

**UNIVERSITE DU DROIT ET DE LA SANTE – LILLE 2  
FACULTE DE MEDECINE HENRI WAREMBOURG**

**Année 2006**

**THESE POUR LE DIPLOME D'ETAT  
DE DOCTEUR EN MEDECINE**

**ETUDE RETROSPECTIVE DE 17 ANS D'ACTIVITE DU  
CENTRE D'EDUCATION POUR LE TRAITEMENT DU  
DIABETE ET DES MALADIES DE LA NUTRITION DU CHG  
DE ROUBAIX :  
TRAITEMENT DE BASE DE DONNEES,  
ANALYSE, INDICATEURS D'UTILITE**

**Présentée et soutenue publiquement le 25 janvier 2006**

**Par**

**Anne-Sophie DESEINE  
Emmanuel CHAZARD  
Bertrand LEGRAND**

**Jury**

**Président :** Monsieur le Professeur SALOMEZ

**Assesseurs :** Monsieur le Professeur BEUSCART  
Monsieur le Professeur FONTAINE  
Monsieur le Professeur GLANTENET  
Monsieur le Docteur GRENIER

**Directeur de Thèse :** Monsieur le Docteur GRENIER

# Sommaire

Sommaire.....	2
Remerciements .....	5
INTRODUCTION.....	9
I.    Propos général .....	9
A.    L'objectif de la thèse .....	9
B.    Tour d'horizon de la bibliographie .....	9
II.   Situation sanitaire de la population concernée.....	10
A.    Situation sanitaire de la population diabétique en France en 2001 [14] .....	10
B.    Situation sanitaire de la population diabétique dans le Nord Pas de Calais.....	13
C.    Situation sanitaire du Versant Nord-Est de la Métropole Lilloise (VNE) [18].....	14
III.  Le centre d'éducation pour le traitement du diabète et des maladies de la nutrition : CETRADIMN .....	18
A.    Historique.....	18
B.    Présentation.....	18
MATERIEL.....	20
I.    Base de données médicales du CETRADIMN .....	20
A.    Mode de recueil des données : l'existant .....	20
B.    Base de travail .....	23
C.    Examen critique des données disponibles .....	23
D.    Evolution de la base informatique .....	29
II.   Données administratives.....	31
A.    Données d'activité.....	31
B.    Données du personnel.....	32
C.    Temps consacré à l'éducation du patient .....	32
D.    Données financières .....	32
METHODE .....	33
I.    Processus de traitement .....	33
A.    Constitution d'une table initiale.....	33

B.	Exploration et correction manuelle des champs.....	33
C.	Rétablissement du patient au centre d'un schéma relationnel.....	35
D.	Définition du diagnostic final.....	38
E.	Cycles exploratoires.....	38
II.	Moyens mis en œuvre .....	39
A.	Stockage des données.....	39
B.	Phase de modification de la base .....	40
C.	Phase d'interrogation : .....	41
RESULTATS.....		48
I.	Eléments médicaux.....	48
A.	Dénombrements .....	48
B.	Validité du dosage de l'hémoglobine glycosylée .....	95
C.	Indicateurs d'utilité .....	98
II.	Eléments administratifs.....	120
A.	Données d'activité.....	120
B.	Données du personnel.....	122
C.	Temps consacré à l'éducation du patient.....	123
D.	Données financières .....	125
DISCUSSION.....		126
I.	Analyse des résultats.....	126
A.	Activité du CETRADIMN .....	126
B.	Médecins traitants.....	127
C.	Présentation des patients : de la totalité de la base, de la file active et des diabétiques de la file active. ....	127
D.	Complications et morbidité.....	129
E.	Indicateurs d'utilité .....	133
II.	Un problème d'informatique médicale.....	147
A.	Les outils de la cohérence d'une base .....	147
B.	Système d'information et motivation des fournisseurs de données ....	150
C.	Activité ou patients ? Place du patient dans le schéma relationnel.....	151
III.	Education thérapeutique et financements .....	152
IV.	Modèle conceptuel de la prise en charge du patient diabétique.....	153
CONCLUSION .....		157
I.	Au sujet du recueil de données.....	157
II.	Au sujet de l'analyse de l'activité du CETRADIMN.....	157
III.	Au sujet des indicateurs d'utilité.....	158

Lexique des sigles.....	159
Bibliographie .....	161
Figures .....	168
Tableaux .....	173

## Remerciements

A notre Président de thèse,

**Monsieur le Professeur Jean-Louis SALOMEZ,**

*Professeur d'Epidémiologie, d'Economie de la Santé et de Prévention,  
Praticien Hospitalier – Chef du Service d'Epidémiologie et de Santé  
Publique, Hôpital Calmette - CHRU de Lille,  
Président de l'Observatoire Régional de la Santé Nord-Pas de Calais,  
Chevalier dans l'ordre des Palmes Académiques*

Vous nous faites le grand honneur de présider le jury de cette thèse. Soyez assuré de notre profond respect et de notre reconnaissance.

A notre Directeur de thèse,

**Monsieur le Docteur Jean-Louis GRENIER,**

*Praticien Hospitalier – Chef de Service du Centre d'Education pour le  
Traitement du Diabète et des Maladies de la Nutrition - CHG de Roubaix*

Vous nous avez fait l'honneur de diriger notre travail. Soyez assuré de notre profond respect et de notre reconnaissance.

Aux Membres du Jury :

**Monsieur le Professeur BEUSCART,**

*Professeur de Bio statistiques, d'Informatique Médicale et de Technologies de la Communication,  
Praticien Hospitalier – Chef de service du Département d'Information Médicale, Coordonnateur de la Clinique de Santé Publique - CHR de Lille,  
Directeur du Centre d'Etudes et de Recherche en Informatique Médicale,  
Officier dans l'Ordre des Palmes Académiques*

**Monsieur le Professeur Pierre FONTAINE,**

*Professeur des Universités  
Praticien hospitalier*

**Monsieur le Professeur GLANTENET,**

*Professeur associé de Médecine Générale  
Directeur du Département de Médecine Générale  
Coordonnateur-adjoint du Résidanat*

Vous nous faites l'honneur d'accepter de faire partie du jury de cette thèse. Soyez assurés de nos sincères remerciements et de notre reconnaissance.

**à Monsieur Cristian PREDA,**

*Maître de Conférence des universités,  
Enseignant-chercheur au Centre d'Etudes et de Recherche en  
Informatique Médicale - Université Lille 2*

Vous nous avez guidés vers les analyses harmoniques et les analyse de données longitudinales. Soyez assurés de nos sincères remerciements et de notre reconnaissance.

**A Messieurs Marc FRANZUK et Eric DEBUIGNE,**

*Informaticiens – CHG de Roubaix*

Vous nous avez aidés à comprendre le système d'information du CHG de Roubaix, vous avez extrait les données initiales utilisées par cette thèse. Soyez assurés de nos sincères remerciements et de notre reconnaissance.

à nos parents, pour l'éducation qu'ils nous ont donnée

à nos conjoints et amis, pour leur indéfectible soutien



# INTRODUCTION

## ***I. Propos général***

### ***A. L'objectif de la thèse***

Notre travail avait trois objectifs :

- Exploiter une base de données rassemblant des données cliniques et biologiques sur 17 années, mais présentant des problématiques liées à sa construction ancienne, et montrer les méthodes utilisées afin de préparer une table utilisable pour :
  - notre étude : constitution d'une base décisionnelle
  - la migration de la base transactionnelle du CETRADIMN
- Proposer un regard rétrospectif sur le rôle du CETRADIMN dans la prise en charge du diabète sur la population du versant Nord-Est de l'agglomération lilloise.
- Proposer et tester in situ des indicateurs rétrospectifs d'utilité d'une prise en charge, dans le cadre d'une structure de soins fonctionnant en réseau.

### ***B. Tour d'horizon de la bibliographie***

Notre recherche bibliographique ne nous a pas permis de retrouver d'étude rétrospective d'évaluation de structures de diabétologie fonctionnant en réseau.

#### ***1. L'étude ENTRED***

L'étude ENTRED (Echantillon National Témoin REprésentatif des personnes Diabétiques) évalue la prise en charge du diabète à partir des données de la CNAM TS. L'étude ENTRED [1] a deux objectifs :

- décrire, évaluer et surveiller l'état de santé des personnes diabétiques traitées, ainsi que les modalités et les résultats de leur prise en charge
- apporter des données représentatives qui permettront l'évaluation des réseaux diabète par comparaison des patients suivis dans le cadre des réseaux à ceux suivis en pratique usuelle [2].

L'étude ENTRED est constituée de 5 sous études [3] :

- un suivi de la consommation médicale et des coûts à partir des remboursements de l'Assurance maladie de 2001 à 2003, sur un échantillon de 10 000 personnes tirées au sort
- un questionnaire patient détaillé (20 pages [3bis] ) envoyé par voie postale, en mars 2002, aux personnes tirées au sort : 3 646 questionnaires et 894 questionnaires simplifiés sont analysables (soit 45,4% de retour)
- un questionnaire médecin traitant, complété par 2073 médecins des patients qui avaient répondu (soit 45,7% de retour), dont 1718 questionnaires détaillés
- une enquête hospitalière effectuée auprès des Départements d'information médicale (DIM) des établissements dans lesquels 314 patients tirés au sort avaient été hospitalisés (les résultats sont disponibles sur le site de l'Ancred [4])
- une enquête de mortalité en cours sur l'ensemble de l'échantillon : en novembre 2004, 828 personnes étaient décédées.

## **2. L'étude UKPDS**

L'étude UKPDS compare l'effet d'un traitement hypoglycémiant intensif à l'effet d'un traitement conventionnel sur le risque de complications microangiopathiques et macroangiopathiques dans le diabète de type 2. Il s'agit là d'une étude prospective randomisée incluant 3867 patients avec un suivi moyen de 10 années [5, 6, 7, 8, 9, 10].

## **3. L'étude de J. Pirart**

L'étude menée par J. Pirart [11, 12, 13] est une étude prospective longitudinale s'étendant sur plus de 25 ans et portant sur 4398 cas dont 2795 suivis sans interruption depuis le début de leur diabète.

## **II. Situation sanitaire de la population concernée**

### **A. Situation sanitaire de la population diabétique en France en 2001 [14]**

La prévalence du diabète traité en France métropolitaine en 1998 a été évaluée à 2.78% par Ricordeau et al. [15]

L'âge moyen des personnes diabétiques traitées en France en 2001 était de 64.6 ans. 22% d'entre elles étaient âgées de plus de 75 ans. 4.1% étaient âgées de moins de 40 ans.

La répartition homme/femme était de 51.8% en faveur des hommes.

La situation sociale peut être évaluée par le taux de personnes diabétiques bénéficiant de la CMU, qui était de 5.6%.

Le taux de prise en charge du ticket modérateur pour une affection ALD30 était de 77.2%. Le taux de personnes qui avaient bénéficié d'au moins 6 consultations chez leur médecin traitant était de 72.2%. Une hospitalisation ou plus était notée chez 27.5% des personnes. Et le taux de personnes ayant consulté un endocrinologue libéral était de 8.5%.

La prise en charge thérapeutique peut être évaluée par le tableau suivant :

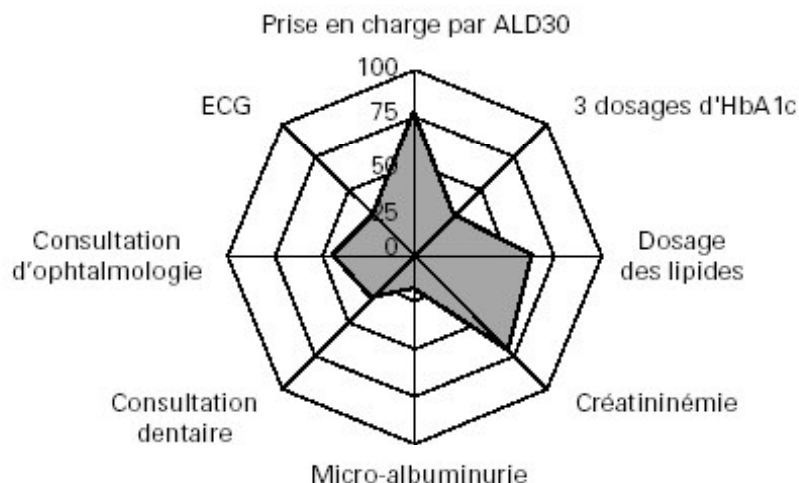
<b>Fréquence des remboursements de traitements antidiabétiques, hypolipémiants et à visée cardiovasculaire parmi les personnes diabétiques traitées. Entred 2001, n = 9 987.</b>		
<b>Remboursements</b>	<b>Fréquence (%)</b>	<b>[intervalle de confiance à 95 %]</b>
<b>Antidiabétiques oraux</b>	<b>80,0</b>	<b>[79,2-80,8]</b>
Sulfamides seuls	27,8	[26,9 - 28,7]
Biguanides seuls	15,1	[14,4 - 15,8]
Inhibiteurs des alpha-glucosidases seuls	2,8	[2,5 - 3,1]
Glinides seuls	1,5	[1,3 - 1,8]
Combinaison d'anti-diabétiques oraux sans insuline	32,8	[31,9 - 33,7]
<b>Insuline avec ou sans antidiabétique oral</b>	<b>20,0</b>	<b>[19,2-20,8]</b>
Insuline et anti-diabétique oral	5,9	[5,4 - 6,3]
Insuline seule	14,1	[13,5 - 14,9]
<b>Médicaments hypolipémiants</b>	<b>38,9</b>	<b>[38,0-39,9]</b>
<b>Médicaments à visée cardiovasculaire*</b>	<b>69,7</b>	<b>[68,8-70,6]</b>

\* Tout médicament à visée cardiovasculaire à l'exclusion des anti-agrégants plaquettaires et des vasodilatateurs.

Tableau 1 - traitement des patients diabétiques en France

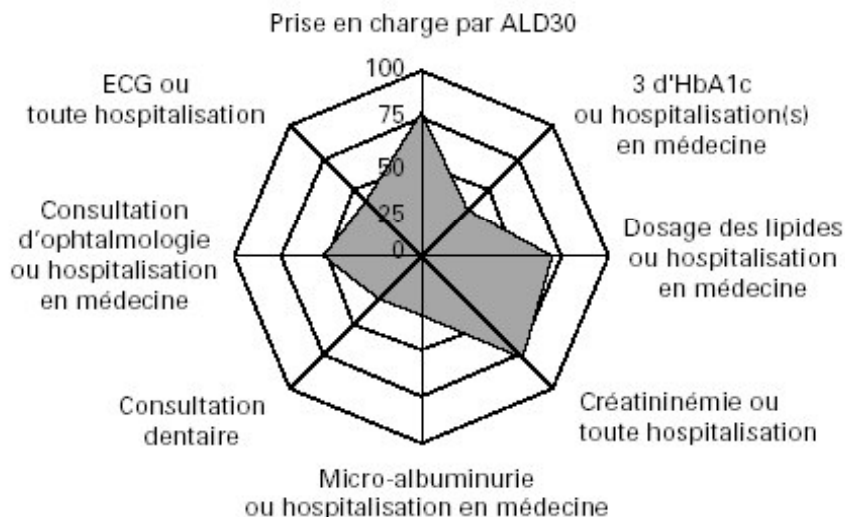
L'adéquation du suivi médical du diabète traité au suivi recommandé par la HAS (ex ANAES) peut être évaluée par le schéma suivant :

**Adéquation du suivi médical du diabète traité au suivi recommandé par l'Anaes, après prise en compte des données ambulatoires seules\*. Entred 2001, n = 9 987.**



\* Cette figure ne prend pas en compte les actes effectués dans un hôpital public, soumis au budget global. Il s'agit donc d'une sous-estimation des actes réellement effectués.

**Adéquation du suivi médical du diabète traité au suivi recommandé par l'Anaes, après prise en compte des données ambulatoires et hospitalières\*. Entred 2001, n = 9 987.**



\* Cette figure repose sur les hypothèses suivantes :

- lors d'une hospitalisation dans un service de médecine, un dosage d'HbA1c, de lipides, et de microalbuminurie, de même qu'une consultation d'ophtalmologie sont systématiquement effectués chez une personne diabétique,
- lors d'une hospitalisation, quel que soit le service concerné, un dosage de créatininémie et un ECG sont systématiquement effectués chez une personne diabétique.

Il s'agit d'une surestimation des actes réellement effectués.

**Figure 1 - adéquation du suivi des diabétiques aux recommandations de la HAS**

## ***B. Situation sanitaire de la population diabétique dans le Nord Pas de Calais***

Ricordeau et al. ont montré que, parmi les bénéficiaires du régime général de l'Assurance maladie, la prévalence du diabète traité variait suivant les régions. La prévalence du diabète traité en 1998 dans la région Nord Pas de Calais était de 2.73%. [15]

La population régionale de personnes diabétiques serait de 122 000 personnes, sur la base de la prévalence nationale [16]. Malheureusement, nous ne disposons pas du décompte par type de diabète.

L'âge moyen des personnes diabétiques traitées dans la région Nord-Pas de Calais en 2001 était de 62 ans, soit 3 points de moins que la population diabétique française.

L'intervalle de confiance à 95% de la répartition hommes/femmes était de 44-51%, soit une proportion plus faible d'hommes en comparaison avec la population diabétique française.

Dans la région Nord-Pas de Calais, le pourcentage de couverture par la CMU, indicateur de niveau socio-économique, était significativement plus élevé à 6-10% (IC95) comparativement à la population diabétique française.

Le taux de prise en charge du ticket modérateur pour une affection ALD30 était de 70-77% (IC95). Le taux de personnes qui avaient bénéficié d'au moins 6 consultations chez leur médecin traitant, était plus élevé à 89-93% (IC95). Par contre le taux de personnes ayant consulté un endocrinologue libéral était plus bas, à 4-7% (IC95). Une hospitalisation ou plus était notée chez 27-34% (IC95) des personnes.

La Haute Autorité en Santé (HAS, anciennement ANAES) préconise un suivi du niveau glycémique par le dosage de l'HbA1c tous les trois à quatre mois, car le contrôle de l'HbA1c est étroitement lié à la prévention des complications du diabète [17]. Or, seules 25-32% (IC95) des personnes, contre 29,0-30,8% (IC95) pour la population diabétique française, avaient bénéficié d'au moins trois remboursements d'HbA1c dans l'année.

Comparativement à la France métropolitaine, dans le Nord-Pas de Calais, au moins un dosage annuel de lipides (cholestérol total, triglycérides ou bilan lipidique), de créatininémie et de micro albuminurie avait été remboursé, respectivement, chez 64-70% des patients contre 62,4-64,3%, chez 75-81% des patients contre 70,5-72,2% et chez 16-22% des patients contre 16,9-18,4%. Ces chiffres sont donnés avec un intervalle de confiance de 95%. L'ANAES préconise un contrôle annuel. [17]

La réalisation d'au moins un ECG annuel (ou d'une consultation chez un cardiologue) avait été remboursée à 24-31% dans le Nord-Pas de Calais contre 28,9-30,7% en France métropolitaine. L'ANAES préconise un contrôle annuel. [17]

Enfin, comparativement à la France métropolitaine dans le Nord-Pas de Calais 38-45% contre 42,1-44,0% des personnes avaient reçu un remboursement pour une consultation d'ophtalmologie annuelle, sans qu'il soit possible de savoir si le fond d'œil avait été effectué comme le recommande l'ANAES [17].

### ***C. Situation sanitaire du Versant Nord-Est de la Métropole Lilloise (VNE) [18]***

D'après une étude réalisée dans le cadre de l'URCAM par le service médical de l'Assurance Maladie sur la ville de Tourcoing puis de Roubaix, le suivi des recommandations de la HAS se décline ainsi :

- 36% et 28,7% des patients avaient bénéficié d'au moins 3 HBA1C avec ou sans glycémie à jeun.
- 56% et 63,9% des patients avaient bénéficié d'un bilan lipidique
- 80% et 80.6% des patients avaient bénéficié d'au moins un dosage de la créatinine.
- 23,8% et 19,8% des patients avaient bénéficié d'au moins un dosage de micro albuminurie.
- 41,7 et 38,8% des patients avaient bénéficié d'une consultation avec un ophtalmologue en libéral
- 39,9% et 27,3% des patients avaient bénéficié d'une consultation avec un cardiologue en libéral.

La situation sociale du versant Nord Est de la métropole sur lequel est implanté le CETRADIMN, semble particulièrement défavorisée.

Sur la Figure 2, on représente le taux de chômage qui a augmenté de manière plus importante dans ce secteur que dans la région.

