconstituer au préalable, l'historique de chaque individu. La technique consiste à reporter à la suite de l'acte de naissance d'un sujet, tous les actes d'état civil : mariage, immigration, émigration, décès et autres renseignements de chaque individu. Cette opération s'effectue sans manipulation de cartes perforées et fait l'objet d'un enregistrement sur bande magnétique. Chaque acte est doté de deux numéros d'identification. Le premier consiste en l'identification du sujet, établi dès la naissance, le second en celui de l'événement en cause. Le numéro d'identification de l'événement est la date de cet événement. Il est constitué de sept chiffres. Ces deux numéros sont inscrits au regard de l'acte proprement dit. Ainsi constitué, l'historique d'un individu se présente sur les imprimés de la manière suivante (cf. exemple p. 49).

Il est très aisé, à partir de ce type de rassemblement des données, de mettre en pratique l'analyse longitudinale et ainsi

de reconstituer l'histoire naturelle des générations, de connaître avec un haut degré de précision le nombre d'habitants vivant dans telle ou telle localité, de connaître à n'importe quel moment sa structure par âge, selon le sexe, selon l'état matrimonial, etc., d'isoler n'importe quel sous-groupe de la population en partant soit des catégories socio-économiques, soit de la nationalité, etc.

En résumé, l'histoire du sujet XXXXX se présente comme suit : c'est un homme, né le 12 octobre 1880, dans la commune de Saint-Mard et inscrit dans les registres de l'état civil sous un numéro d'ordre se terminant par 21, il est de nationalité belge.

Son père est né le 3 mars 1849, dans la commune de Saint-Mard, inscrit dans les registres d'état civil sous le numéro d'ordre se terminant par 12. Il est de nationalité belge et est ouvrier dans la branche d'activité économique des transports et communications, il est du niveau d'étude primaire.

Sa mère est née le 7 septembre 1856 dans le département de la Meuse (France), le numéro sous lequel elle est inscrite dans les registres d'état civil nous est inconnu. Elle est de nationalité française et n'exerce aucune profession; son niveau d'études est inconnu.

Le couple est domicilié à Saint-Mard.

Le sujet se marie le 3 avril 1903 à Arlon. Son niveau d'études : secondaire moyen inférieur. Au moment de son mariage, il est toujours domicilié à Saint-Mard. Il est employé dans la branche d'activité économique des transports et communications.

Son épouse est née le 5 juin 1883 à Heinsch et est inscrite dans les registres de l'état civil sous le numéro d'ordre se terminant par 34. Elle est de nationalité belge, n'exerce aucune profession, et est du niveau d'études primaire. Au moment du mariage, elle est domiciliée à Arlon.

Ils sont tous deux célibataires et leur union ne fait l'objet d'aucun contrat de mariage par acte notarié.

Le sujet émigre le 5 avril 1903 vers Arlon.

Il rentre dans sa commune d'origine le 24 novembre 1910 en provenance d'Arlon.

Le 18 juillet 1913, le divorce de son union est prononcé.

Il se remarie à Virton, le 30 mars 1920 avec une veuve. Sa seconde femme est née le 18 mars 1887 à Ethe et est inscrite dans

Exemple de l'historique d'un individu  
 Nombre d'actes différents = 7      sujet n° XXXXXX

N° de l'individu	N° des actes	Contenu des actes
8.5.035.17.10.880.21	17.10.880	01.85035171088021.404.1.85035030384912.404.175.000.5070985600.410.000.85035
8.5.035.17.10.880.21	03.04.903	02.85035171088021.404.274.85035.1.81001.0304903.81011050688334.404.100.81001.0
8.5.035.17.10.880.21	05.04.903	05.85035171088021.404.1.274.2.85035.0504903.81001
8.5.035.17.10.880.21	24.11.910	04.85035171088021.404.1.274.2.85035.2411910.81001
8.5.035.17.10.880.21	18.07.913	06.85035171088021.404.274.1807913.85035.81011050688334.404.00
8.5.035.17.10.880.21	30.03.920	02.85035171088021.404.274.85035.6.85045.3003920.85010180388709.404.200.85045.1
8.5.035.17.10.880.21	21.02.955	03.85035171088021.404.1.85035.21016.3.274.2102955.2

les registres de l'état civil sous le numéro d'ordre se terminant par 9. Elle est de nationalité belge, n'exerce aucune profession et est du niveau d'études secondaire moyen inférieur.