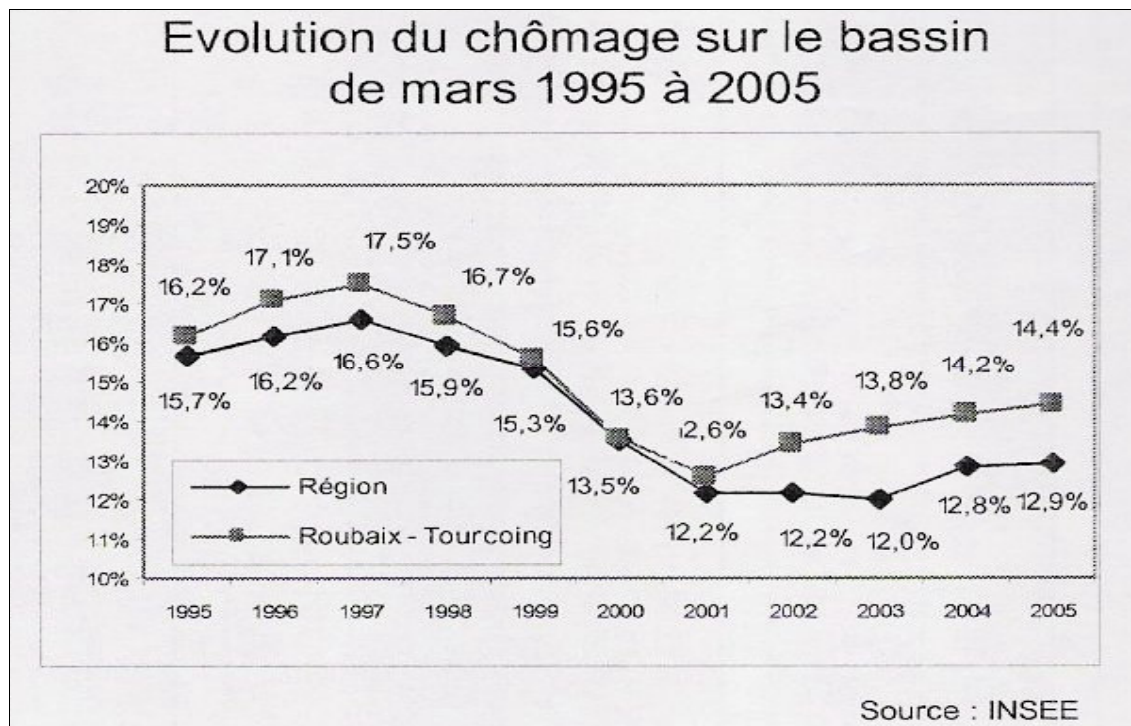


Figure 2 - évolution du chômage sur le territoire du Versant Nord-Est de la métropole lilloise (VNE)

La Figure 3 représente le taux de personnes bénéficiant de minima sociaux, qui sur le VNE est de 26.3 % contre 17.8 sur l'ensemble de la France. Ces données proviennent de la CAF sur l'année 2004.

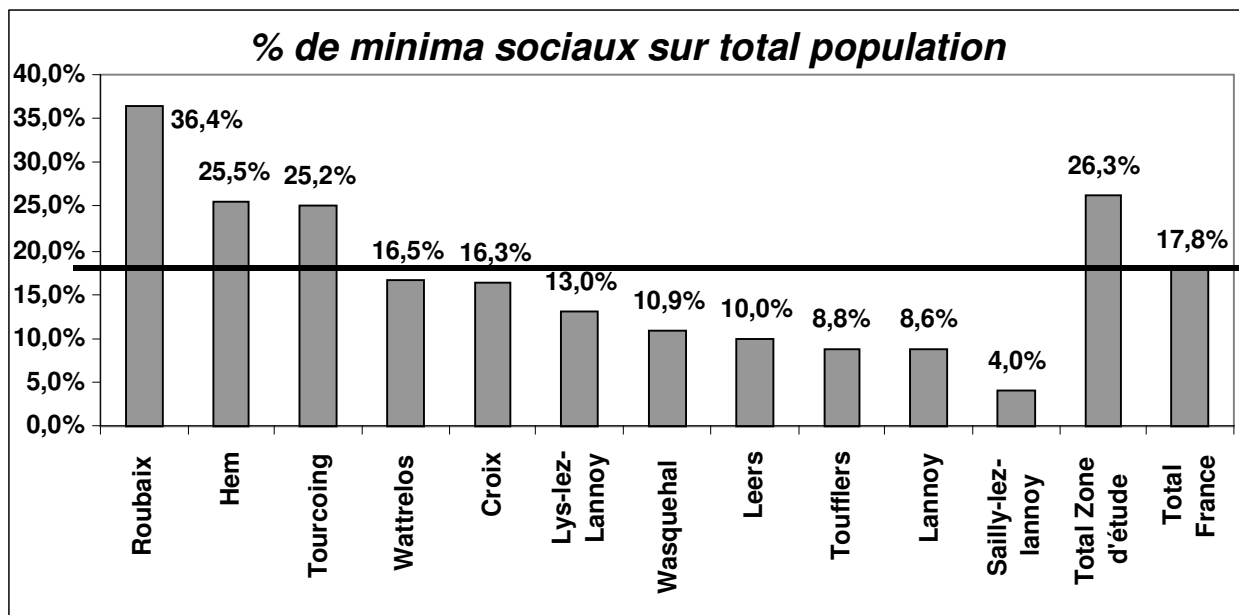


Figure 3 - proportion de minima sociaux sur le VNE (source CAF 2004)

La carte reproduite en Figure 4 présente un indice de défaveur sociale, établi par l'ORS en décembre 2004. On remarque que la zone du VNE présente un indice de défaveur sociale nettement plus élevé.

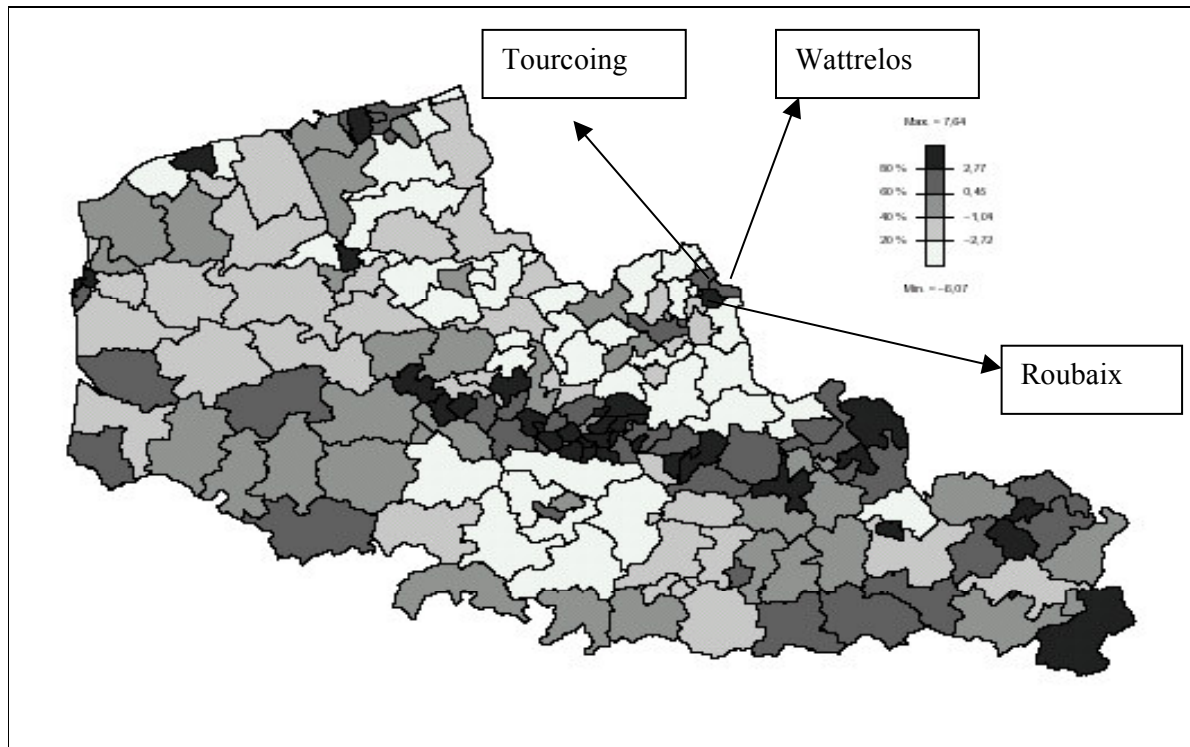


Figure 4 - Carte de l'indice de défaveur sociale sur 13 variables (ORS déc. 2004)

La lecture de cette Figure 4 doit se faire à la lumière de la densité (au Km<sup>2</sup> ) moyenne de ces villes, qui est exposée dans le Tableau 2.



Densité au km <sup>2</sup> Recensement 1999	Population	Superficie en kilomètres <sup>2</sup>	Densité au km <sup>2</sup>
Lannoy	1 726	0,18	9 589
Roubaix	96 984	13	7 331
Tourcoing	93 540	15	6 158
Croix	20 638	4	4 648
Lys-lez-Lannoy	13 018	3	3 993
Wattrelos	42 753	13	3 181
Mouvaux	13 177	4	3 160
Wasquehal	18 541	7	2 703
Hem	19 675	10	2 039
Leers	9 651	5	1 787
Toufflers	3 864	2	1 617
Neuville en Ferrain	9 527	6	1 542
Halluin	18 997	13	1513
Forest sur Marque	1 562	1	1488
Roncq	12 705	11	1 200
Werwicq Sud	4 288	5	842
Bondues	10 680	13	818
Comines	11 952	16	746
Linselles	7 876	12	673
Bousbecque	4 157	6	645
Willems	2 799	6	483
Sailly-lez-Lannoy	1 763	4	398
Deûlémont	1 461	10	147
Warneton	178	4	43
<b>Total Zone d'étude</b>	<b>421 512</b>	<b>183</b>	<b>2 301</b>
<b>Total Nord</b>	<b>2 555 020</b>	<b>5 743</b>	<b>445</b>
<b>Total Région</b>	<b>3 996 588</b>	<b>12 414</b>	<b>322</b>
<b>Total France</b>	<b>58 518 966</b>	<b>543 966</b>	<b>108</b>

Tableau 2 - densité au Km<sup>2</sup> des villes du versant Nord est (source INSEE)

La revue de la littérature permet d'avoir une idée plus précise de l'état sanitaire d'une population défavorisée. Ainsi nous savons que cet état social favorise les comportements à risque et le manque de compliance au suivi médical. Ainsi la défaveur sociale est un facteur de risque de :

- déséquilibre alimentaire [19, 20, 21, 22].
- surcharge pondérale [23, 24]
- recours aux soins retardés, et d'absence de recours à la prévention [21, 22]

On peut donc supposer que l'état sanitaire du versant Nord-Est de la métropole doit être fortement dégradé.

### **III. Le centre d'éducation pour le traitement du diabète et des maladies de la nutrition : CETRADIMN**

#### **A. Historique**

Ce Centre qui était en gestation depuis 1979 a commencé à fonctionner en Février 1982 et s'est individualisé totalement à partir d'Octobre 1983 sous la responsabilité du Docteur Jean-Louis GRENIER.

Sa construction a été guidé par une volonté d'un fonctionnement totalement en ambulatoire. De plus, l'approche du CETRADIMN était avant tout fondée sur l'éducation thérapeutique du patient.

#### **B. Présentation**

Comme nous l'avons mentionné précédemment, ce Centre était en gestation depuis 1979, il a commencé à fonctionner en 1983 mais s'est installé dans ses locaux actuels en 1985.

Il est situé juste en face du Centre Hospitalier de Roubaix dans une maison particulière, ce qui lui permet d'une part de ne pas être considéré comme un véritable service d'hospitalisation, et d'autre part de bénéficier du plateau technique du Centre Hospitalier Général (CHG) Victor Provo de Roubaix. C'est en fait à part entière le service de diabétologie endocrinologie et nutrition du CHG de Roubaix. Le CETRADIMN est unique en son genre en France puisqu'il ne fonctionne qu'en ambulatoire et en hospitalisation programmée à durée déterminée.

Les objectifs du CETRADIMN ont toujours été de favoriser la gestion par les patients de leur pathologie en améliorant leur compétence et leur autonomie à long terme ainsi que leur qualité de vie au quotidien. [25]

Le CETRADIMN a un rôle de soins centré sur l'éducation thérapeutique du patient, sans bien sûr négliger les aspects purement bio-médicaux ou de bilan notamment dans le diabète. A ce rôle de soins, s'associe un rôle de formation mais également un rôle d'information et de documentation tant pour les soignants que pour le grand public et ce, par le biais d'associations qui lui sont rattachées. L'activité de soins du CETRADIMN s'exprime sous trois modalités : consultation, hôpital de jour, hôpital de semaine ; l'hospitalisation classique n'existant pas dans ce Centre.

Qu'il s'agisse d'éducation polyvalente en hospitalisation de semaine ou de prise en charge en ambulatoire, la séquence éducative se prolonge par une confrontation médico-paramédicale qui permettra de proposer aux patients un projet de suivi éducatif à long terme. [26]

A cette activité de soins, comme nous le disions précédemment, s'associe une activité de formation qui intéresse tant les personnels spécialisés en diabétologie que les para-médicaux non spécialisés et les généralistes.

Dès 1985, un travail a été engagé sur la rentrée de données minimales non pas dans un but épidémiologique mais dans un but pratique car, ces données permettait d'une part d'avoir une facilitation dans la gestion des dossiers et ce tant au niveau bio-médical qu'au niveau de l'éducation thérapeutique du patient ; et d'autre part, de faciliter la transcription des données médicales notamment de laboratoire pour l'élaboration de courriers informatisés. C'est grâce à l'élaboration de cette base que dans un deuxième temps, le questionnaire a été simplifié pour aboutir à une masse de données conséquente qui fait l'objet de notre travail.

# MATERIEL

## ***I. Base de données médicales du CETRADIMN***

### ***A. Mode de recueil des données : l'existant***

#### ***1. Architecture générale***

##### ***a- Généralités***

En attendant l'avènement du projet NESSI (décrit plus bas), le Centre Hospitalier de Roubaix dispose d'une base de données SESAME. SESAME est un système de gestion de base de données (SGBD) implémenté en 1982 en Cobol/Assembleur/APL, tournant sur un serveur CICS d'IBM [27]. Chose surprenante pour l'époque, SESAME implémente une technologie orientée objet. Outre les fonctionnalités classiques d'un SGBD, SESAME propose des fonctionnalités répondant directement aux besoins de l'hôpital sans nécessiter de développement applicatif supplémentaire. Nous citerons pour exemples :

- la conception de formulaires complets de visualisation, saisie et modification des données
- l'invocation de ces formulaires à distance
- l'impression de courriers types complétés automatiquement par les valeurs des champs enregistrés

##### ***b- Utilisation des formulaires***

Les formulaires sont invoqués à distance, la connexion est réalisée via un navigateur internet doté d'un plugin, mais il ne s'agit pas de HTML : l'aspect rappelle les formulaires en interface Shell.

Les formulaires ainsi conçus sont pensés pour permettre une saisie rapide par des secrétaires. Une attention toute particulière est accordée à l'enchaînement des champs et pages de formulaires. Certains contrôles de saisie sont réalisables, mais ils restent toutefois sommaires et insuffisants.

En outre, toute modification d'un champ sur l'écran modifie immédiatement la base de données, sans même que l'utilisateur ne puisse manifester l'intention d'enregistrer ou d'annuler les modifications.

### ***c- Implantations locales***

SESAME a permis aux informaticiens du CH de Roubaix de développer pour chaque service demandeur une base de données propre originale et isolée, et les formulaires connexes. Comme nous l'énoncions plus haut, SESAME est un SGBD objet, ce n'est donc pas un SGBD relationnel. Cela signifie deux choses :

- un inconvénient notable : l'intégrité des données contenues dans des champs multivalués n'est pas assurée
- un avantage majeur : il est possible de définir librement une structure de données extensible à volonté, tel un arbre, originale et autonome pour chaque service, sans jamais devoir définir un lourd schéma relationnel.

C'est cette deuxième caractéristique qui a permis le développement rapide et adapté de bases de données par service. La base du CETRADIMN est une des bases mises en place par ce système.

Dans un deuxième temps, l'ajout du numéro d'identification permanent du patient (IPP) vise à pouvoir rapatrier pour un patient donné toutes ces « vues service », sans parler pour autant de mise en commun des informations. Chaque source de données locale reste totalement autonome. Nous développerons cette évolution plus loin.

## ***2. Mise en pratique au CETRADIMN***

### ***a- Cas d'utilisation***

Chaque fois qu'un nouveau patient se présente au CETRADIMN, les secrétaires saisissent certaines informations d'état civil et administratives concernant le patient et sa venue présente. Lors de la consultation, le médecin examinateur remplit un formulaire papier, qui est ensuite saisi par les secrétaires ou l'infirmière de nuit. La biologie réalisée à l'occasion de l'hospitalisation de jour est également saisie par les secrétaires.

Lorsque le patient est déjà venu (et est retrouvé), il est possible d'étendre son enregistrement initial avec un nouveau bloc (notion détaillée plus bas).

### ***b- Une utilisation au fil de l'eau / papier***

L'utilisation est avant tout conçue au fil de l'eau : l'intérêt de la saisie est de permettre d'imprimer, en fin de consultation ou d'hospitalisation de jour, des courriers type, des compte-rendus type ; l'information prend son sens une fois sur papier. On conçoit bien que, dans ce système, il importe peu de préserver l'identité exacte du patient, car le courrier, une fois imprimé, est facilement attribué à l'un ou l'autre des patients physiquement présents ce jour-là, et facilement classé dans le dossier papier correspondant. En revanche, jusqu'à

récemment, il n'était pas nécessaire d'être en mesure de récapituler l'antériorité du patient. C'est la nature de l'utilisation qui explique les difficultés rencontrées aujourd'hui, rétrospectivement, autour de l'identification du patient.

### c- Un système de blocs et de recopies automatiques

Par souci didactique, nous présenterons ci-après les données dans une table, avec une ligne par enregistrement et une colonne par champ, bien que la notion de table soit une notion purement relationnelle et donc hors sujet.

Supposons une table enregistrant 5 champs, un champ identité et 4 champs d'information, complétée ici pour 3 patients participant au total à 6 consultations. L'ajout de nouvelles consultations suppose deux choses : ressaisir sans erreur l'identité du patient, et renseigner les 4 champs d'information... quitte à recopier certains d'entre eux s'ils varient peu.

<u>Identité</u>	<u>Date consultation</u>	<u>Médecin traitant</u>	<u>Poids</u>	<u>HbA1c</u>
dupont	01/03/1998	Dr XXX	82	8.4
dupont	08/09/1998	Dr XXX	82	9.2
martin	02/07/2001	Dr YYY	70	10.1
léger	02/04/2003	Dr ZZZ	62	9.1
léger	23/01/2004	Dr ZZZ	62	8.9
léger	31/12/2004	Dr ZZZ	60	8.7

Tableau 3- recopie automatique et blocs - exemple 1/3

En réalité, comme on peut le supposer, l'identité du patient est composée de plusieurs champs, dont la saisie est plus ou moins longue. De plus, de nombreux champs, comme ici le médecin traitant, ou -c'est un choix fait au CETRADIMN- le poids, sont a priori plus ou moins constants. On conçoit alors combien une recopie pourrait alléger le travail des secrétaires.

Nos 4 champs d'information constituent un premier « bloc ». Il a été choisi de pouvoir répéter les 4 champs d'information autant de fois que nécessaire, en créant de nouveaux blocs *dans le même enregistrement*. Le fait de rester dans le même enregistrement permet à la fois de ne pas ressaisir l'identité, et de bénéficier de certaines recopies automatiques, comme ici le médecin traitant et le poids, ce qui n'interdit pas de les modifier manuellement.

<u>unique</u>	<u>Bloc 1</u>				<u>Bloc 2</u>				<u>Bloc 3</u>			
<u>Identité</u>	<u>Date Consult.</u>	<u>Médecin traitant</u>	<u>Poids</u>	<u>HbA1c</u>	<u>Date Consult.</u>	<u>Médecin traitant</u>	<u>Poids</u>	<u>HbA1c</u>	<u>Date Consult.</u>	<u>Médecin traitant</u>	<u>Poids</u>	<u>HbA1c</u>
dupont	01/03/1998	Dr XXX	82	8.4	08/09/1998	Dr XXX	82	9.2				
martin	02/07/2001	Dr YYY	70	10.1								
léger	02/04/2003	Dr ZZZ	62	9.1	23/01/2004	Dr ZZZ	62	8.9	31/12/2004	Dr ZZZ	60	8.7

Tableau 4 - recopie automatique et blocs - exemple 2/3

En admettant que chaque enregistrement ne puisse pas contenir plus de 3 blocs, si le patient Léger vient une quatrième fois le 20/04/2005, il faudra alors créer un deuxième enregistrement. On perd alors le bénéfice des recopies automatiques. Dans le cas du CETRADIMN, la limite est de 10 blocs par enregistrement.

<i>unique</i>	<i>Bloc 1</i>				<i>Bloc 2</i>				<i>Bloc 3</i>			
<i>Identité</i>	<i>Date Consult.</i>	<i>Médecin traitant</i>	<i>Poids</i>	<i>HbA1c</i>	<i>Date Consult.</i>	<i>Médecin traitant</i>	<i>Poids</i>	<i>HbA1c</i>	<i>Date Consult.</i>	<i>Médecin traitant</i>	<i>Poids</i>	<i>HbA1c</i>
dupont	01/03/1998	Dr XXX	82	8.4	08/09/1998	Dr XXX	82	9.2				
martin	02/07/2001	Dr YYY	70	10.1								
léger	02/04/2003	Dr ZZZ	62	9.1	23/01/2004	Dr ZZZ	62	8.9	31/12/2004	Dr ZZZ	60	8.7
léger	20/04/2005	Z-Z-Z	61	10.2								

**Tableau 5 - recopie automatique et blocs - exemple 3/3**

Enfin, en l'absence de schéma relationnel, il n'existe aucune garantie d'intégrité des champs multi-valués, tels l'identité et le médecin traitant par exemple, dont rien ne garantit que l'orthographe restera la même entre deux enregistrements. Finalement, dans une certaine mesure, les recopies entre blocs du même enregistrement constituaient une sorte de garde-fou.

Nous exposerons plus bas les difficultés liées à cette extension en blocs.

## ***B. Base de travail***

Nous travaillons sur une table unique extraite depuis la base, sous forme de fichier texte tabulé. Cette table contient toute l'activité du CETRADIMN jusqu'au 30 juin 2005. Cette table unique ne s'inscrit pas dans un schéma relationnel. Il s'agit d'un tableau dans lequel une ligne représente un événement (consultation, hôpital de jour, hôpital de semaine confondus). Les identifiants du patient sont éventuellement saisis à chaque fois, leur identité n'est donc pas garantie. De la même manière, toutes les informations susceptibles d'être permanentes pour un patient sont éventuellement saisis à chaque fois, et donc éventuellement variables.

La table fournie contient 70318 lignes (événements), initialement répartis entre 18886 patients différents.

## ***C. Examen critique des données disponibles***

### ***1. Les champs d'identité du patient***

Il était primordial de s'assurer l'intégrité de ces données pour pouvoir répondre à la seule question « combien de patients différents avons-nous ? ». La réponse suppose que, pour un patient donné, il n'existe qu'une seule identité, c'est-à-dire un seul quadruplet de valeurs {nom;prénom;date\_naissance;sexe}. Ce prérequis

était loin d'être acquis. En effet, dans la base initiale, il était fréquent d'observer une dizaine d'identités différentes pour le même patient. La correction des champs ne pouvait se faire que manuellement sur les 70318 enregistrements, en raison de la variabilité des causes d'erreur, dont certaines sont indétectables à l'échelle du champ mais le deviennent à l'échelle du quadruplet. Comme nous l'illustrerons plus bas, la tâche qui paraît simple mais fastidieuse à un cerveau humain ne pouvait être automatisée.

**Après correction, nous estimons que ces champs sont fiables.**

#### ***a- Erreurs communes aux champs textuels***

Nous citerons rapidement les problèmes suivants, qui entravent l'intégrité au sens où la variable semble présenter faussement de nouvelles modalités :

- variabilité liée aux caractères accentués
- variabilité liée aux caractères de ponctuation, en particulier, pour les noms composés, le choix de l'espace, du tiret, de la virgule ou même du vide pour séparer les mots
- variabilité liée aux erreurs de frappe, notamment les inversions de deux lettres frappées par deux doigts différents de la même main (« coquille »)
- variabilité liée à une volonté délibérée de s'économiser : troncature ou abréviation

Comme nous le verrons plus tard, toutes ces variations sont d'autant plus fréquentes en l'absence de schéma relationnel et en l'absence de motivation.

#### ***b- Le nom de famille***

Outre les erreurs sus-mentionnées, le nom de famille subit des variations toutes particulières liées au recrutement spécifique du centre :

- les noms d'origine maghrébine sont souvent écorchés car leur consonance n'est pas commune d'une part, et d'autre part car le patient ne sait parfois pas l'écrire, encore moins l'épeler
- les noms d'origine espagnole souffrent de leur surcomposition au grès de l'arbre généalogique, les secrétaires ont souvent tendance à les écourter ou à faire varier les séparateurs. A titre d'exemple fictif, le même nom pourrait s'écrire ainsi :

*Amaro Da Silva Ribeiros, Amaro-Dasilva-Ribeiros , Amaro-Dasilva, Da Silva Ribieros, Dasilva...*

Ces variations sont multipliées par les nombreuses coquilles, altérant parfois fortement la phonétique. Cet exemple illustre combien la correction ne pouvait être automatisée par des algorithmes du type SOUNDEX de Margaret K. Odell et Robert C. Russel [28, 29], ou METAPHONE de



Lawrence Philips [30, 27], pourtant implémentés nativement dans MySQL [31] ou disponibles dans PHP [32].

Pour ce qui est de la correction manuelle, on conçoit bien ici un problème supplémentaire : la même personne peut se retrouver aux deux extrémités d'une liste alphabétique...

- les noms d'origine belge posent le problème très particulier des motifs [consonne-voyelle double] ou des lettres H à proximité de consonnes, alors que le français est familier des consonnes doubles et voyelles simples. A titre d'exemple fictif mais très réaliste, le même nom pourrait s'écrire ainsi :

*Schamheere, Schaneere, Schanere, Schanhaire, Schanheere,  
Schanhere, Schannehere, Schannheere, Schaaneere, Schaaner,  
Schaanere, Schaanheere, Schaanhere*

- enfin, il est important de noter que pour les noms français dont l'homonymie est réputée, nous n'avons noté aucune erreur. On peut supposer que les secrétaires adoptent des stratégies de contrôle systématique :

« Lefèvre, avec ou sans B ? » « Dupont, avec un T ou un D ? »

### **c- le prénom**

Outre les variations précédentes, nous rappellerons que de nombreux prénoms sont transposés d'un alphabet non latin, et n'ont donc pas une orthographe « officielle », comme par exemple *Mohamed* et ses multiples variations. Les prénoms français quant à eux souffrent dans leurs versions composées : outre la variation de ponctuation, on observe des stratégies d'économie de la part des secrétaires. A titre d'exemple le même prénom peut s'écrire ainsi :

*Jean-Pierre, Jean Pierre, JPierre, j-pierre, je-pierre, jp, j.p., j-p ...*

### **d- la date de naissance**

La date de naissance est soumise à plusieurs phénomènes :

- de fréquentes inversions, comme par exemple 17/03/1971 et 13/07/1971
- les erreurs classiques de confusion de chiffres manuscrits, tels 1-7, 5-6, 8-9
- un nombre impressionnant de patients nés le 1<sup>er</sup> janvier, la secrétaire saisissant seulement l'année de naissance, ou parfois l'année et le mois
- plus rarement une inversion entre date de naissance et date d'examen, toute ambiguïté étant levée par la date d'ouverture du CETRADIMN et le classement chronologique de ces deux dates.

### **e- le sexe**

Le sexe est souvent erroné, qu'il s'agisse de prénoms mixtes (*Camille, Dominique, Aicha*), de prénoms étrangers peu connus (*Aimkhaltone, Zenobjia...*), ou de simples erreurs (*Madame Robert X*).

## **2. Les autres champs administratifs**

### **a- La date de venue, le type de venue**

La date de venue est essentielle, elle était la plupart du temps cohérente. Le nombre d'incomplets fut suffisamment faible pour permettre une recherche dans le dossier. Le type de venue (consultation, hôpital de jour, hôpital de semaine) était la plupart du temps renseigné.

### **b- le médecin traitant et le médecin examinateur**

Ces champs, peu complétés, subissent les mêmes distorsions que le nom du patient, un fort polymorphisme étant ajouté par les préfixes tels :

*Docteur, dr., dr, Monsieur le docteur, M. le docteur, M. le Dr...*

Nous avons utilisé ces champs à l'échelle du patient et non de l'événement, la ré-agrégation a donc permis d'augmenter la complétude. Le dé-doublonnage a suivi le même processus que pour les identités de patients.

## **3. Les champs d'épidémiologie**

### **a- Listes déroulantes**

La hiérarchie ci-dessous rappelle la dépendance réelle ou logique des champs, mais chaque item représente bien un champ et non une rubrique.

- [diagnostic]
  - [type de diabète] (si diagnostic=diabète)
  - [traitement du diabète] (en général, si diagnostic=diabète)
    - [diététique du diabète]
    - [antidiabétiques oraux] (famille de molécule)
      - [sulfamides] (nom de molécule)
      - [biguanides] (nom de molécule)
    - [insuline : type d'administration]
    - [insuline : schéma]
- [hypertension artérielle] (oui/non)
- [sang] (appareil de mesure de la glycémie)
- [urines] (nom de la bandelette réactive utilisée)
- [type de complications]

Ces champs sont cohérents au sens où les réponses possibles sont fixées par liste déroulante. En revanche, leur complétude est très faible à l'échelle de l'événement. Toutefois, à l'échelle du patient, elle est améliorée par ré-agrégation des événements l'impliquant.

D'autre part, nous noterons que les champs concernant le traitement consistent en une enquête épidémiologique *avant* l'événement (consultation, hospitalisation) et non à son terme. Cela signifie que, lors du premier événement, il s'agit du traitement *avant* prise en charge par le CETRADIMN. Lors du deuxième événement et les suivants, il s'agit du traitement *après* prise en charge par le CETRADIMN.

Enfin, il importe de noter que les listes proposées ne permettent pas de choisir explicitement « pas de traitement » par exemple. De ce fait, constatant un champ vide, il est impossible de savoir si ce vide a une signification ou est une omission.

#### ***b- Poids, taille, BMI***

Les poids, taille et BMI sont des champs proposés en saisie libre. Les erreurs sont nombreuses :

- caractères interdits (autres que [0123456789,.-])
- séparateur décimal variable et fantaisiste
- unités variables : taille en mètres, décimètres, centimètres, et même millimètres parfois...
- inversion de taille et poids, difficile à détecter chez des obèses lorsque la taille est en centimètres
- valeurs aberrantes inexplicables
- BMI très peu complété et souvent discordant avec le poids et la taille

De plus, afin de faciliter la saisie, un système de recopie automatique du poids et de la taille a été mis en place (cf. ante) : si un événement (bloc) existe dans le même enregistrement, alors la taille et le poids sont automatiquement recopiés dans le nouvel événement (nouveau bloc). Cette recopie empêche tout suivi du poids, il est même possible qu'elle ait diminué l'effort de renseignement. Enfin, elle produit des situations indécryptables : une première saisie est fautive par inversion de poids et taille, les saisies suivantes corrigent le poids mais pas la taille, on dispose alors simultanément de deux poids...

Après explorations et tentatives de correction, nous avons décidé de ne pas exploiter ces champs, jugés irrécupérables.

#### ***4. Les champs de biologie***

La table contient les champs suivants :

- |  |  |                     |
|--|--|---------------------|
| ➤ glycémie à jeun                          |  | ➤ LDL               |
| ➤ hémoglobine glyquée / glycosylée (HbA1c) |  | ➤ urée              |
| ➤ triglycérides                            |  | ➤ créatininémie     |
| ➤ HDL                                      |  | ➤ uricémie          |
|  |  | ➤ micro albuminurie |

Hormis de nombreux écarts typographiques très stéréotypés et donc rapidement corrigés, la biologie présente les caractéristiques suivantes :

- toutes les données sont cohérentes
- toute biologie réalisée en hôpital de jour a bien été saisie
- les unités employées sont standards et permanentes

Nous ajouterons toutefois que chaque passage de patient diabétique est systématiquement accompagné d'un dosage de l'HbA1c :

- en hôpital de jour, l'HbA1c est dosée sur place et saisie dans la base
- en consultation, l'HbA1c n'est jamais saisie dans la base :
  - parfois le patient la fait doser en ville avant de venir
  - dans les autres cas elle est dosée sur place

### **5. Les enregistrements (événements, lignes) manquants**

Il faut noter que certaines lignes manquent dans la table. Cet incident ne s'est produit que chez les patients venus onze fois ou plus avant 1999. Il peut manquer toutes les venues qui sont simultanément :

- postérieures à la neuvième visite
- antérieures à l'année 2000
- différentes de la dernière visite, qui est toujours préservée

Nous expliquions plus haut le parti pris d'autoriser la création de 1 à 10 blocs de champs dans un même enregistrement. Il est donc nécessaire, après le dixième bloc, de changer d'enregistrement. Dans le cas contraire, le bloc suivant remplace le dixième, et donc seuls les neuf premiers sont conservés. Au fil de l'eau, chaque nouveau bloc écrase le précédent et devient dixième à son tour. Une telle erreur de manipulation a été répétée pendant plusieurs années par l'infirmière de nuit chargée de la saisie. Cette erreur a débuté à une date inconnue, et a duré jusqu'en 1999, date de découverte du problème.

La figure ci-dessous illustre quelles peuvent être les conséquences :

- le patient 2 n'est venu que 3 fois, il n'y a donc pas de perte
- le patient 3 est venu plus de 10 fois, mais puisque son onzième passage est postérieur à 1999, il n'y a pas de perte
- pour le patient 1, nous avons perdu certaines données, encadrées

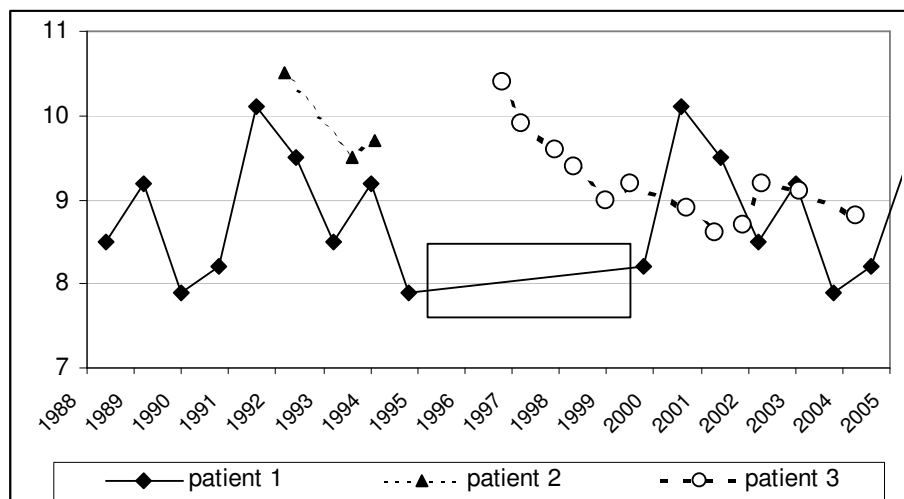


Figure 5 - illustration des pertes de données avant 1999 (ici avec l'HbA1c)

## D. Evolution de la base informatique

### 1. Interrogation centralisée des différentes bases service

Dès 2003, un rapatriement automatique et périodique a été mis en place, afin de cloner toutes les bases SESAME existantes dans une autre base qui sera utilisée par le nouveau système d'information : la base TAMPON, tournant sous le SGBD CACHÉ [33]. Toutefois, les applications service continuent pour l'instant de fonctionner sur la base SESAME. CACHÉ, tout comme SESAME, est un SGBD orienté objet. Il permet donc assez simplement de conserver des structures originales et hétéroclites pour chaque service.

Pour un patient donné, il est désormais possible d'accéder à tous ses dossiers services sur la base TAMPON, sous réserve d'identification correcte.

De plus, tous les patients créés de novo après 1995 sont dotés d'un identifiant permanent du patient (IPP), autrefois baptisé IPUM au CH de Roubaix. La cohérence de l'IPP est assurée par la confrontation au fichier tenu à jour par l'administrateur des IPP. Actuellement, sur les patients en base du Centre Hospitalier :

- 220 000 patients sont corrects
- 110 000 patients n'ont pas d'IPP (créés avant 1995)
- 2 200 patients ont un IPP incohérent

### 2. Le nouveau système d'information du CH de Roubaix

L'installation du nouveau système d'information du CH de Roubaix, citée dans le compte rendu d'accréditation de la HAS [34], a fait l'objet d'une vaste opération de

communication interne, sous le nom du projet NESSI, pour « nouveau système d'information ». Ce nouveau système d'information devrait permettre notamment :

- la tenue en ligne du dossier du patient, ou plutôt des dossiers services du patient, avec les courriers et compte-rendus divers
- un véritable circuit automatisé du médicament : prescription nominative en ligne, traçabilité du circuit
- le recueil et l'exploitation des données utilisées par le programme de médicalisation des systèmes d'information (PMSI) : codage des actes et diagnostics, groupage, simulation tarifaire
- le recueil des données administratives, l'interfaçage avec les applications existantes
- ...

L'appel d'offre a finalement été remporté par la société Siemens Health Services et son outil de gestion médico-administrative intégrée CLINICOM® , actuellement utilisé par de nombreux établissements [35]. CLINICOM s'appuie sur la base de données orientée objet CACHÉ, produit de la société Intersystems [33].

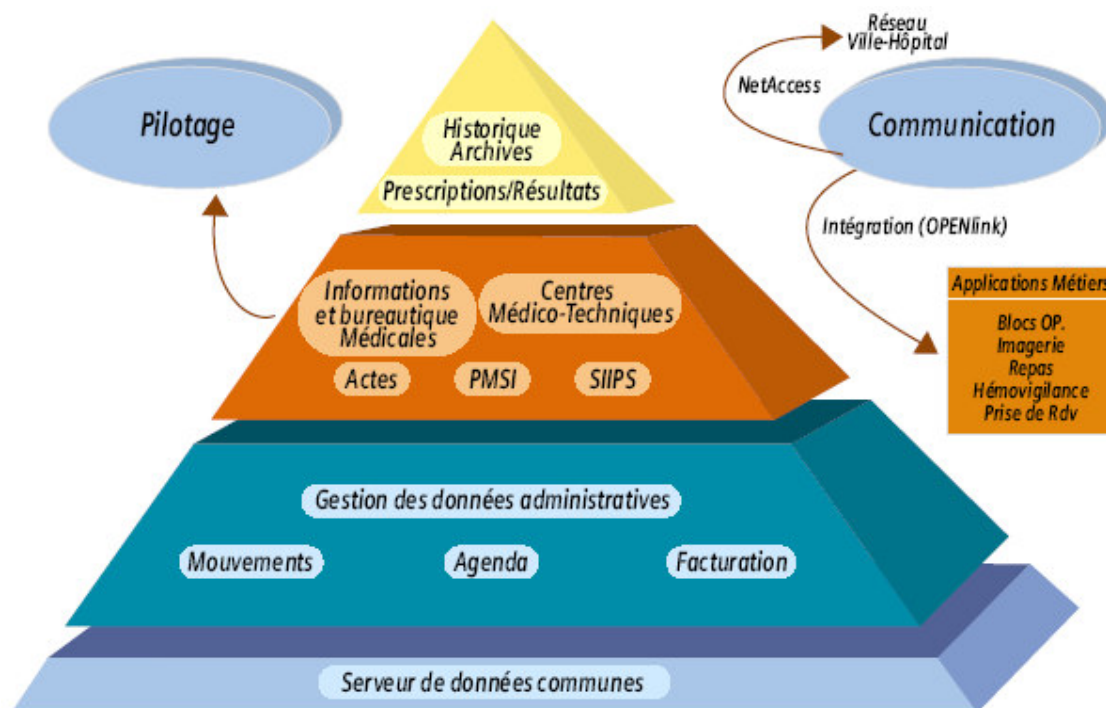


Figure 6 - architecture de CLINICOM, © Siemens Health Systems

Pour l'anecdote, CLINICOM a été pour l'occasion interfacé avec l'application d'aide au codage ROUBLAC, produite par le CH de Roubaix, et à laquelle les

utilisateurs étaient particulièrement attachés en raison de la rapidité des recherches, de la gestion performante des changements de version de la CCAM, de la gestion simple et décentralisée des catalogues d'actes [36, 37].

En se bornant à la tenue d'un dossier patient en ligne, un des enjeux majeurs est de traiter les nombreuses et hétéroclites bases propres à chaque service :

- rapatrier les dossiers afin de conserver toute l'antériorité
- ressusciter les formulaires dans de nouvelles applications qui ne devront si possible rien changer aux habitudes de recueil

### ***3. L'opération de migration***

Le 29 décembre 2005, l'image de la base de donnée SESAME, contenant toutes les bases services, sera rapatriée telle quelle sur la base EURÉKA du SGBD CACHÉ, par le protocole FTP. La migration devrait durer 30 heures et la mise en place des fichiers 48 heures.

Ce rapatriement permettra de regrouper tous les dossiers du patient, si celui-ci est doté d'un IPP et si cet IPP est cohérent. Sinon, il sera possible de forcer manuellement le rapatriement d'un dossier dont la propriété semble prouvée, car tous les dossiers non conformes resteront, comme actuellement, disponibles sur la base TAMPON.

Dans le même temps, les formulaires propres à chaque service ont été reproduits de manière similaire sous CACHÉ. Tout comme SESAME, CACHÉ autorise la création de formulaires adaptés à une base. Toutefois, la particularité de CACHÉ est d'utiliser le langage interprété côté serveur Caché Server Pages (CSP), qui permet de générer des formulaires complexes en HTML et JavaScript. Ces formulaires permettent de mettre en place des contrôles de saisie particulièrement performants qui, bien utilisés, devraient permettre d'améliorer la normalisation des saisies. Cependant, un soin tout particulier a été apporté à ne pas modifier les habitudes des services, en reproduisant la structure des anciens formulaires.

## ***II. Données administratives***

### ***A. Données d'activité***

Les données concernant l'activité du CETRADIMN nous ont été fournies par le service administratif du CETRADIMN sur la base du carnet de rendez-vous.

### ***B. Données du personnel.***

Les données concernant l'évolution des effectifs médicaux et paramédicaux nous ont été fournies par le Cadre Supérieur de Santé du CETRADIMN.

### ***C. Temps consacré à l'éducation du patient***

La présentation et le dénombrement des modalités d'éducation du patient ont été fournis par le CETRADIMN.