Leur union fait l'objet d'un contrat de mariage par acte notarié.

Finalement, le sujet meurt le 21 février 1955 à Bruxelles, veuf, dans un hôpital. Il était toujours domicilié dans sa commune de naissance. Et toute sa vie, il a pratiqué la même profession.

c. — La codification du système proposé permet par une série de tris de dresser un pedigree aussi complet que le rend possible l'ancienneté du système d'enregistrement. En effet, sur la fiche de naissance d'un enfant, se trouve le code du père. En triant ce code mais cette fois sur la fiche de naissance du père, on trouve le code du grand-père de cet enfant. En triant l'ensemble des cartes sur le code du grand-père, on trouve tous les frères et sœurs du père, c'est-à-dire les oncles et les tantes du sujet initial. La méthode peut se répéter à l'infini puisque, des oncles, on passe aux cousins germains du sujet initial, etc ...

L'étude mécanographique des généalogies et de la fécondité exige de recueillir les données de base d'une manière rigoureuse. Il s'agit, dans la conjoncture actuelle, du travail certainement le plus long et le plus fastidieux. La reconstitution des familles s'élabore suivant le même principe que celui mis au point pour reconstituer l'historique d'un individu.

Le numéro d'identification de la famille est obtenu dans l'acte de mariage par la juxtaposition des codes définissant les deux conjoints. Ce même code se retrouve dans l'acte de naissance de l'enfant issu de ce couple, ce sont les codes du père et de la mère.

Pour les enfants, le numéro d'identification de la famille est obtenu à partir de l'acte de naissance du fichier regroupant l'historique de chaque individu. De cette manière, tous les événements relatifs aux enfants issus d'un mariage sont automatiquement classés dans la reconstitution de chaque famille. L'opération d'identification au niveau des parents s'opère dans le même fichier mais à partir de l'acte de mariage.

La phase de reconstitution des familles nécessite de nombreuses manipulations de données. La capacité des ordinateurs de la «troisième génération» rend accessible l'élaboration de ces séries statistiques, et simultanément le prix de revient de

l'utilisation des machines électroniques s'en trouve réduit. L'opération de reconstitution des familles s'exécute en plusieurs phases. En premier lieu, on dresse un tableau à quatre colonnes où sont inscrites les coordonnées d'un individu. La première de ces quatre colonnes est réservée à l'identification du sujet, la seconde à celle de sa femme (ou de son mari), la troisième à celle de son père, la quatrième à celle de sa mère. Le tableau des coordonnées des sujets est mis en mémoire. Le fichier des historiques des individus est imprimé sur un disque. Cette solution ne nécessitera qu'un déroulement de la bande magnétique durant la phase de rassemblement des données. Cette technique est appelée : Random Access Files. Les phases de triage sont opérées à partir du tableau. On rassemble tous les enfants ayant le même père, tous les enfants ayant la même mère, et ensuite tous les enfants ayant le même père et la même mère. Après ces différentes étapes de triage, on imprime les familles, mais les historiques des individus sont retrouvés sur le disque.

Le couplage des données, regroupées par famille offre la possibilité de mener l'analyse de la fécondité, du mode de constitution des familles ce qu'aucune autre méthode ne permet.

### **5. Perspectives d'utilisation.**

C'est dans le domaine des perspectives que les richesses d'exploitation d'un tel système sont les plus prometteuses. De telles expériences ont déjà vu le jour un peu partout dans le monde. C'est ainsi qu'en Islande (BJARNASON O., 1968) sont enregistrées le maximum de caractéristiques démographiques et médicales de la population dans le but d'établir des statistiques précises et variées ; en Angleterre (ACHESON, E. D., 1968) dans le but d'études épidémiologiques ; à Pavie (BARRAI, I., 1968) dans le but d'études de génétique de la population ; à Maryland (BAHN, A. K., 1968) pour des études de cas psychiatriques.