Les mesures chronométrées de temps consacré à l'éducation du patient ont été réalisées par le Cadre Supérieur de Santé du CETRADIMN.

### ***D. Données financières***

Les données financières nous ont été fournies par le service financier du CHG de Roubaix.



# METHODE

## ***I. Processus de traitement***

### ***A. Constitution d'une table initiale***

La table initiale est une table contenant une ligne par événement (consultation, hôpital de jour, hôpital de semaine), en format texte tabulé. Elle résulte d'une extraction réalisée depuis la base TAMPON par Eric Debuigne, informaticien au CH de Roubaix. Il s'agit du contenu de la base locale « questionnaire n°8 Dr Grenier ».

Cette extraction a rencontré les difficultés suivantes :

- les champs non renseignés prenaient pour valeur NULL, ce qui entraînait un décalage vers la gauche des champs suivants, dans les mauvaises colonnes
- les numéros des champs correspondant aux questions du questionnaire ont changé au fil du temps
- le système des blocs répétés paraissait opaque au départ (nous l'expliquons plus haut)

### ***B. Exploration et correction manuelle des champs***

#### ***1. Correction par colonne***

Dans un premier temps nous avons dû examiner la validité des champs un par un, en répondant aux questions suivantes :

- les caractères utilisés sont-ils légaux ?
- pour les champs multivalués : la réponse apportée entre-t-elle dans la liste de réponses attendues ?
- pour les champs numériques : la valeur, même anormale, reste-t-elle dans une fourchette crédible (en particulier pour les valeurs biologiques) ?
- si non, est-ce un problème d'unité, comment le corriger systématiquement ?

Chaque fois que nous avons trouvé des valeurs anormales :

- nous les avons corrigées lorsque cela était possible
- dans le cas contraire nous les avons supprimées
- enfin, lorsque le champ dans son entièreté semblait corrompu, nous avons refusé de traiter la variable

## **2. Cohérence entre colonnes**

La prise en compte de la cohérence inter-colonne intervient dans deux cas de figure :

- La recherche d'inversions systématiques de champs ; ce problème s'est posé dans deux cas :
  - Entre la date de naissance et la date de venue : il est alors simple de détecter et résoudre ces cas sans aucune ambiguïté
  - Entre la taille et le poids : comme expliqué plus haut, nous avons renoncé à traiter ces champs, en raison des trop nombreuses erreurs et de la recopie automatique
- La recherche d'incohérence entre des champs présentant une dépendance logique : par exemple [diagnostic] et [type de diabète] et [traitement du diabète], ou encore [traitement du diabète] et [antidiabétiques oraux] et [sulfamides]. Nous n'avons pas modifié la base. Nous avons pris note de ces divergences et, lorsque nous cherchions par exemple à compter le nombre de diabétiques traités par ADO (Anti-Diabétiques Oraux), nous recherchions cette information dans tous les champs évoquant plus ou moins directement la question. L'incohérence probable est donc prise en compte dans la requête d'interrogation.

## **3. Cohérence entre lignes**

La cohérence inter-lignes répond à la question suivante : « lorsqu'on regroupe tous les événements d'un même patient, un même champ est-il cohérent tout au long du suivi du patient ? ». Plusieurs cas de figure peuvent faire répondre par la négative :

- La relativement faible exhaustivité de renseignement des champs dits épidémiologiques : on conçoit volontiers en revanche que l'agrégation de ces champs au niveau du patient résout très volontiers le problème pour les patients venus au moins deux fois.
- Dans le même ordre d'idée, une précision plus ou moins grande entre les réponses provoque une variation qui reste simple à interpréter
- La modification de la valeur dans le cadre de l'évolution naturelle de la maladie : DNID (diabète non insulino-dépendant), puis DIT (diabète insulino-traité)
- Une réelle contradiction entre les valeurs : ce cas est plus rare mais parfois observé.

Ce problème très complexe concerne en réalité essentiellement les champs en rapport direct avec le diagnostic du patient. Après examen minutieux, nous avons

décidé de ne pas modifier la base en soi, en revanche nous avons pris en compte ce problème en créant de nouveaux champs globaux à l'échelle du patient (et non plus de l'événement). A titre d'exemple, le calcul du diagnostic final est développé plus bas.

## ***C. Rétablissement du patient au centre d'un schéma relationnel***

### ***1. principe***

Très schématiquement, le schéma relationnel est fondé sur des tables liées entre elles par des champs clef (clefs primaires *primary keys*, et clefs étrangères *foreign keys*). Sans entrer dans les détails des formes normales et de la forme de Boyce-Scott, on peut dire en résumé que si un champ est multi-valué, il doit engendrer une nouvelle table. Ainsi, l'identité du patient est multi-valuée car plusieurs enregistrements contiennent la même identité. Cette répétition pose un problème d'intégrité des données. Alors, une table des patients doit être créée. Nous reprenons à titre didactique l'exemple précédent, nous remplaçons par souci de réalisme l'identité par un doublet {nom, prénom} (en réalité il s'agit d'un quadruplet) (Tableau 6 - exemple de normalisation - contenu des tables). Nous présentons successivement le contenu de la table initiale, puis le même contenu remis dans un schéma relationnel.

**Sans normalisation :**

## Table événements

<i>Identité :nom</i>	<i>Identité :prénom</i>	<i>Date consultation</i>	<i>Médecin traitant</i>	<i>Poids</i>	<i>HbA1c</i>
Dupont	Jean	01/03/1998	Dr XXX	82	8.4
Dupont	Jean	08/09/1998	Dr XXX	82	9.2
Martin	Pierre	02/07/2001	Dr ZZZ	70	10.1
Léger	Paul	02/04/2003	Dr ZZZ	62	9.1
Léger	Paul	23/01/2004	Dr ZZZ	62	8.9
Léger	Paul	31/12/2004	Dr ZZZ	60	8.7

**Avec normalisation :**

## Table Patients

<i>id</i>	<i>nom</i>	<i>prénom</i>
101	Dupont	Jean
102	Martin	Pierre
103	Léger	Paul

## Table événements

<i>id</i>	<i>id Patient</i>	<i>Date consultation</i>	<i>id médecin</i>	<i>Poids</i>	<i>HbA1c</i>
1	101	01/03/1998	501	82	8.4
2	101	08/09/1998	501	82	9.2
3	102	02/07/2001	502	70	10.1
4	103	02/04/2003	502	62	9.1
5	103	23/01/2004	502	62	8.9
6	103	31/12/2004	502	60	8.7

## Table médecin

<i>id</i>	<i>Médecin traitant</i>
501	Dr XXX
502	Dr ZZZ

Tableau 6 - exemple de normalisation - contenu des tables

Le schéma relationnel des tables se formalise alors comme suit (Tableau 7 - exemple de normalisation - structure des tables). Les clefs primaires sont indiquées en gras, les clefs étrangères sont indiquées en souligné. A titre d'exemple, on peut dire que la jointure entre la table *Patients* et la table *événements* est réalisée par la clef primaire *Patients.id* et la clef étrangère *événements.id\_Patient*, dans une relation 1-∞.

### **Sans normalisation :**

Table événements	
Identité_nom	Varchar(255)
Identité_prénom	Varchar(255)
Date_consultation	Date
Médecin_traitant	Varchar(255)
Poids	Float
HbA1c	Float

### **Avec normalisation :**

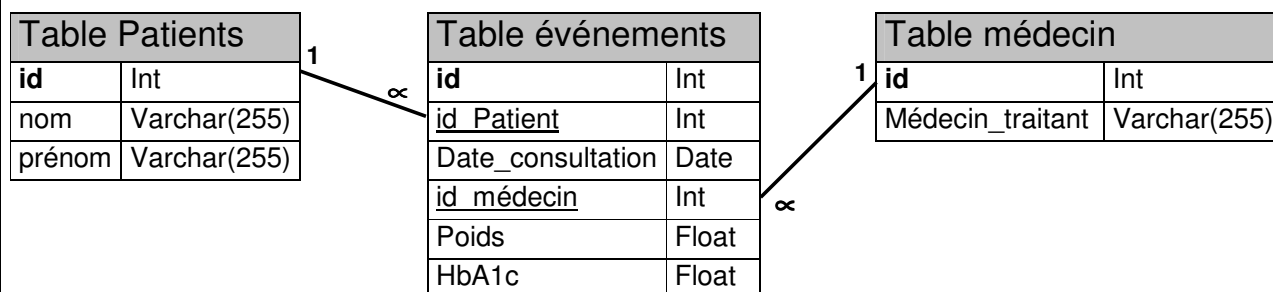


Tableau 7 - exemple de normalisation - structure des tables

Incontournable depuis la fin du XX<sup>e</sup> siècle, seul le schéma relationnel permet de garantir l'intégrité des données, au prix d'une complexification des schémas de bases de données, et de jointures ralentissant l'extraction de l'information.

## **2. identité du patient**

Seuls les patients postérieurs à 1995 ont un numéro d'IPP. Pour les autres, l'identité est constituée par le quadruplet {nom, prénom, date de naissance, sexe}. Comme nous l'avons expliqué plus haut, le polymorphisme de chaque champ provoque un fort polymorphisme du quadruplet complet. Ainsi un même patient peut-il exister dans la base sous une dizaine d'identités différentes. Compte tenu de la complexité du problème, il était impossible d'automatiser le processus de correction. Nous avons donc réalisé cette correction manuellement.

Ainsi, sur 70318 enregistrements, la correction manuelle fait passer le nombre de patients de 18886 à 16772, soit -11%.

## **3. champs agrégés**

La construction d'une table de patients est l'occasion de créer des champs de portée plus globale. Nous disposons d'une information à l'échelle de l'événement, nous souhaitons disposer d'une information à l'échelle du patient. Nous citons à titre d'exemple quelques champs « agrégés », sans détailler leur mode de calcul :

- Identité (nom, prénom, date de naissance, sexe)
- Nombre d'événements, nombre d'événements depuis 2001
- Dates du premier et du dernier événement
- Liste des médecins examinateurs
- Liste des médecins traitants
- Premier médecin traitant
- Diagnostics agrégés, type de diabète agrégé, complications agrégées, traitements agrégés...
- Diagnostic final (nous expliquons sa formation plus bas)
- Glycémie à jeun : première valeur, valeur la plus haute
- Hémoglobine glyquée : première valeur, valeur la plus haute
- Valeurs les plus péjoratives de divers items biologiques
- ...

### ***D. Définition du diagnostic final***

La détermination du diagnostic final est un sujet épineux pour les raisons suivantes :

- L'information diagnostic figure en réalité sur plusieurs champs, y compris les champs thérapeutiques
- Les champs épidémiologiques sont assez peu remplis à l'échelle de l'événement
- Un patient peut être pris en charge pour un motif médical puis pour un autre
- Un patient peut présenter simultanément plusieurs pathologies
- La pathologie du patient peut s'aggraver

Nous avons dû arrêter un ordre précis de détermination du diagnostic, cet ordre est lié à l'orientation de la présente étude uniquement : nous nous intéressons aux patients diabétiques. Si un patient présente les critères du diagnostic K, alors K devient le diagnostic final du patient. Sinon, nous testons ensuite le diagnostic K+1. Enfin, les patients non classés échoient dans la catégorie « non renseigné ».

- |   |                             |
|---|-----------------------------|
| 01- diabète type 1                      | 06- surcharge pondérale     |
| 02- diabète type 2                      | 07- pathologie thyroïdienne |
| 03- diabète gestationnel                | 08- dislipidémie            |
| 04- diabète autre                       | 09- autres                  |
| 05- intolérance aux hydrates de carbone | 10- non renseigné           |

### ***E. Cycles exploratoires***

Sur les fondements précédemment évoqués, nous avons entrepris l'exploration de la base, en commençant par les dénombrements les plus simples. Chaque

dénombrement a été confronté aux autres sources de données et à l'intuition des médecins du CETRADIMN pour tester sa cohérence. Ces échanges nous ont permis de déceler des incohérences de la base ou des erreurs de calcul. Les explorations ont également suscité de nouvelles interrogations. De la sorte, la fourniture de résultats s'apparente réellement à une succession de cycles exploratoires de plus en plus précis.

## **II. Moyens mis en œuvre**

### **A. Stockage des données**

#### **1. Pourquoi une base de données**

L'utilisation d'une base de donnée suivant un schéma relationnel a répondu aux contraintes suivantes [31] :

- **complexité des calculs** : certains calculs, en particulier lors de la constitution de la table des patients, n'auraient pu se faire autrement
- **taille de la base** : en raison du nombre élevé d'enregistrements (70318), un simple tableur n'aurait pu suffire
- **Schéma relationnel** : la base de départ était une succession d'enregistrements (consultations, hospitalisations). Il nous était indispensable de redonner sa place au patient, en créant une table des événements, mais aussi une table des patients, contenant des données agrégées, et mise en relation avec la première par une liaison de type {1-n}. La simple question « combien avons-nous de patients » ne put trouver de réponse avant un lourd travail informatique.

Nous avons choisi le serveur de bases de données MySQL [38]. Contrairement à nombre de ses concurrents, MySQL est un SGBD relationnel gratuit, qui implémente certaines fonctionnalités SQL indispensables à des traitements complexes [31] :

- fonctions classiques de traitement des nombres, chaînes et dates
- requêtes imbriquées en profondeur illimitée  
SELECT .. FROM (SELECT ... FROM ...) as table\_temporaire ...
- la fonction atypique GROUP\_CONCAT qui, utilisée avec GROUP BY, permet de concaténer les différentes valeurs d'un même champ dans un même groupe
- interrogation en ligne de commande, simplicité de l'administration, rapidité de l'exécution, bonne optimisation des index.

Pour mémoire, l'installation de MySQL sous Windows peut être facilitée par l'installation d'EasyPHP [39], son interrogation exploratoire peut être simplifiée grâce à un client graphique léger : MySQL Query Browser [40].

## **2. Réalisation**

Nous avons simplement incorporé le fichier à plat extrait de la base SESAME dans une table reproduisant la même succession de champs. Il s'agit donc d'une table d'événements (hôpital de jour, hôpital de semaine, consultation), et non de patients.

### **B. Phase de modification de la base**

#### **1. Pourquoi des scripts SQL en modification**

Schématiquement, les modifications de la base furent de trois ordres :

- correction des champs
- calcul de champs supplémentaires
- calcul d'une table agrégée au niveau du patient

Sur une base de taille réduite, il est habituel de réaliser ces modifications à la main via une interface graphique (GUI). Cela n'était pourtant pas souhaitable. L'utilisation de scripts SQL a répondu aux contraintes suivantes :

- **traçabilité des corrections** : pour prévenir toute erreur, il était indispensable de pouvoir consulter tout l'historique des modifications
- **réversibilité des corrections**
- **reconstruction de la base** : toutes les modifications apportées devaient pouvoir s'exécuter de nouveau d'un simple clic, en cas de perte ou de changement de stratégie concernant une modification en amont.

Chaque modification s'appuie sur les modifications précédentes, il fallait trouver un système respectant une chronologie d'exécution. Ainsi, si dans un premier temps nous transformons tous les « j-jacques » en « jean-jacques », la modification de la date de naissance de « j-jacques dubois » doit nécessairement en tenir compte. Les requêtes modificatrices doivent donc être exécutées *en cascade*.

## **2. Réalisation**

Les requêtes de modifications, dont le nombre approche de 2000, ont été enregistrées au fur et à mesure dans des fichiers texte SQL. Chaque fichier texte a été associé à un fichier batch permettant le lancement de toutes les requêtes d'un simple double-clic. Cela concerne tant les requêtes de correction des champs, que les requêtes de calcul et d'agrégation des données du patient.



Les requêtes de contrôle de cohérence, une centaine, qui ont guidé nos corrections, n'ont pas été sauvegardées.

### ***C. Phase d'interrogation :***

#### ***1. Pourquoi des scripts SQL en interrogation***

L'interrogation faisait au départ appel à un script SQL exécuté en ligne de commande, tout comme pour la phase de modification de la base. Nous avons opté pour un script SQL stocké intégralement pour les raisons suivantes :

- Dans un ensemble aussi complexe de requêtes, il arrive souvent que l'investigateur change de point de vue, par choix scientifique, ou bien par la découverte d'une erreur. Ce changement de point de vue peut retentir sur de nombreux calculs.
  - Il devient alors nécessaire de tout recalculer. De ce fait, il est totalement impensable de ré-extraire manuellement toutes les données en cours d'analyse, ce serait une perte de temps énorme et une source d'erreurs supplémentaires
  - La deuxième conséquence est qu'il est indispensable de vérifier de nombreuses fois le choix adopté dans tel ou tel calcul. La traçabilité des données est un vœu intangible, qui ne peut être satisfait en utilisant les logiciels classiques GUI (Graphic User Interface)
- A tout moment nous étions susceptibles de découvrir une nouvelle erreur non corrigée dans la base de données, il fallait alors pouvoir corriger la base et relancer automatiquement tous les états.

#### ***2. Pourquoi un programme spécifique d'interrogation en PHP***

Toutefois, certaines limites sont apparues rapidement avec les « simples » scripts SQL :

- La mise en page était tout à fait insuffisante, altérant la compréhension et l'exploitation de résultats de plus en plus nombreux :
  - Il était difficile d'insérer des commentaires méthodologiques
  - Il était impossible d'intégrer nos nombreuses requêtes dans une arborescence logique d'investigation
  - De manière générale, il était impossible de produire autre chose que des tableaux mis à la suite les uns des autres
- Nous ne pouvions pas afficher au même endroit le code source SQL, pour contrôle visuel

- Lors d'une erreur de syntaxe SQL, l'ensemble du script s'interrompait, entraînant une perte de temps ; le débogage était laconique, requérant des investigations supplémentaires.

Nous avons donc du programmer pour l'occasion une application résolvant tous ces problèmes. Nous la présentons ci-après.

### **3. Réalisation : notre application**

#### **a- Principe, matériel**

Nous avons programmé un analyseur de *templates* et de requêtes SQL entremêlées. Il s'agit d'un script programmé en PHP 5 (Php Hypertext Preprocessor, [41]), langage de programmation orienté-objet très utilisé sur Internet. PHP est un langage de scripts interprétés, c'est à dire que les programmes ne sont pas compilés, ils sont donc lisibles à l'œil nu et modifiables à volonté, et sont interprétés à chaque exécution par le moteur ZEND de PHP lui-même. Cette exécution se fait le plus souvent dans un contexte client-serveur. Ici, nous avons choisi d'exécuter ce script en ligne de commande grâce au CLI (Command Line Interpreter [32]) de PHP, ce qui nous a affranchi de toute contrainte architecturale. Pour simplifier l'exécution, un petit programme batch permet un lancement automatique en un simple double-clic.

Pour que le programme fonctionne, il faut naturellement que le serveur MySQL soit en cours de fonctionnement. En revanche, aucun serveur Apache/PHP n'est nécessaire.

#### **b- Utilisation**

Lors de son lancement, le programme lit un fichier source contenant les requêtes, les exécute, et fournit un fichier de sortie mis en forme, directement consultable. Cette exécution ne fait appel à aucune interaction et fournit un fichier de sortie très rapidement, ce qui est appréciable au vu du nombre d'exécutions successives réalisées.

Le fichier source utilisé est un fichier contenant à la fois des éléments de mise en forme en HTML (HyperText Markup Language, [42]), et des requêtes SQL brutes insérées au beau milieu du code HTML. Le programme sera capable de les extraire, de les exécuter, et d'insérer à la place le résultat de la requête.

Le fichier de sortie généré est un fichier HTML, donc visualisable agréablement à l'aide d'un navigateur web, permettant en outre une exploration interactive grâce au plan automatique agrémenté de liens hypertexte.

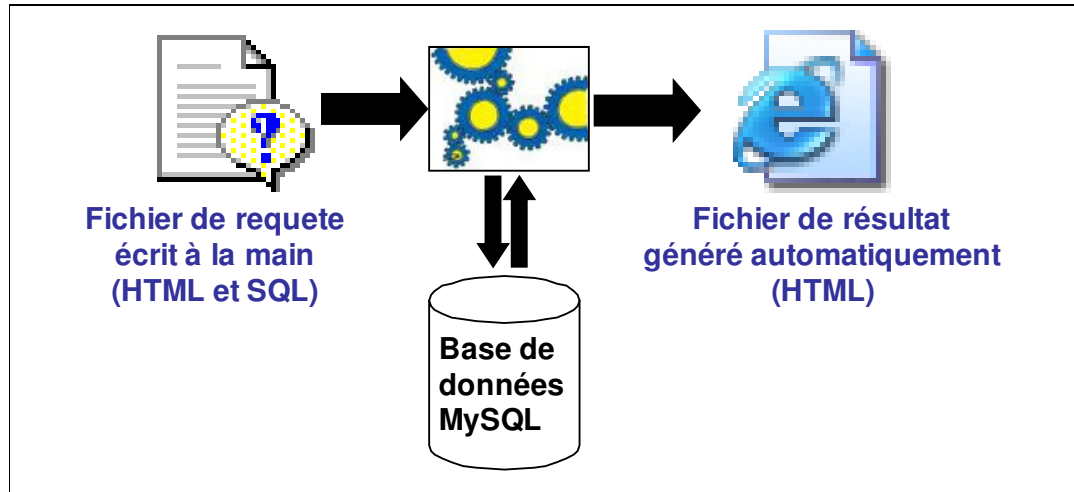


Figure 7 - Fonctionnement de l'application

### ***c- Actions détaillées***

Les actions suivantes sont automatiquement réalisées :

- Lecture du fichier source
  - extraction des requêtes SQL (balises `<q_*></q_*>`). Puis pour chaque requête :
    - affichage du code source SQL pour contrôle visuel
    - exécution de la requête (4 modalités selon le code SQL)
    - affichage des erreurs le cas échéant
    - affichage du résultat de la requête, avec mise en forme variable selon le cas (4 modalités)
    - éventuellement sauvegarde du résultat dans un nouveau fichier en texte brut tabulé
  - extraction de tous les titres HTML, de niveau 1 à 6, entre les balises `<h1></h1>` à `<h6></h6>`
    - calcul d'un plan automatique
    - insertion du plan à la place de la balise `<plan />`
    - insertion automatique de liens hypertexte, permettant une navigation aisée
- Ecriture du fichier HTML de sortie
- Ce fichier de sortie peut ensuite être exploré avec un navigateur Internet, importé dans Excel ou SPSS. Les fichiers en texte brut peuvent également être importés dans Excel ou SPSS.