Nous pourrions envisager des réalisations semblables en Belgique sans grandes difficultés. Déjà depuis plusieurs années des renseignements font l'objet d'un enregistrement systématique. Il suffit de coordonner les efforts des différents services chargés de relever ces renseignements et de les enregistrer sous un format standard de telle manière qu'ils puissent être traités en commun suivant les nécessités.

— Renseignements récoltés à la naissance.

Dès la naissance, les nouveau-nés et les nourrissons sont de plus en plus souvent examinés par un médecin. En maternité, on relève systématiquement le poids, la taille, le périmètre crânien ; l'état de santé du nouveau-né fait l'objet de deux mesures, la première après une minute de vie, la seconde après cinq minutes de vie. Cette quantification, dénommée indice d'Apgar, est établie à partir de cinq renseignements ; coloration de la peau, tonus, réflexe, respiration, cœur. L'indice est calculé sur un total de 10 points.

Tous ces renseignements peuvent être portés sur une carte perforée. On notera le poids de l'enfant en kilogrammes et hectogrammes, la taille en centimètres et également le périmètre crânien. Ainsi un nouveau-né d'un poids de 3,200 kg. sera codifié 3,2.

Ces renseignements associés aux données démographiques récoltées à la naissance nous permettront d'établir des tableaux statistiques importants. On pourra connaître, par exemple, la distribution de la valeur du poids à la naissance avec des variations dans le temps et dans l'espace, l'influence du poids à la naissance sur la longévité ou la morbidité...

— Renseignements récoltés au moment du décès.

Au moment de la mort, la cause du décès doit être reconnue par un médecin. Ces certificats médicaux sont systématiquement centralisés. Il serait possible, outre la cause immédiate du décès, d'inscrire deux autres diagnostics médicaux. De la sorte, il est aisé de noter deux affections héréditaires ou non qui auraient pu échapper aux examens précédents. Ces renseignements peuvent être aisément portés sur la carte de l'acte de décès. Ainsi, les colonnes 43 à 54 seront réservées à l'inscription de ces trois diagnostics.

Ces renseignements nous permettront de connaître la distribution de la mortalité selon les causes dans le temps et dans l'espace en fonction par exemple de l'âge, de l'état-civil, de la catégorie socio-économique. A l'aide des généalogies, on pourra savoir s'il existe une corrélation entre les membres d'une même famille au sujet de la longévité, de la morbidité et des causes de décès.

— Renseignements récoltés à différentes étapes de la vie du sujet.

Pour chaque enfant en âge d'école sont relevés les mensurations anthropométriques, l'acuité visuelle, le poids, la taille : il suffit de créer une nouvelle carte perforée comportant au préalable l'identification du fichier et du sujet, soit en tout 20 cases, les 60 cases restantes sont libres pour l'inscription des données rassemblées. Il ne faut jamais omettre de dater l'époque d'observation.

— Les observations réalisées sur la population masculine à la fin de la puberté au centre de recrutement de l'armée pourraient être également collectées et intégrées dans cet immense fichier de la population.

A ces examens médicaux, une fraction importante de la population est soumise, il s'y ajoute une série de groupes humains également soumis à des observations médicales. Il s'agit essentiellement du personnel des administrations publiques et des grandes entreprises privées.

Associés aux données démographiques, ces renseignements nous permettraient de dresser un bilan de l'état de santé de la population et d'établir des statistiques détaillées de santé publique mettant en évidence des carences et de cette manière facilitant la tâche des autorités compétentes pour mener une politique de santé publique plus efficace.

— L'exploitation de l'enregistrement des données regroupées en familles offre à l'analyse démographique et à la génétique des populations humaines ce qu'aucune autre méthode ne permet d'entreprendre.

En effet, les statistiques nationales de population manifestent surtout des insuffisances lorsqu'on désire analyser la fécondité et la mode de constitution des familles. Les carences sont diverses et les principales sont :

— la fécondité des mariages et le mode de constitution des familles, en particulier, ne sont révélés que de façon grossière par les recensements ;

— l'âge des enfants à l'époque d'un recensement renseigne mal sur les intervalles qui les séparent au sein de la famille ;

— on ne peut tenir aucun compte de l'âge des enfants, ni de leur sexe lorsque ceux-ci ne résident plus chez leurs parents ;

— il est impossible de distinguer les jumeaux de deux frères (ou de deux sœurs) nés à moins d'une année d'intervalle ;

— les renseignements sur les enfants mort-nés et les enfants décédés en bas âge font défaut.