Nous avons défini 4 modalités de traitement des requêtes selon le cas :

- La requête vide :
  - elle est exécutée sans retourner de résultat
  - exemple : définition d'une constante, modification de données, insertion dans une table...
  - elle est déclarée entre les balises `<q_vider>` et `</q_vider>`.
- La requête brute :
  - elle est exécutée, et son résultat est affiché en texte brut
  - exemple : calcul d'un nombre unique à insérer dans une phrase
  - elle est déclarée entre les balises `<q_brute>` et `</q_brute>`.
- La requête tabulée :
  - elle est exécutée et son résultat est affiché sous forme de tableau mis en forme
  - c'est le cas de la plupart des requêtes
  - elle est déclarée entre les balises `<q_table>` et `</q_table>`.
- La requête fichier :
  - elle est exécutée et son résultat est enregistré dans un fichier texte brut tabulé, sans affichage.
  - nous avons ressenti le besoin de l'ajouter car nos requêtes tabulées retournaient des résultats lisibles et parlants, tels des moyennes. En revanche, pour traiter ces données dans un logiciel de statistique, il était nécessaire de lui fournir des données plus brutes et volumineuses, dont l'affichage n'aurait pas été informatif.
  - elle est déclarée entre les balises `<q_fichier>` et `</q_fichier>`

#### ***d- Démonstration***

Un simple double-clic sur le fichier *execution.bat* génère automatiquement le fichier *sortie.htm* après lecture du fichier *source.htm*, comme l'illustre la capture d'écran ci-dessous. En outre, sur les 128 requêtes, 22 génèrent des sorties en texte tabulé pour traitement externe.

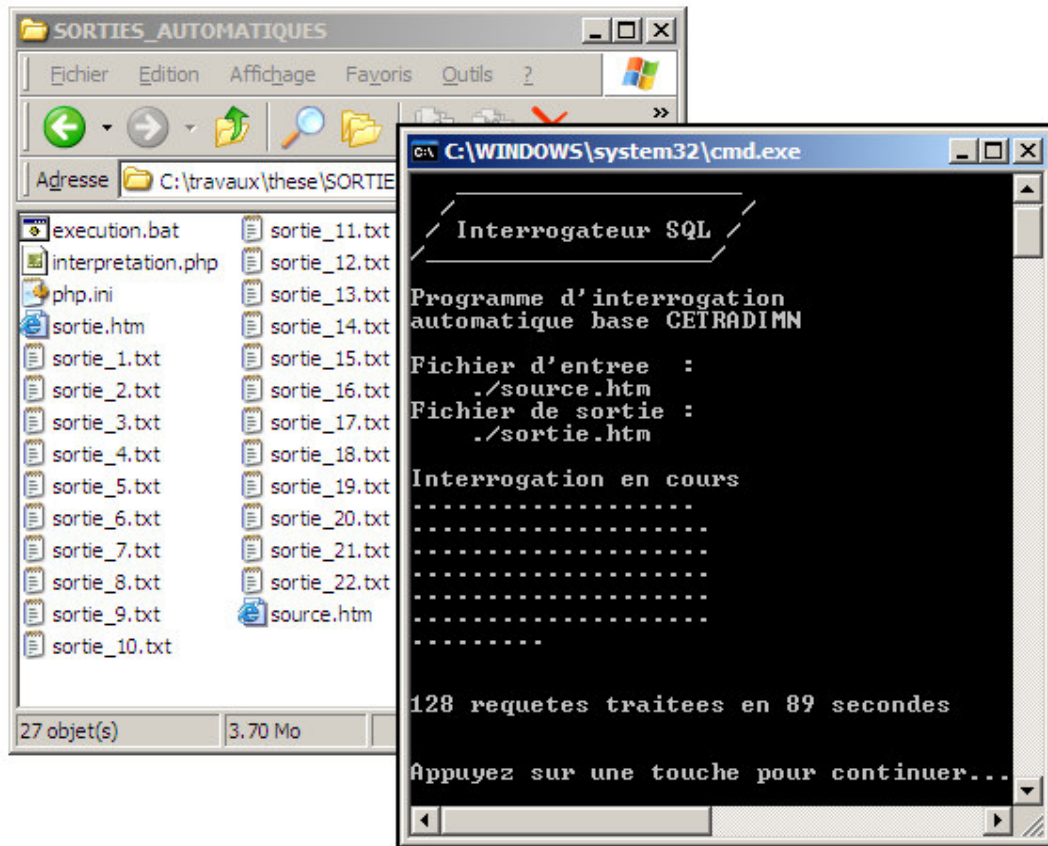


Figure 8 - Notre application : lancement

La page ci-après présente, à l'aide d'un exemple très simple, un fichier source comprenant 3 requêtes, puis le fichier résultat après analyse par notre programme. Pour ces deux fichiers, nous présentons à gauche le code source et à droite l'aspect dans un navigateur web.

```

<plan />

<h1>Essai</h1>

<h2>Premier essai : Requête vide</h2>
<q_vide>set @variable = ( select count(*) from patients ) ;</q_vide>

<h2>Deuxième essai : Requête brute</h2>
<q_brute>
select "la table contient " , round(count(*)/@variable*100,2) , "% de patients diabétiques type 2"
from patients where diag_final like "02%" ;
</q_brute>

<h2>Troisième essai : Requête tabulée</h2>
<q_table>
select diag_final, count(*) as nombre,
      round(count(*)/@variable*100,2) as pourcentage
from patients group by diag_final order by diag_final ;
</q_table>

<h1>Conclusion</h1>
<p>Tout fonctionne !</p>

```

**Essai**

**Premier essai : Requête vide**  
set @variable = ( select count(\*) from patients ) ;

**Deuxième essai : Requête brute**  
select "la table contient " ,  
round(count(\*)/@variable\*100,2) , "% de patients  
diabétiques type 2" from patients where diag\_final  
like "02%"

**Troisième essai : Requête tabulée**  
select diag\_final, count(\*) as nombre,  
round(count(\*)/@variable\*100,2) as pourcentage  
from patients group by diag\_final order by diag\_final

**Conclusion**  
Tout fonctionne !

Code source en entrée



Affichage en entrée

Code source en sortie

Affichage en sortie

```

<a name=top />
<p style="font-size:11px;margin-top:0;margin-bottom:0;text-indent:0;"><a href="#1"> Essai</a></p><p style="font-size:11px;margin-top:0;margin-bottom:0;text-indent:30;"><a href="#2"> Premier essai : Requête vide</a></p><p style="font-size:11px;margin-top:0;margin-bottom:0;text-indent:30;"><a href="#3"> Deuxième essai : Requête brute</a></p><p style="font-size:11px;margin-top:0;margin-bottom:0;text-indent:30;"><a href="#4"> Troisième essai : Requête tabulée</a></p><p style="font-size:11px;margin-top:0;margin-bottom:0;text-indent:0;"><a href="#5"> Conclusion</a></p>

<a name="1"></a><h1>Essai</h1>
<a name="2"></a><h2>Premier essai : Requête vide</h2>
<pre class=vide>
set @variable = ( select count(*) from patients ) ;
</pre>

<a name="3"></a><h2>Deuxième essai : Requête brute</h2>
<pre class=brute>
select &quot;la table contient &quot; , round(count(*)/@variable*100,2) , &quot;% de patients
diabétiques type 2&quot;
from patients
where diag_final like &quot;02%&quot;
</pre>
<p>la table contient 32.53 % de patients diabétiques type 2 </p>

<a name="4"></a><h2>Troisième essai : Requête tabulée</h2>
<pre class=table>
select diag_final, count(*) as nombre, round(count(*)/@variable*100,2) as pourcentage
from patients
group by diag_final order by diag_final
</pre>
<table border=1><tr><th>diag_final</th><th>nombre</th><th>pourcentage</th></tr><tr><td>01- diabete type 1</td><td>897</td><td>5.35</td></tr><tr><td>02- diabete type 2</td><td>5456</td><td>32.53</td></tr><tr><td>03- diabete gestationnel</td><td>100</td><td>0.60</td></tr><tr><td>04- diabete autre</td><td>196</td><td>1.17</td></tr><tr><td>05- intolerance aux hydrates de carbone</td><td>211</td><td>1.26</td></tr><tr><td>06- surcharge pondérale</td><td>4499</td><td>26.82</td></tr><tr><td>07- pathologie thyroïdienne</td><td>2044</td><td>12.19</td></tr><tr><td>08- dislipidemie</td><td>220</td><td>1.31</td></tr><tr><td>09- autres</td><td>1229</td><td>7.33</td></tr><tr><td>10- non renseigné</td><td>1920</td><td>11.45</td></tr></table>

<a name="5"></a><h1>Conclusion</h1>
<p>Tout fonctionne !</p>

```

[Essai](#)  
[Premier essai : Requête vide](#)  
[Deuxième essai : Requête brute](#)  
[Troisième essai : Requête tabulée](#)  
[Conclusion](#)

**Essai**

**Premier essai : Requête vide**  
set @variable = ( select count(\*) from patients ) ;

**Deuxième essai : Requête brute**  
select "la table contient " , round(count(\*)/@variable\*100,2) , "% de patients  
diabétiques type 2"  
from patients  
where diag\_final like "02%";  
la table contient 32.53 % de patients diabétiques type 2

**Troisième essai : Requête tabulée**  
select diag\_final, count(\*) as nombre, round(count(\*)/@variable\*100,2) as  
pourcentage  
from patients  
group by diag\_final order by diag\_final

diag_final	nombre	pourcentage
01- diabete type 1	897	5.35
02- diabete type 2	5456	32.53
03- diabete gestationnel	100	0.60
04- diabete autre	196	1.17
05- intolerance aux hydrates de carbone	211	1.26
06- surcharge pondérale	4499	26.82
07- pathologie thyroïdienne	2044	12.19
08- dislipidemie	220	1.31
09- autres	1229	7.33
10- non renseigné	1920	11.45

**Conclusion**  
Tout fonctionne !

Figure 9 - Notre application : parsing et exécution automatisée de requêtes d'interrogation

### **e- Résultats**

Notre application a permis :

- l'édition de tableaux agrégés en HTML : ces tableaux ont guidé notre réflexion, puis certains d'entre eux ont été importés dans Excel pour la réalisation des graphiques.
- L'édition de tables de données en texte brut : ces tables ont pu être importées dans SPSS, logiciel de statistiques, afin de permettre l'édition de graphiques plus complexes (box plots) et de statistiques (distribution, régressions). Certaines de ces tables ont été importées dans d'autres logiciels de statistiques pour réaliser des explorations ou traitements plus complexes, comme le logiciel d'Analyse Harmonique programmé par Cristian Preda, Maître de Conférences à l'Université Lille 2

# RESULTATS

## ***I. Eléments médicaux***

### ***A. Dénombrements***

#### ***1. Vue globale***

L'analyse de la base de données du CETRADIMN s'est faite sur l'ensemble des événements enregistrés jusqu'au 31/06/2005. Nous avons étudié une *file active* qui comprend l'ensemble des patients ayant au moins un événement (consultation, hôpital de jour, hôpital de semaine) compris dans la période du 1/01/2001 au 31/06/2005.

Nous avons dénombré :

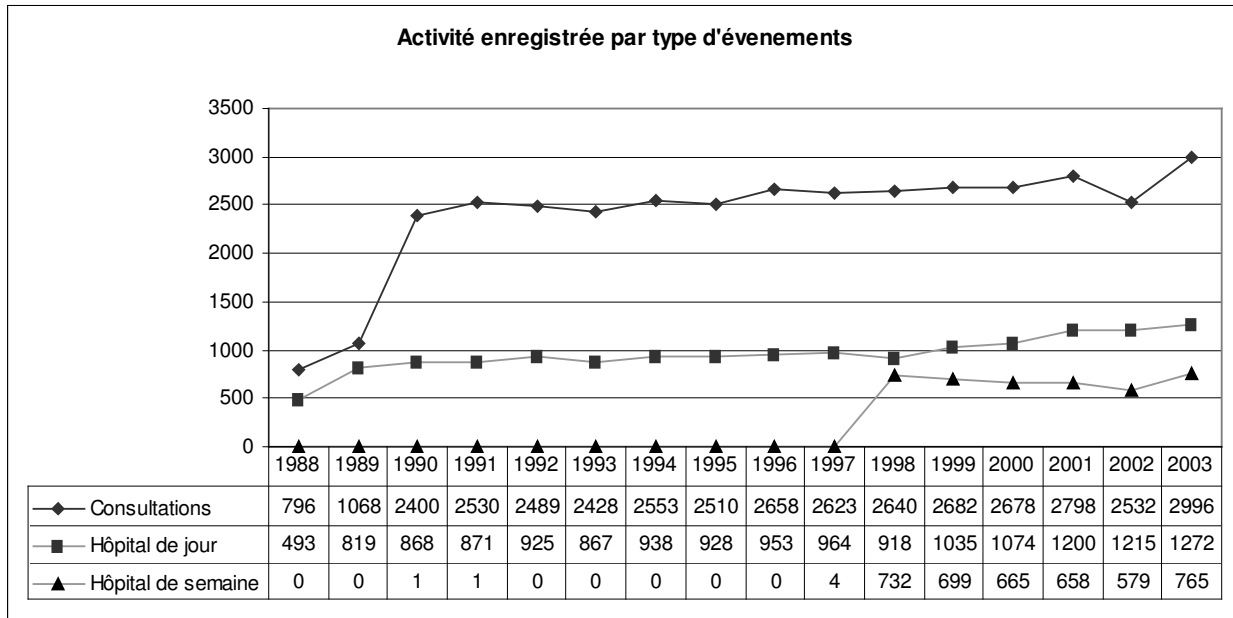
- 16772 patients distincts dont 6765 inclus dans la file active
- 70318 événements dont 23090 entre 2001 et 2005

#### ***2. Activité du service***

##### ***a- Activité enregistrée***

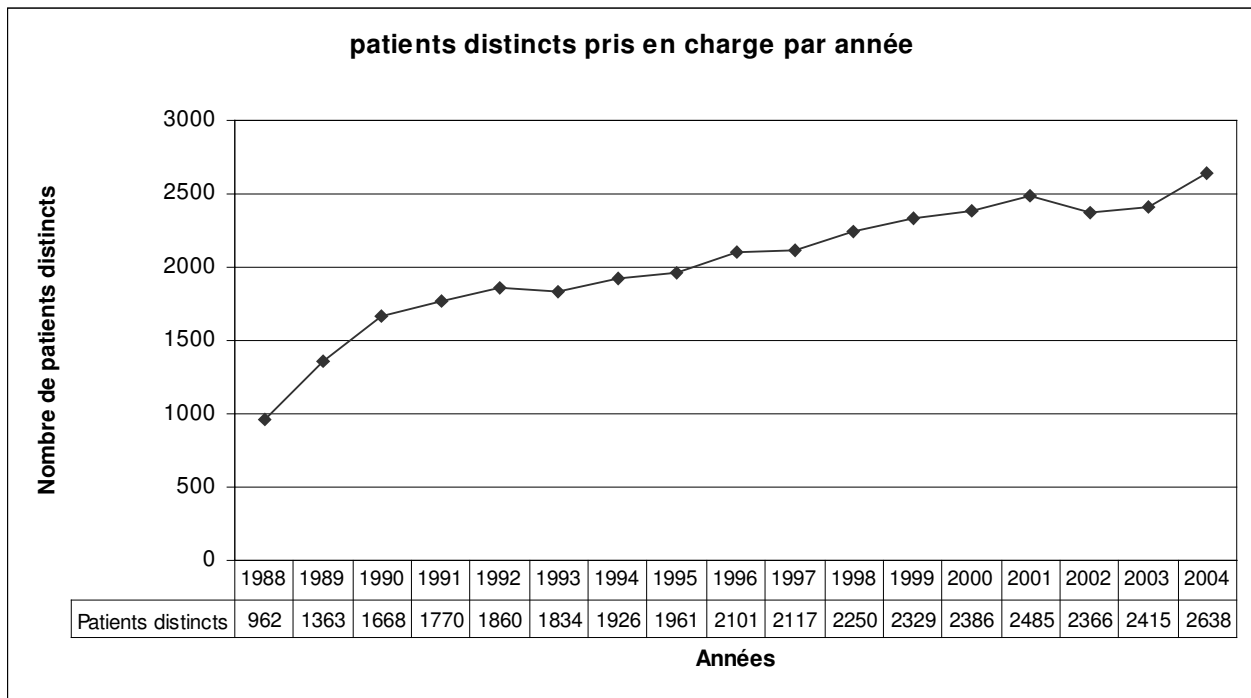
Nous notons une augmentation constante du nombre de consultations et d'hospitalisations de jour, avec une montée en charge sur deux ans (Figure 10).





**Figure 10 - Evolution du nombre de consultations, d'hôpital de jour et d'hôpital de semaine.**

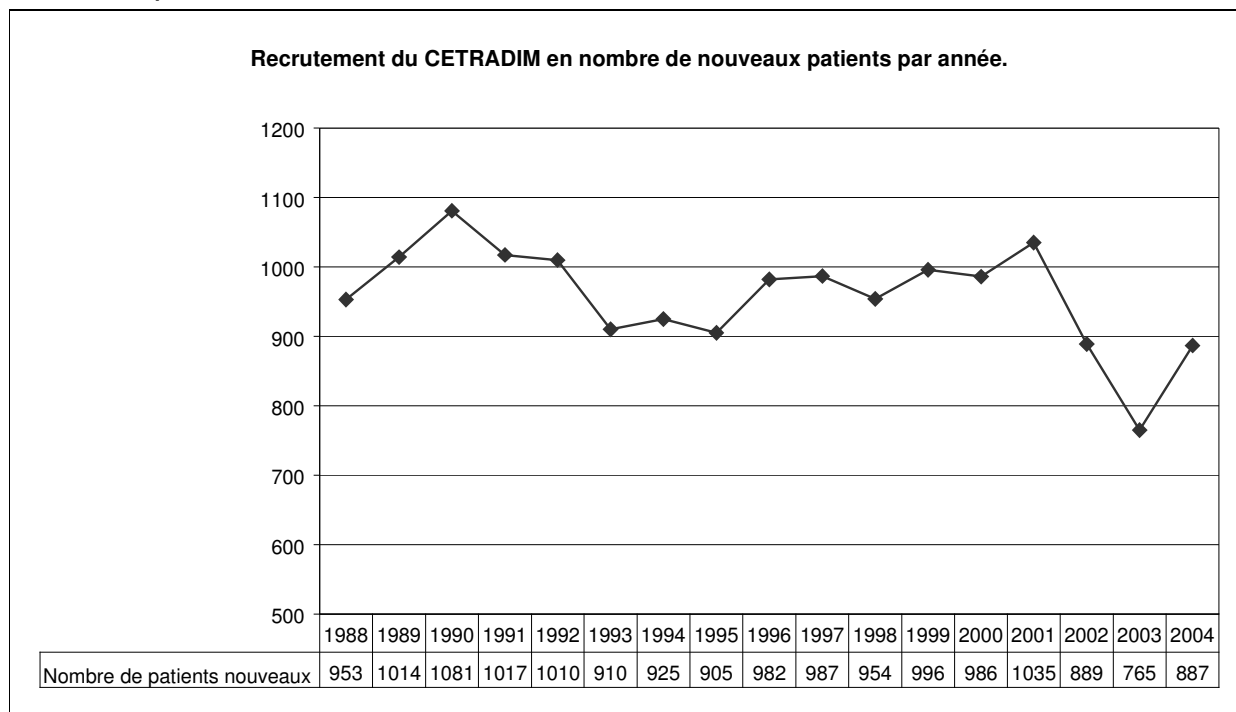
Sur la Figure 11, nous montrons l'évolution du nombre de patients distincts pris en charge chaque année. Il s'agit donc des différents patients venus au moins une fois dans l'année. L'augmentation est constante, témoignant d'un renouvellement à solde positif de la patientèle.



**Figure 11 - Evolution du nombre de patients distincts pris en charge chaque année.**

### ***b- Recrutement de patients***

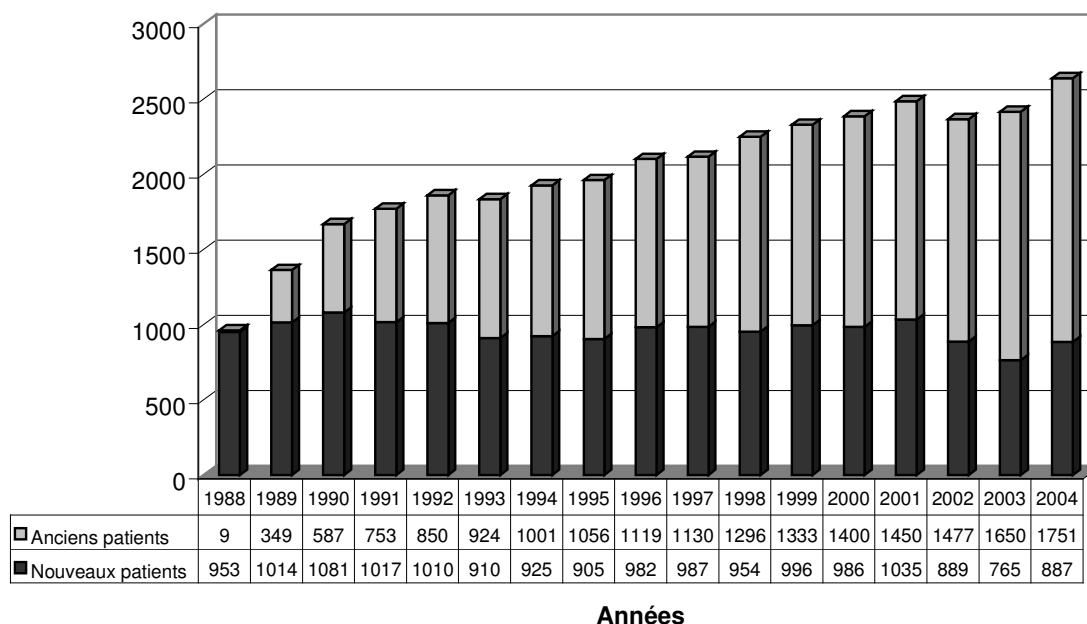
Sur la Figure 12, nous représentons le nombre de nouveaux patients pris en charge chaque année. Il s'agit donc des patients n'ayant pas d'antériorité. On note des phases ascendantes et descendantes.



**Figure 12 - Nombre de nouveaux patients chaque année**

Sur la Figure 13, nous représentons sous forme d'histogrammes empilés, la part des nouveaux patients et des anciens patients sur le nombre de patients distincts vus chaque année. Nous avons synthétisé les données de la Figure 11 et de la Figure 12.

### Evolution de la part de nouveaux patients chaque année



**Figure 13 - Part des nouveaux patients sur les patients distincts vus chaque année**

Sur la Figure 14 nous avons décomposé le nombre de nouveaux patients par diagnostic afin de voir l'évolution du recrutement selon le diagnostic au cours des années.

Le recrutement des nouveaux patients présentant un diabète de type 1 est marqué par une diminution avec stabilisation autour de 30 patients par an, après un recrutement important dans les premières années du CETRADIMN.

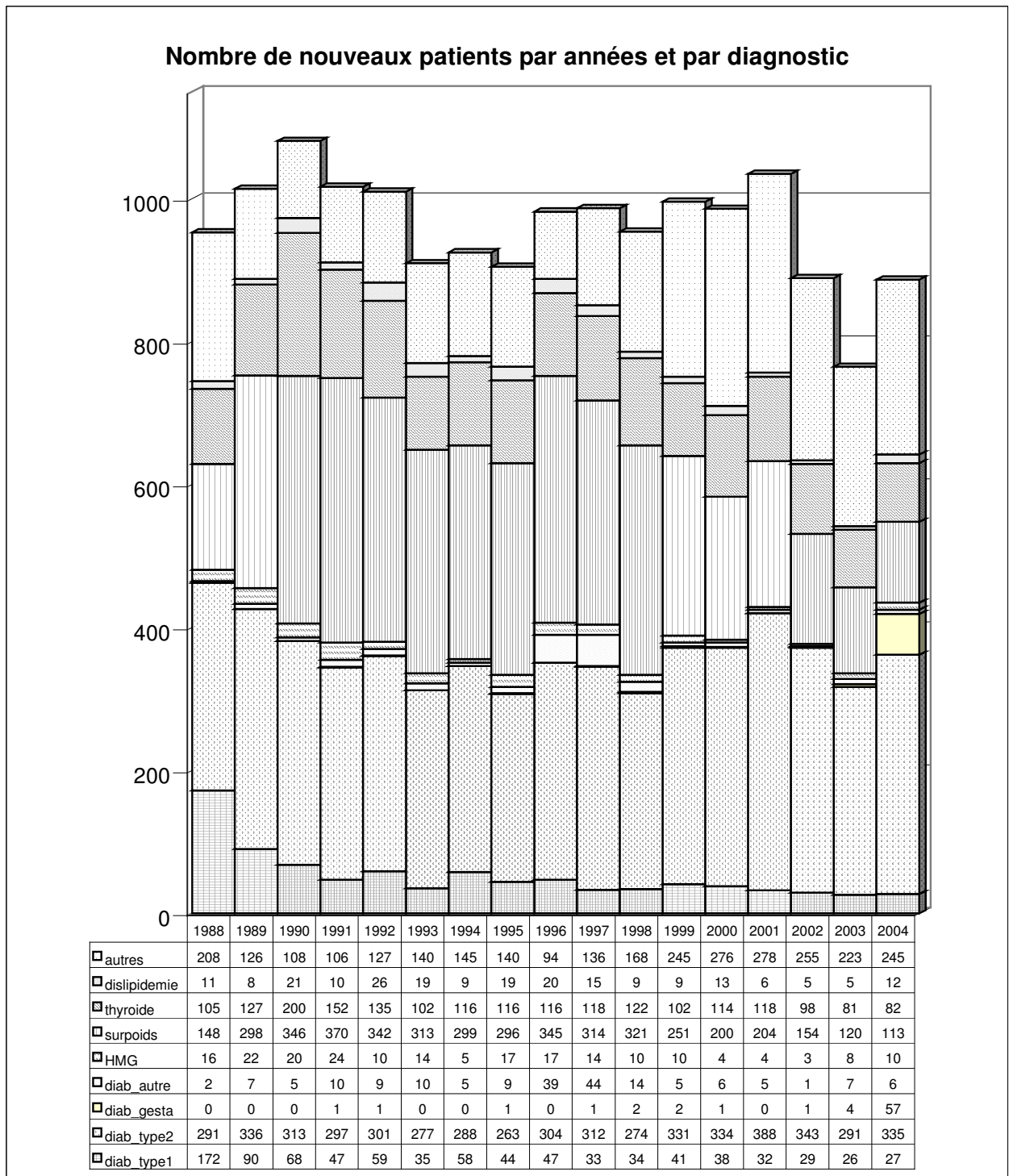
Le recrutement de nouveaux patients diabétiques de type 2 est relativement constant dans le temps.

Le recrutement de nouveaux patients en surpoids isolé était important jusqu'en 1990 puis a subi une diminution progressive et importante.

Les nouveaux patients présentant des pathologies thyroïdiennes sont relativement stables, autour de 100 par an.

Le diabète gestationnel n'apparaît dans la base qu'en 2004 : ceci n'est pas lié à une modification du recrutement, mais à l'apparition de la valeur « diabète gestationnel » dans le champ diagnostic du formulaire de saisie. Auparavant, les femmes atteintes de diabète gestationnel échoyaient dans la rubrique « diabètes autres ».

Le recrutement de nouveaux patients présentant une autre pathologie est variable, avec une nette augmentation à partir de 1998.



**Figure 14 - recrutement des nouveaux patients par année et par diagnostic**

***c- Fidélisation des patients***

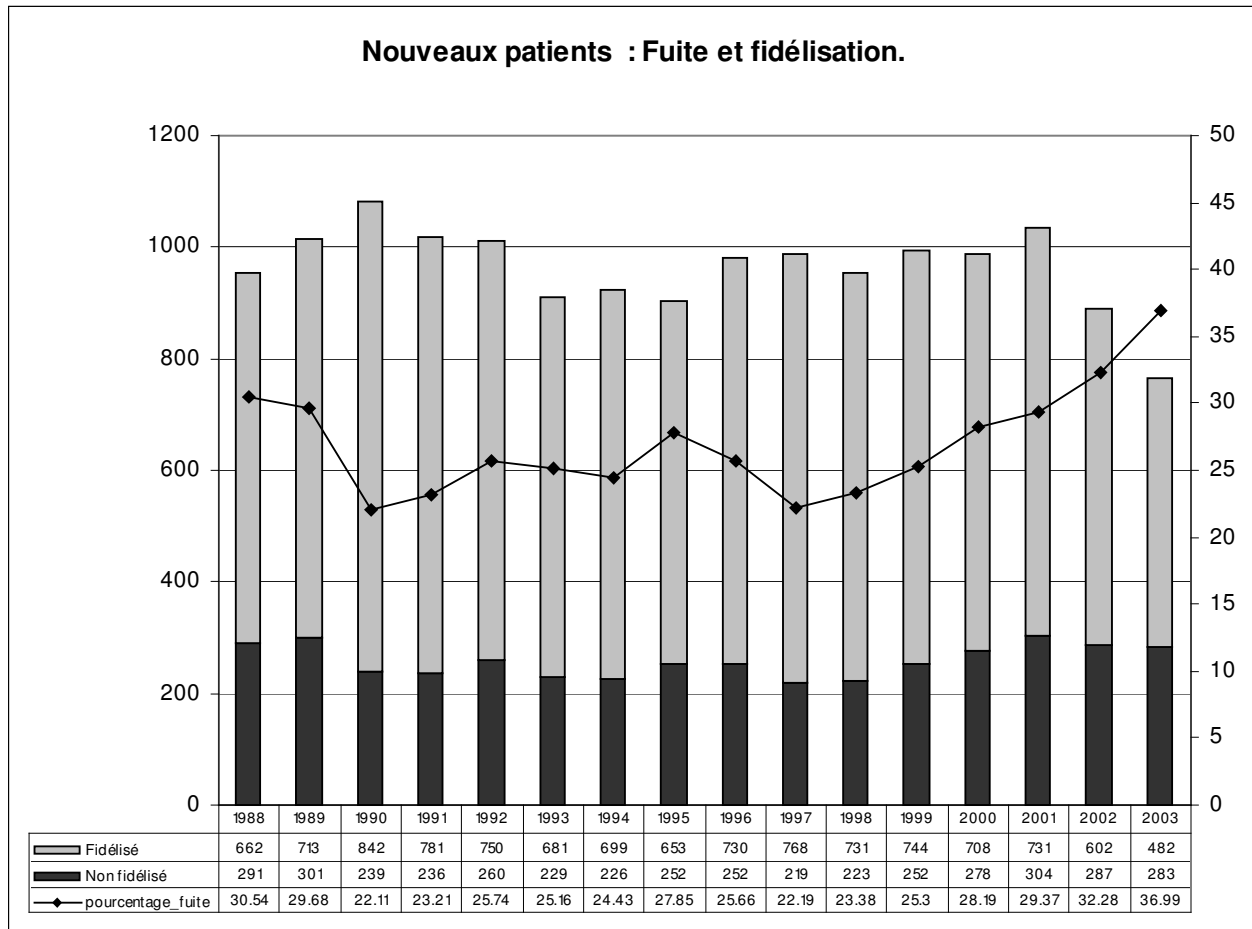
Les patients non fidélisés sont les patients qui viennent une fois, et ne reviendront ensuite jamais.

Nous nous arrêtons à l'année 2003, car il est légitime de penser qu'un patient venu en 2003 et jamais revu au 01/06/2005 est bien perdu de vue. Ce raisonnement serait faux pour les patients venus en 2004.

Sur la Figure 15, nous avons représenté par année :

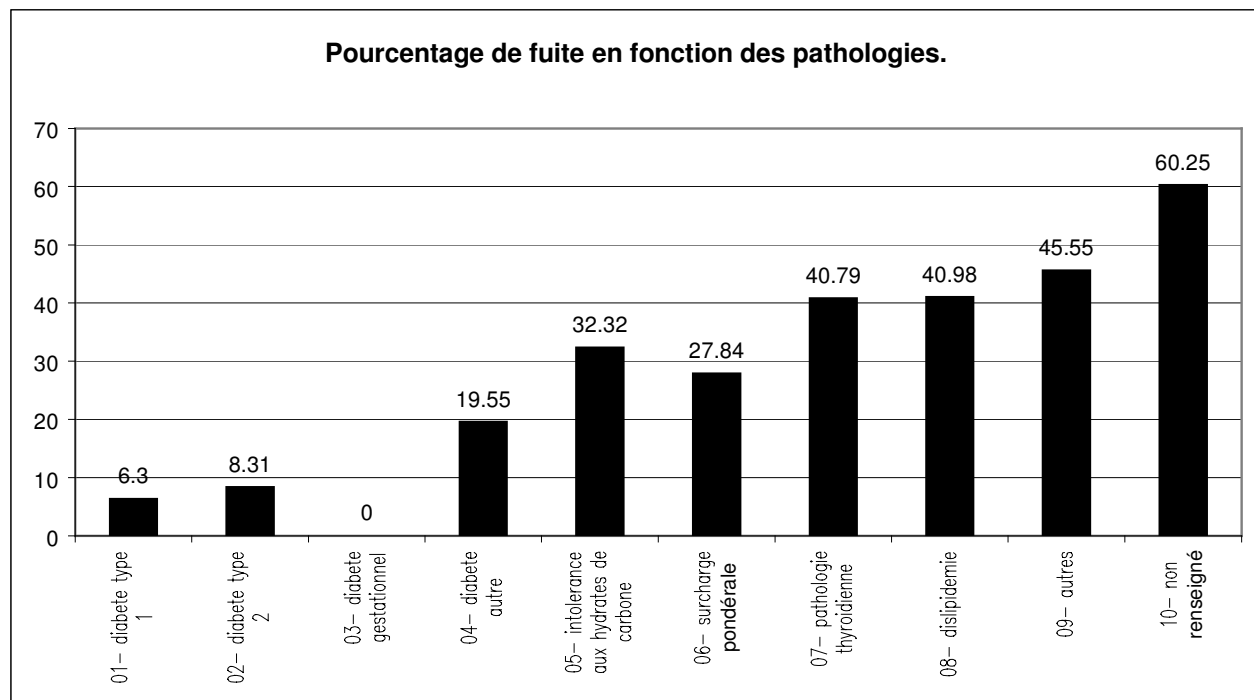
- échelle de gauche : le nombre de patients nouveaux ayant poursuivi le suivi (patients fidélisés), et le nombre de patients ayant arrêté le suivi (patients non fidélisés).
- échelle de droite : le taux de fuite calculé.

Il semble y avoir une remontée très nette du taux de fuite à partir de l'année 1997.



**Figure 15 - Fuite et fidélisation des nouveaux patients**

Sur la Figure 16, nous avons représenté le taux de fuite en fonction de la pathologie, sur l'ensemble des patients.



**Figure 16 - taux de fuite par pathologie**

#### ***d- Durée de suivi des patients***

Nous calculons la durée de suivi des patients en années, définie par la différence (date\_dernier\_événement - date\_premier\_événement). Si un patient ne vient qu'une seule fois, cette variable prend pour valeur zéro.

Nous ne prenons en compte que les patients dont le suivi semble terminé, c'est-à-dire les patients venus pour la dernière fois en 2003. De la sorte, certains patients toujours fidèles ne participent pas à ce calcul. Pour l'interprétation de ces résultats, nous rappelons que la base de données est en place depuis 1988 seulement.

Nous représentons la distribution de cette variable (Figure 17), puis nous représentons la même variable mais après élimination des patients venus une seule fois (Figure 18). A titre d'information, le Tableau 8 représente les effectifs des patients représentés par figure.

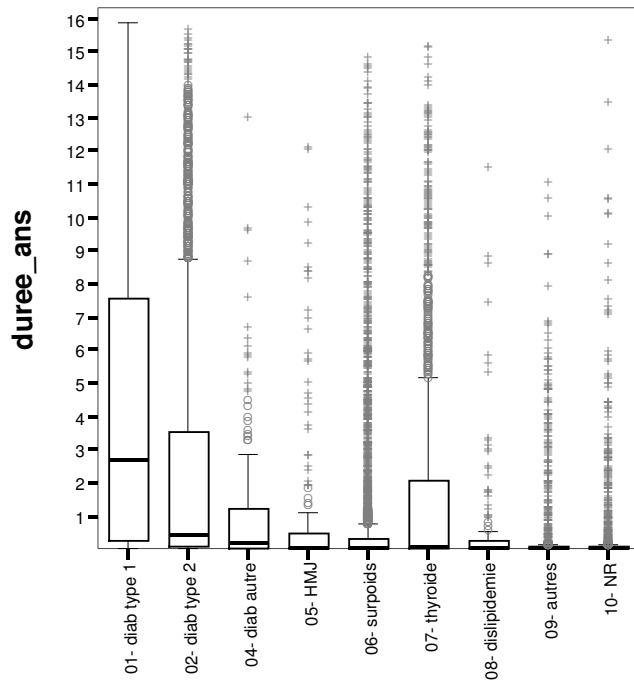


Figure 17 – box plots - distribution par diagnostic des durées de suivi

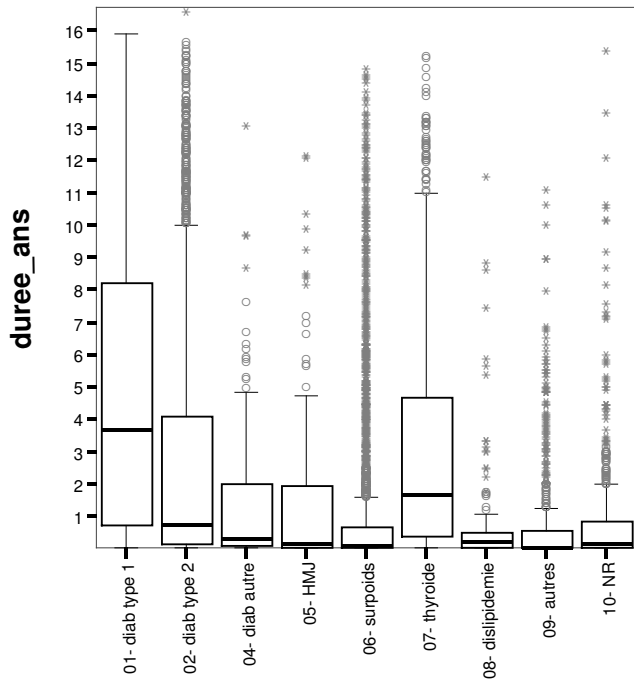


Figure 18 – box plots - distribution par diagnostic des durées de suivi, patients venus plus d'une fois

	Nb patients total (A)	Nb patients dont le suivi est achevé (B)	Rapport B/A (%)	Nb patients dont le suivi est achevé venus plusieurs fois (C)	Rapport C/B (%)
01- diabète type 1	897	516	57.53%	462	89.53%
02- diabète type 2	5456	3631	66.55%	3220	88.68%
04- diabète autre	196	158	80.61%	123	77.85%
05- intolérance aux hydrates de carbone	211	181	85.78%	117	64.64%
06- surcharge pondérale	4499	4212	93.62%	3009	71.44%
07- pathologie thyroïdienne	2044	1750	85.62%	966	55.20%
08- dyslipidémie	220	204	92.73%	120	58.82%
09- autres	1229	1136	92.43%	598	52.64%
10- non renseigné	1920	1501	78.18%	537	35.78%

**Tableau 8 - effectifs utilisés pour les représentations de durées de suivi**

### **3. Les médecins traitants**

#### **a- Dénombrement**

On retrouve 732 médecins traitants correspondants sur l'ensemble des patients de la base, et 533 sur les patients de la file active.

#### **b- Indice de familiarité**

Un médecin traitant peut envoyer plusieurs patients différents au CETRADIMN. Ce nombre de patients est un *indice de familiarité* du médecin traitant. Nous avons effectué un dénombrement des médecins traitants en fonction de leur indice de familiarité.

Les deux graphiques Figure 19 et Figure 20, représentent le nombre de médecins par indice de familiarité, et ce respectivement sur l'ensemble des patients puis sur les patients de la file active.