Outre les paramètres démographiques traditionnels, ce système d'enregistrement des populations humaines nous permettra d'analyser et de répondre, par exemple, aux questions suivantes :

— de considérer les conceptions préuptiales comme influençant les risques de divorce ;

— la baisse de la fécondité s'accompagne-t-elle plutôt d'une réduction de la période de procréation que d'un allongement des intervalles entre naissances ;

— de mener à bien les analyses longitudinales de la nuptialité et de la mortalité ;

— de connaître avec précision quelle est l'influence du rang de naissance sur la mortalité infantile ;

— quels sont les effets des naissances de mort-nés et de décès d'enfants en bas âge sur le comportement ultérieur des parents ;

— y a-t-il diffusion d'un groupe à un autre de certains modèles de famille ;

— l'avancement de l'âge au mariage est-il la conséquence d'une plus forte proportion de conceptions préuptiales ou pas ;

— quels sont les effets des naissances multiples sur le comportement ultérieur des parents ;

— quels sont les risques associés aux naissances multiples. Les risques de mortinatalité et de décès sont-ils plus grands pour les jumeaux que pour les autres naissances. Ces risques sont-ils plus élevés pour les jumeaux de même sexe que pour les jumeaux de sexes différents ;

— y a-t-il une corrélation entre les événements survenus dans une même famille, c'est-à-dire, quels sont les risques courus par les frères ou les sœurs d'enfants mort-nés ;

— les naissances illégitimes sont-elles plus sujettes au décès que leurs homologues légitimes ;

— quel est l'apport des étrangers à la population, la contribution donnée par la descendance de ces sujets à la population ultérieure de la population ;

— quelles sont les conséquences des mariages consanguins ;

— influence de l'éloignement géographique des lieux de naissance sur la conclusion des mariages et sur les caractéristiques démographiques ;

— calcul de la fraction fertile de la population. La connaissance de cette donnée a un très grand intérêt social et hérédologique.

### Conclusions

Il nous a paru possible de concevoir un dispositif d'ensemble permettant d'enregistrer systématiquement les caractéristiques bio-démographiques de la population. Le système proposé est lié aux possibilités des ordinateurs. Ceux-ci ont progressé à un tel rythme depuis quelques années qu'il faut s'attendre à une simplification du travail de manipulation et du couplage des données dans un avenir relativement proche. Il y a des raisons de s'intéresser aux problèmes de l'emmagasinage des données. Dès la mise en pratique de l'enregistrement, des renseignements précieux pourront être attendus. Le but lointain vise à la reconstitution des familles. La grande force de la reconstitution des familles réside dans son aptitude à pénétrer plus réellement la démographie d'une communauté que les méthodes d'analyse globale.

Le but poursuivi a été de doter d'un instrument de base permettant d'établir des statistiques de population aussi variées que précises à des disciplines diverses comme la génétique, la médecine, la biologie des populations, les sciences économiques, la sociologie, leur offrant la possibilité d'aborder des problèmes dont l'accès est à présent difficile, si pas fermé.

Nous tenons à souligner que les codifications des patronymes ainsi que les prénoms n'ont plus aucun intérêt. Le système proposé est anonyme. L'identification patronymique des individus est remplacée par des données immédiatement utilisables en vue de statistiques démographiques usuelles.

### Références

- ACHESON, E. D.  
1968 The Oxford record linkage study. The first five years.  
E. D. Acheson, éd. Record linkage in medicine. London,  
Livingstone : 46-51.
- BAHN, K. et coll.  
1968 Longitudinal studies using psychiatric case registres.  
E. D. Acheson, éd. Record linkage in medicine. London,  
Livingstone : 226-251.