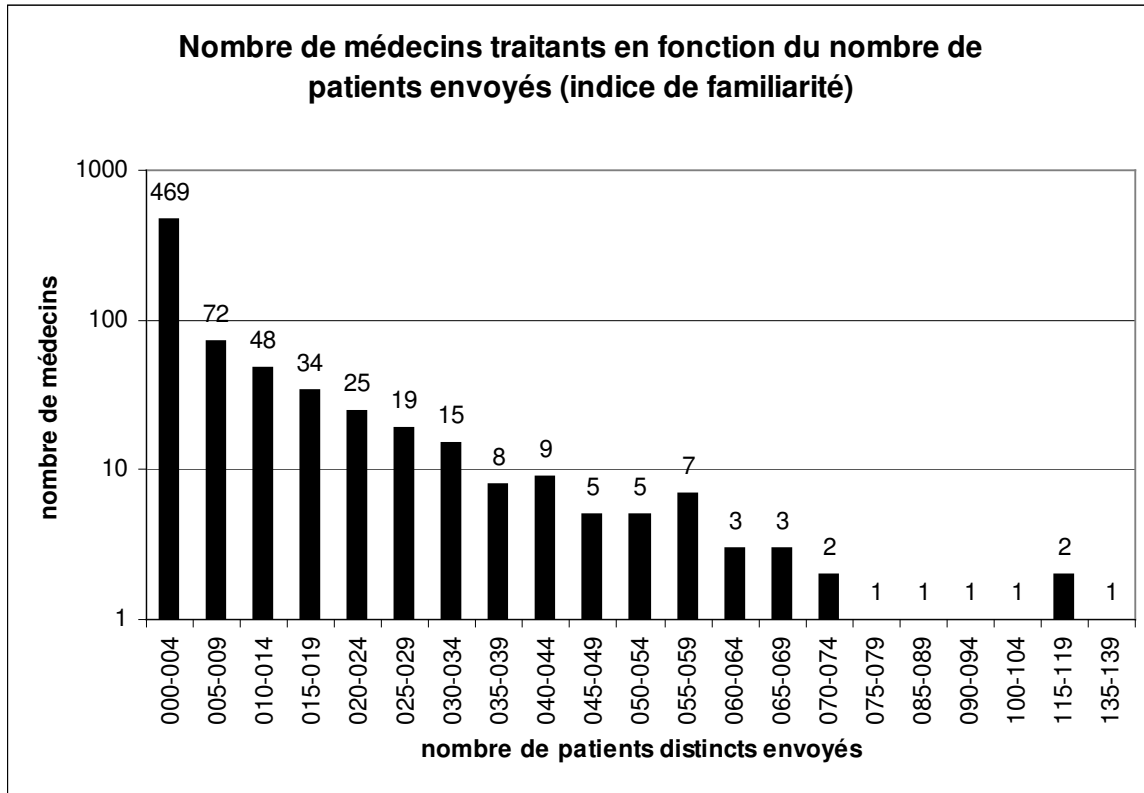


Figure 19 - dénombrements des médecins traitants par indice de familiarité sur l'ensemble des patients

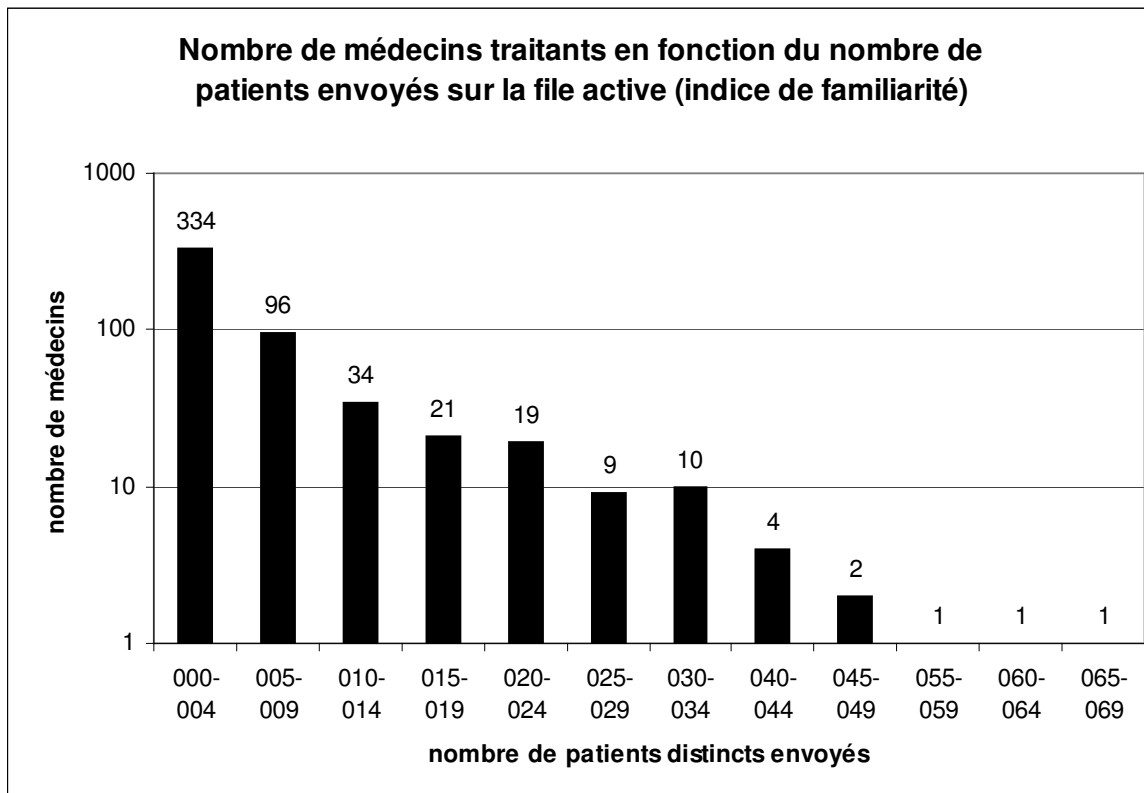


Figure 20 - dénombrements des médecins traitants par indice de familiarité sur la file active.

## 4. Présentation des patients

### a- Démographie

L'âge ne peut être représenté pour des raisons méthodologiques. En effet, l'âge des patients varie selon le moment choisi pour l'analyse rétrospective. Par exemple, si un patient est suivi sur 20 ans, doit-on prendre son âge de première consultation, ou à un autre moment donné ? Nous ne réaliserons donc des pyramides des âges que pour la file active.

Sur la Figure 21, la représentation par camembert de la répartition hommes/femmes montre un net déséquilibre en faveur des femmes. (11474 femmes pour 5298 hommes).

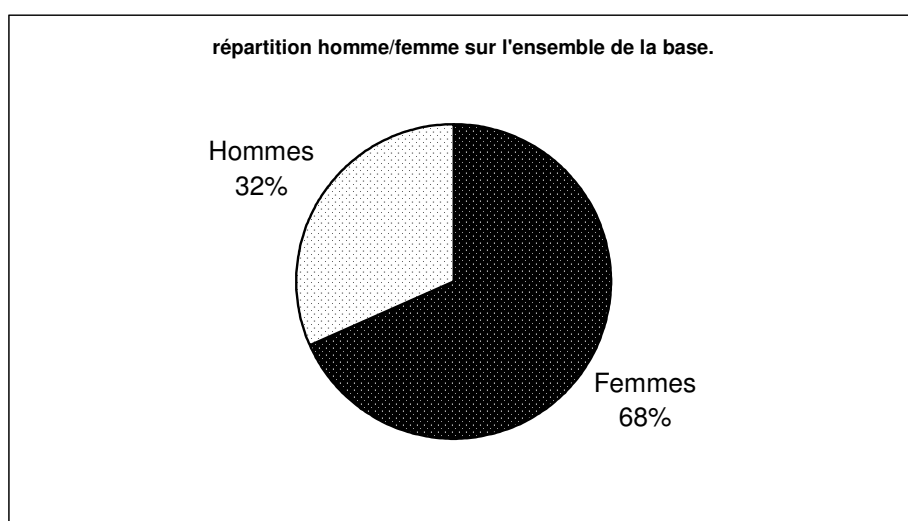


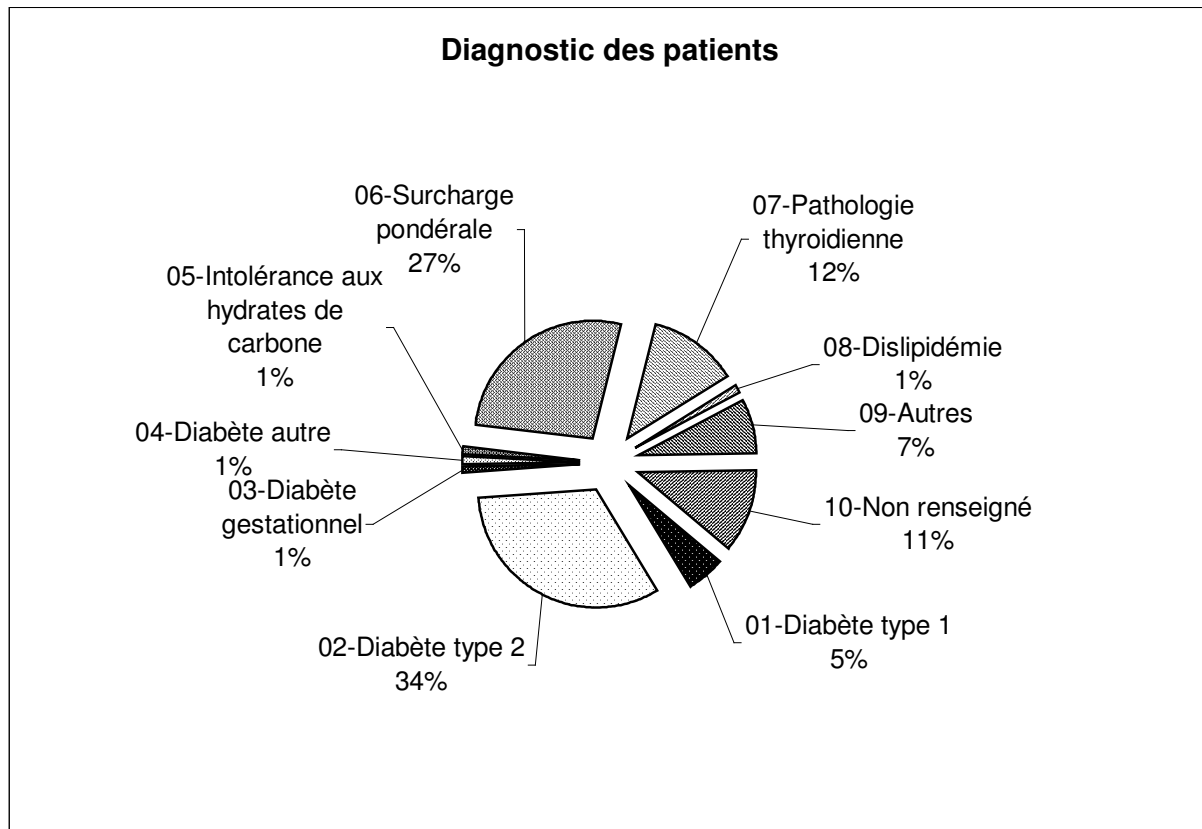
Figure 21 - répartition hommes/femmes sur la totalité de la base.

### b- Diagnostic

Nous représentons en Figure 22 la répartition des patients par diagnostic sur la population globale. La détermination du diagnostic est explicitée dans la partie I.D Définition du diagnostic final (p38).

Diagnostic final	nombre
01 Diabète type 1	897
02 Diabète type 2	5456
03 Diabète gestationnel	100
04 Diabète autre	196
05 Intolérance aux hydrates de carbone	211
06 Surcharge pondérale	4499
07 Pathologie thyroïdienne	2044
08 Dislipidémie	220
09 Autres	1229
10 Non renseigné	1920

**Tableau 9 - Dénombrement des patients par diagnostic**



**Figure 22 - Répartition par diagnostic sur l'ensemble de la base.**

### ***c- Nouvelles recrues : évolution du diagnostic***

Nous avons représenté l'évolution des nouveaux patients par pathologie sous la forme de quatre courbes évoluant en fonction du temps (Figure 23).

Nous remarquons :

- une augmentation très nette à partir de 1995 des nouveaux patients de pathologie autre.

- Une diminution très nette du recrutement de nouveaux patients en surcharge pondérale à partir de 1998
- Une certaine stabilité du recrutement de nouveaux patients diabétiques et des pathologies thyroïdiennes.

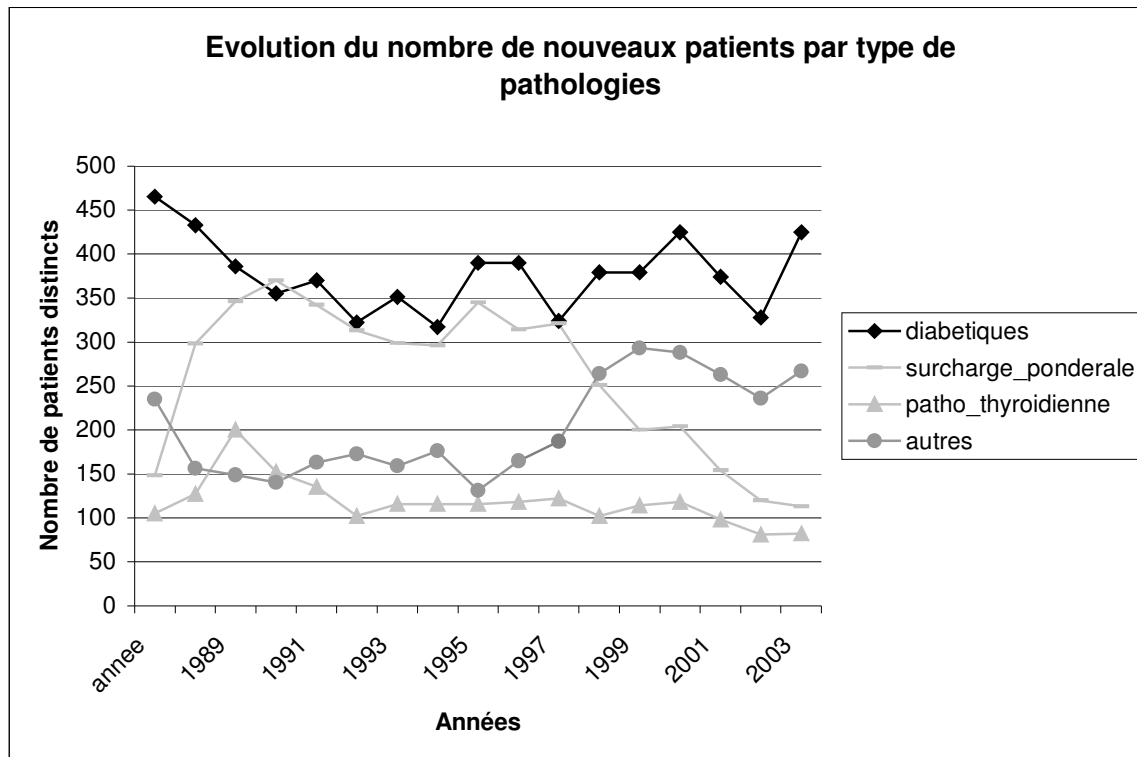


Figure 23 - évolution du nombre de nouveaux patients par type de pathologie.

Nous avons représenté l'évolution du nombre de nouveaux diabétiques en fonction de leur type de diabète par quatre courbes évoluant en fonction des années. (Figure 24).

Nous remarquons :

- Une baisse puis stabilisation du recrutement des patients diabétiques de type 1 après un recrutement important dans les premières années.
- Un début de recrutement des patientes porteuses d'un diabète gestationnel en 2004 – il s'agit en réalité d'un artéfact informatique.
- Une relative stabilité du recrutement des patients diabétiques de type 2.

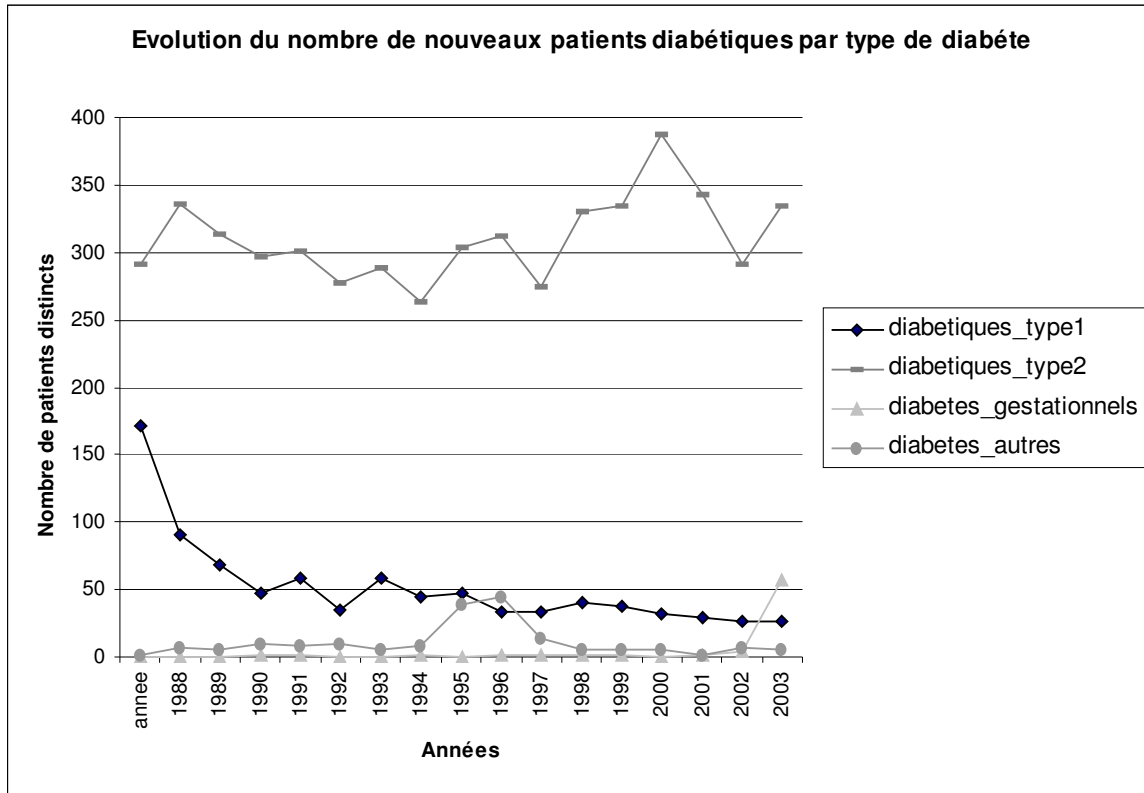


Figure 24 - évolution du nombre de nouveaux patients par type de diabète

## 5. Présentation des patients de la file active

### a- Démographie

Sur la Figure 25 la représentation par camembert de la répartition hommes/femmes note un net déséquilibre en faveur des femmes (4371 femmes pour 2394 hommes).

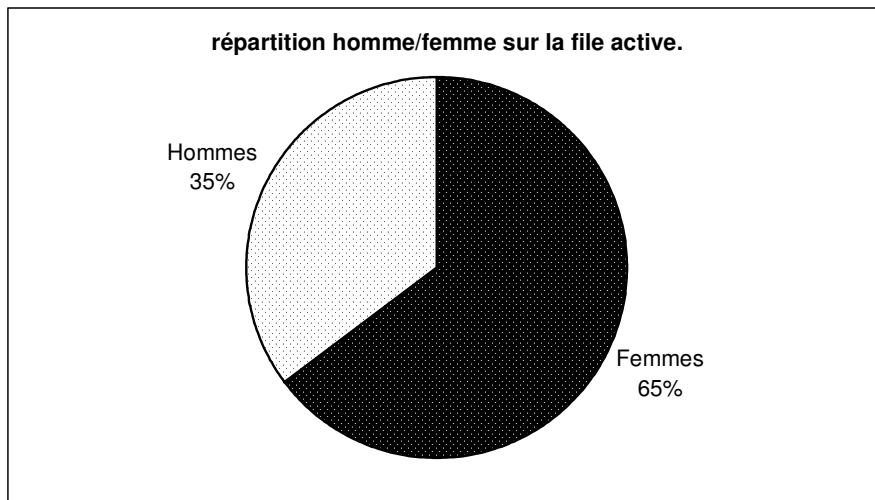


Figure 25 - répartition hommes/femmes de la file active

Sur la Figure 26 nous avons représenté une pyramide des âges des patients de la file active au 1<sup>er</sup> janvier 2005. Ce choix se justifie par l'hypothèse que les patients sont encore tous suivis.

Pour l'ensemble de nos pyramides des âges, nous avons arbitrairement choisi de :

- découper les tranches d'âges par 5 ans, elles sont représentées en ordonné
- Les hommes seront dénombrés à droite (pour des raisons techniques de représentations ils ont une valeur négative mais il faut bien lire le nombre en valeur absolue)
- Les femmes seront représentées à droite.