- BARRAI, I. et coll.  
 1968 Further studies on record linkage from parish books.  
 E. D. Acheson, éd. Record linkage in medicine. London,  
 Livingstone : 270-282.
- BJARNASON, O. et coll.  
 1968 Record linkage in a self contained community.  
 E. D. Acheson, éd. Record linkage in medicine. London,  
 Livingstone : 62-69.
- CUISENIER, J., SEGALEN, M. et DE VIRVILLE, M.  
 1970 Pour l'étude de la parenté dans les sociétés européennes ; le  
 programme d'ordinateur ARCHIV.  
*L'Homme*, 10 (3) : 27-74.
- FLEURY, M. et HENRY, L.  
 1965 Nouveau manuel de dépouillement et d'exploitation de l'état civil  
 ancien.  
 Paris, I.N.E.D.
- GANIAGE, E.  
 1963 Trois villages d'Ile-de-France au XVIII<sup>e</sup> siècle.  
 Paris, I.N.E.D.
- GAUTHIER, E. et HENRY, L.  
 1958 La population de Crulai, paroisse normande.  
 Paris, I.N.E.D.
- GIRARD, P.  
 1959 Aperçus de la démographie de Sotteville-lez-Rouen vers la fin du  
 XVIII<sup>e</sup> siècle.  
*Population* : 14 : 485-508.
- GOMILA, J. et GUYON, L.  
 1969 Etude comparative de petites communautés rurales. Méthodes et  
 premiers résultats à propos de la consanguinité à Bois-Vert  
 (Quebec).  
*Population*, 24 : 1127-1153.
- GOMILA, J. et PHILIPPE, P.  
 1971 Structure de population et mariages à l'Isle-aux-Coudres (Québec).  
*Population*, 26 : 707-716.
- HENRY, L.  
 1953 Une richesse démographique en friche : les registres paroissiaux.  
*Population*, 8 : 281-290.
- JACQUARD, A. et NADOT, R.  
 1968 Mariages consanguins et fertilité différentielle.  
*Population*, 23 : 325-332.
- JACQUARD, A.  
 1970 Structures génétiques des populations.  
 Paris, Masson.
- MCKUSICK, V. A. and CROSS, H. E.  
 1968 Genealogical linkage of records in two isolate populations.  
 E. D. Acheson, éd. Record linkage in medicine. London,  
 Livingstone : 263-270.

- MUGNIER, M., SUTTER, J. et GOUX, J. M.  
1965 Constitution des généalogies à partir de l'état civil par une méthode mécanographique.  
*Bull. Mém. Soc. Anthropol.* Paris, **8**, XI<sup>o</sup> série : 39-44, cahiers du C.R.A. n<sup>o</sup> 3.
- 1966 Organigrammes pour l'étude mécanographique de la parenté et de la fécondité dans une population.  
*Population*, **21** : 75-98.
- NEWCOMBE, H. B., KENNEDY, J. M., AFORD, S. J., JAMES, A. P.  
1959 Automatic linkage of vital records.  
*Science*, **130** (3381) : 954-959.
- NEWCOMBE, H. B. and KENNEDY, J. M.  
1962 Record linkage. Making maximum use of the discrimination power in identifying information.  
*Communications of the association for computing machinery*, **5** (II) : 563-566.
- NEWCOMBE, H. B.,  
1967a Present state and long term objectives of the British Columbia population study.  
Proceedings of the 3rd. international Congress of Genetics. Chicago-Baltimore, the Johns Hopkins press : 291-313. ID.
- 1967b Record linking ; the design of efficient systems for linking records into individual and family histories.  
*American Journal of human Genetics*, **19** (3, 1) : 335-259.
- 1968 Multigeneration pedigrees from linked records.  
E. D. Acheson, éd. Record linkage in medicine. London, Livingstone : 295-301.
- 1969 Couplage des données pour les études démographiques.  
*Population*, **24** : 653-684.
- RAMBAUD, P.  
1962 Economie et sociologie de la montagne ; Albiez-le-Vieux en Maurienne.  
Paris, A. Collin.
- SUTTER, J. et TABAH, L.  
1956 Méthode mécanographique pour établir la généalogie d'une population-Application à l'étude des Esquimaux polaires.  
*Population*, **11** : 507-530.
- TWIESSELMANN, F.  
1962 Présentation d'un procédé d'enregistrement systématique des maladies et des tares des populations humaines.  
*Bull. Acad. royale de Médecine de Belgique*, VII<sup>o</sup> série, **2** (1) : 93-111.

Adresse de l'auteur : Y. LEPAGE,  
rue César Franck, 8,  
B1050 Bruxelles.