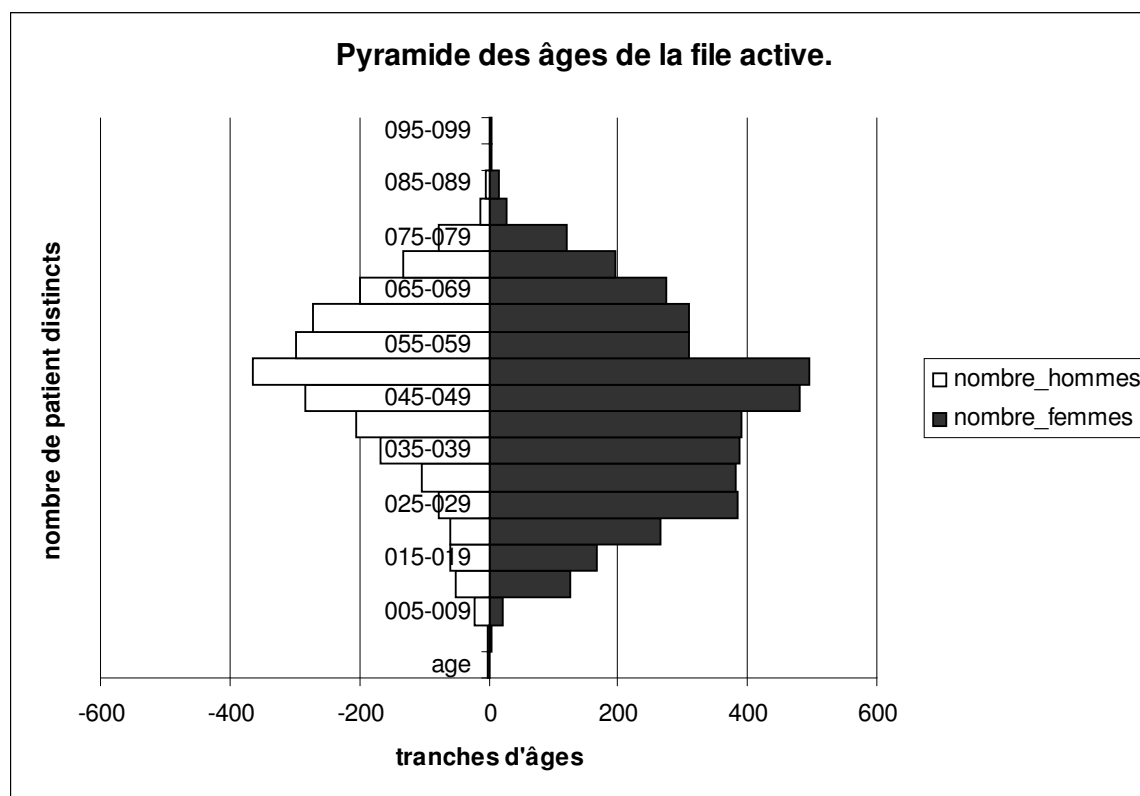


Figure 26 - Pyramide des âges de la file active

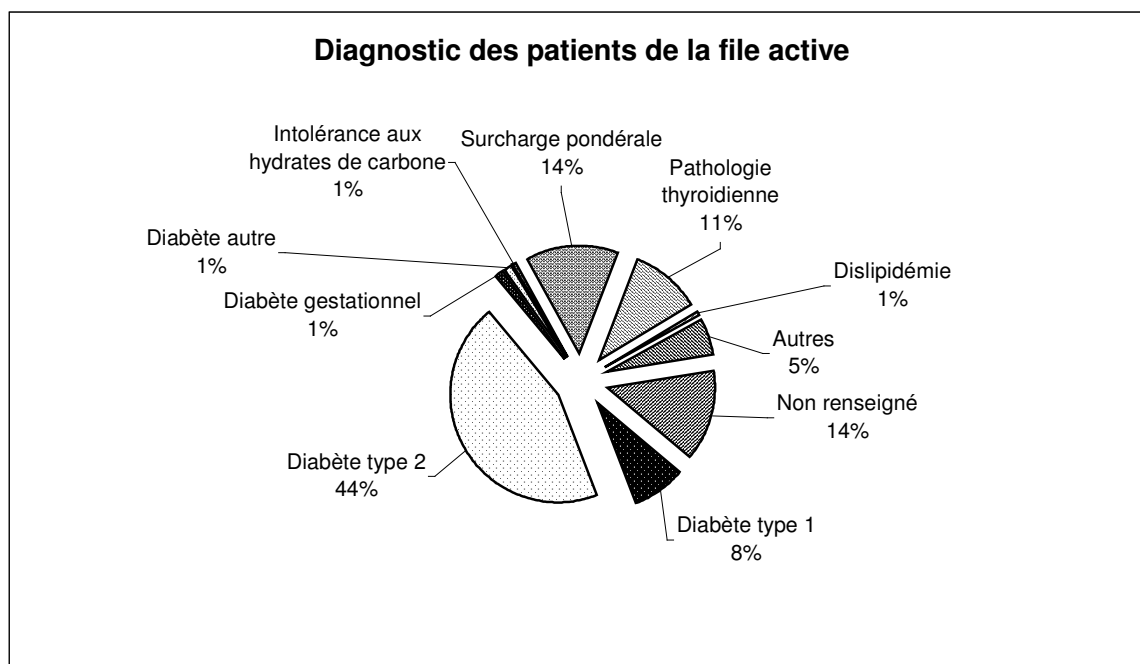
### ***b- Diagnostic***

Nous représentons en Figure 26 la répartition des patients de la file active par diagnostic.

La part de diabète de type 2 est en forte augmentation sur la file active par rapport à la population globale.

Diagnostic final	nombre
Diabète type 1	536
Diabète type 2	3033
Diabète gestationnel	100
Diabète autre	64
Intolérance aux hydrates de carbone	47
Surcharge pondérale	932
Pathologie thyroïdienne	723
Dislipidémie	38
Autres	362
Non renseigné	930

**Tableau 10 - dénombrement des patients de la file active par diagnostic.**

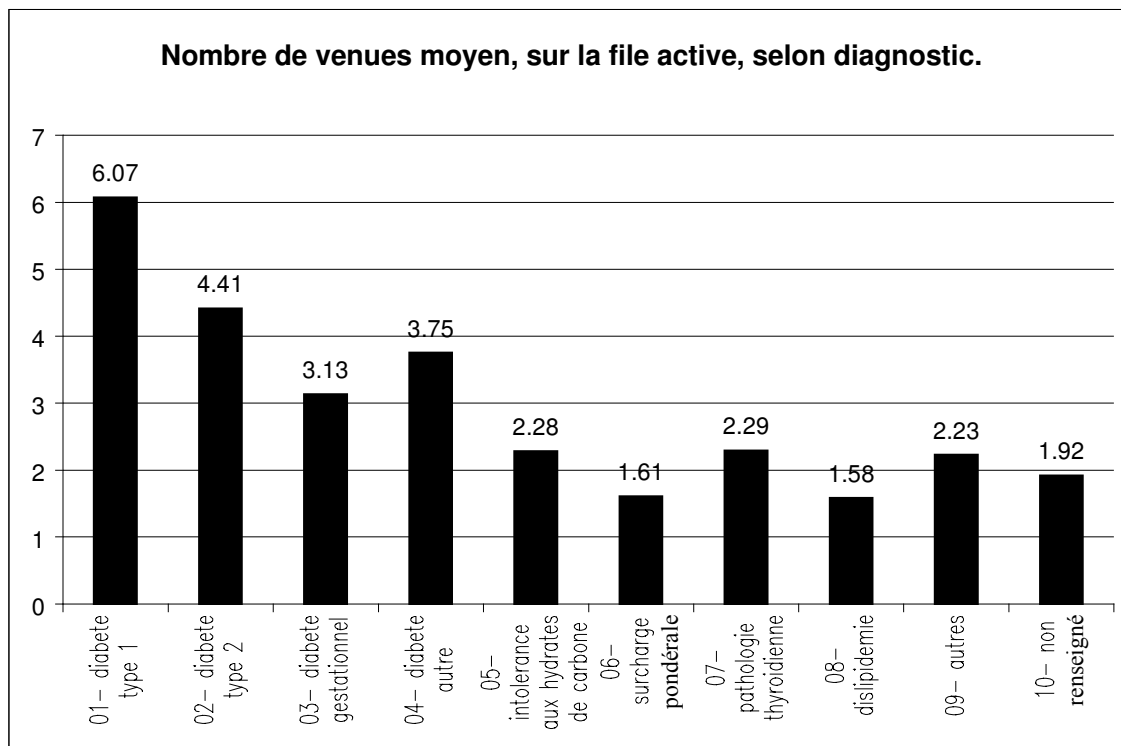


**Figure 27 - répartition des patients de la file active par diagnostic.**

### ***c- Habitudes de fréquentation***

Les habitudes de fréquentation permettent de cerner des profils de consommation des soins.

La Figure 28 représente sous forme d'histogramme le nombre d'événements moyen pour chaque patient selon la pathologie sur les quatre années et demie de suivi de la file active.



**Figure 28 - Nombre de venues moyen par type de pathologie**

La Figure 29 représente sous forme de box plot la distribution du nombre total d'événements par patient selon le diagnostic. Ce graphique met en évidence deux modes de consultation au CETRADIMN : les consultations pour un suivi régulier (diabète type 1, type 2, et dans une certaine mesure le diabète gestationnel sur une durée limitée), et les consultations pour un début de prise en charge ou un avis d'expert (les autres).



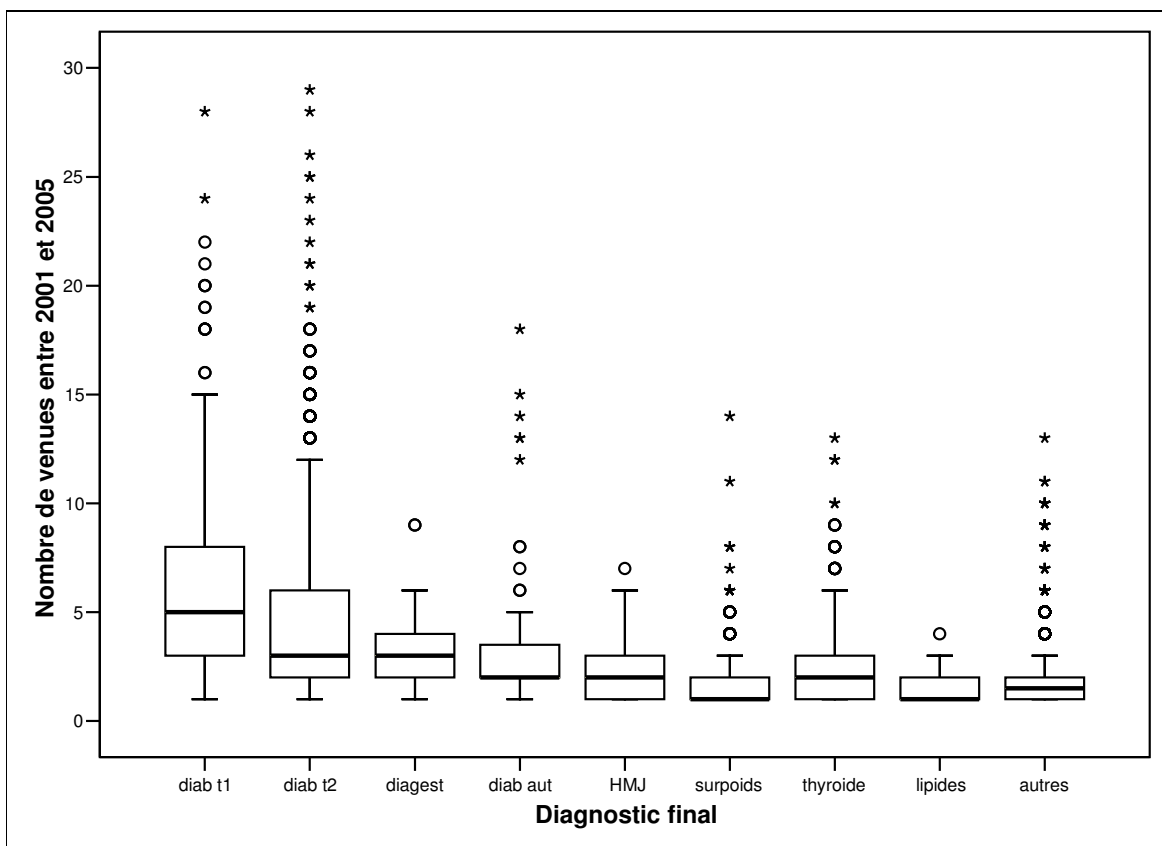


Figure 29 - box plots - distribution du nombre de venues par patient selon le diagnostic.

Dans la Figure 30, nous calculons, pour chaque patient venu au moins deux fois entre 2001 et 2005, l'intervalle moyen entre les venues. Ensuite, nous représentons la distribution de cette variable selon le diagnostic, à l'aide de Box Plots.

En guise de notice de lecture, on peut lire sur ce graphique que les patients diabétiques de type 1 espacent leurs venues de 4 mois (intervalle médian), tandis que les patients diabétiques de type 2 espacent leurs venues de 3 mois seulement. Les patientes suivies pour un diabète gestationnel viennent tous les mois, cet intervalle ne varie pratiquement pas entre les patientes. Certaines distributions très asymétriques sont difficilement interprétables sans connaissance de la pratique. Ainsi, les patients en surpoids isolé, sont convoqués plusieurs fois de manière très rapprochée. Cependant, certains d'entre eux s'arrêtent rapidement puis reviennent longtemps après pour une « deuxième tentative ».

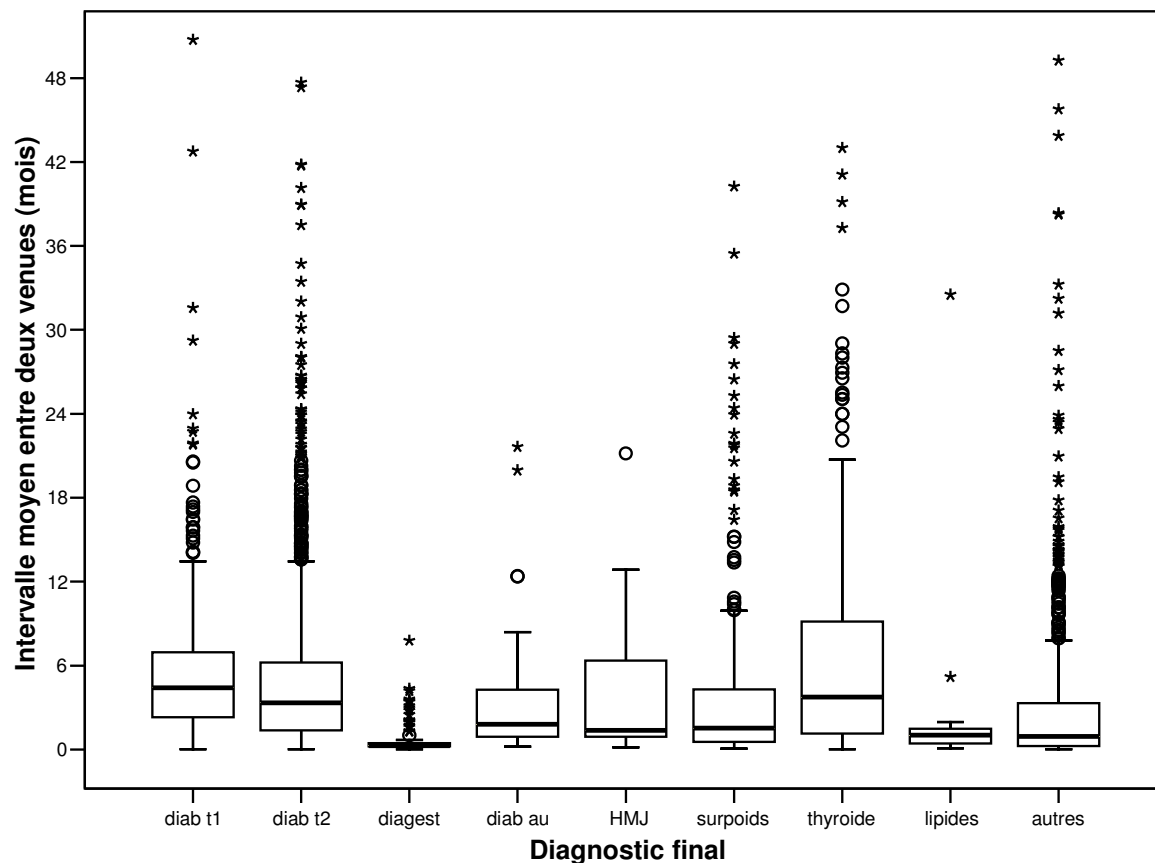


Figure 30 – box plots - distribution, par diagnostic, de l'intervalle moyen entre les venues du patient

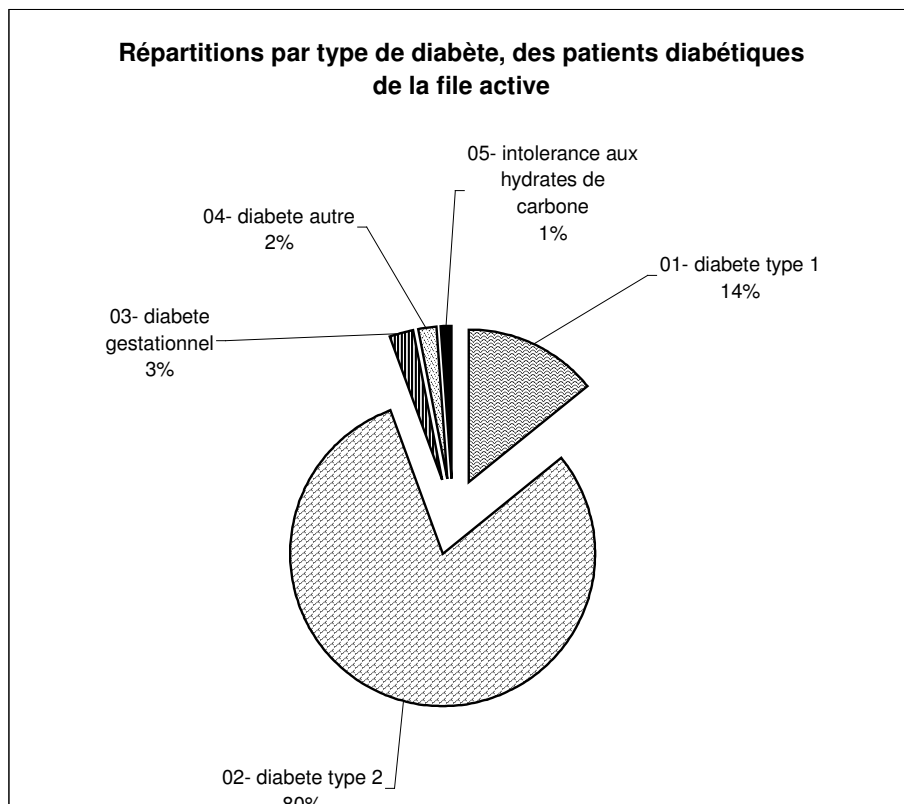
## 6. Diabétiques et HMJ de la file active - démographie

### a- Tous diabétiques et HMJ file active

Nous représentons dans la Figure 31 les effectifs par type de diabète, sous forme de camemberts. On remarquera la place prépondérante du diabète de type 2.

Diagnostic final	nombre
01- diabète type 1	536
02- diabète type 2	3033
03- diabète gestationnel	100
04- diabète autre	64
05- intolérance aux hydrates de carbone	47

Tableau 11 - dénombrements par type de diabète des patients diabétiques de la file active

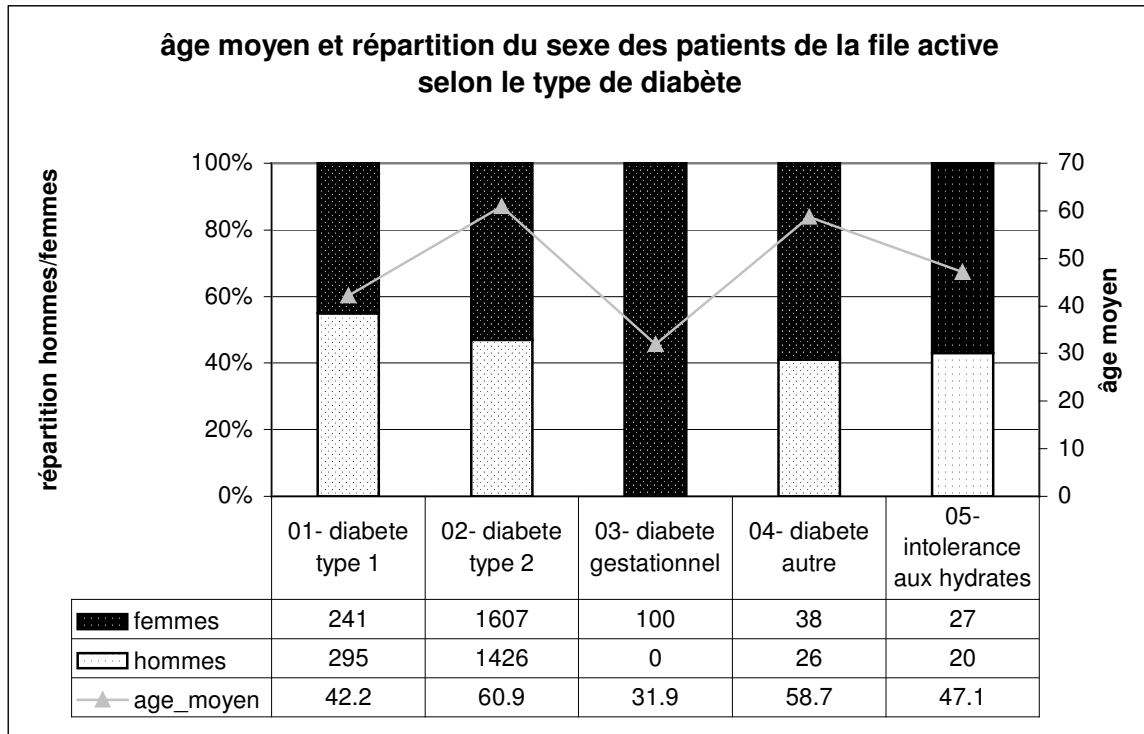


**Figure 31 - Répartitions par type de diabète des patients diabétiques de la file active.**

Sur la Figure 32, nous avons représenté la répartition hommes/femmes par pathologie sous forme d'histogramme empilé à 100% (échelle de gauche) ainsi que l'âge moyen des patients par pathologie (échelle de droite).

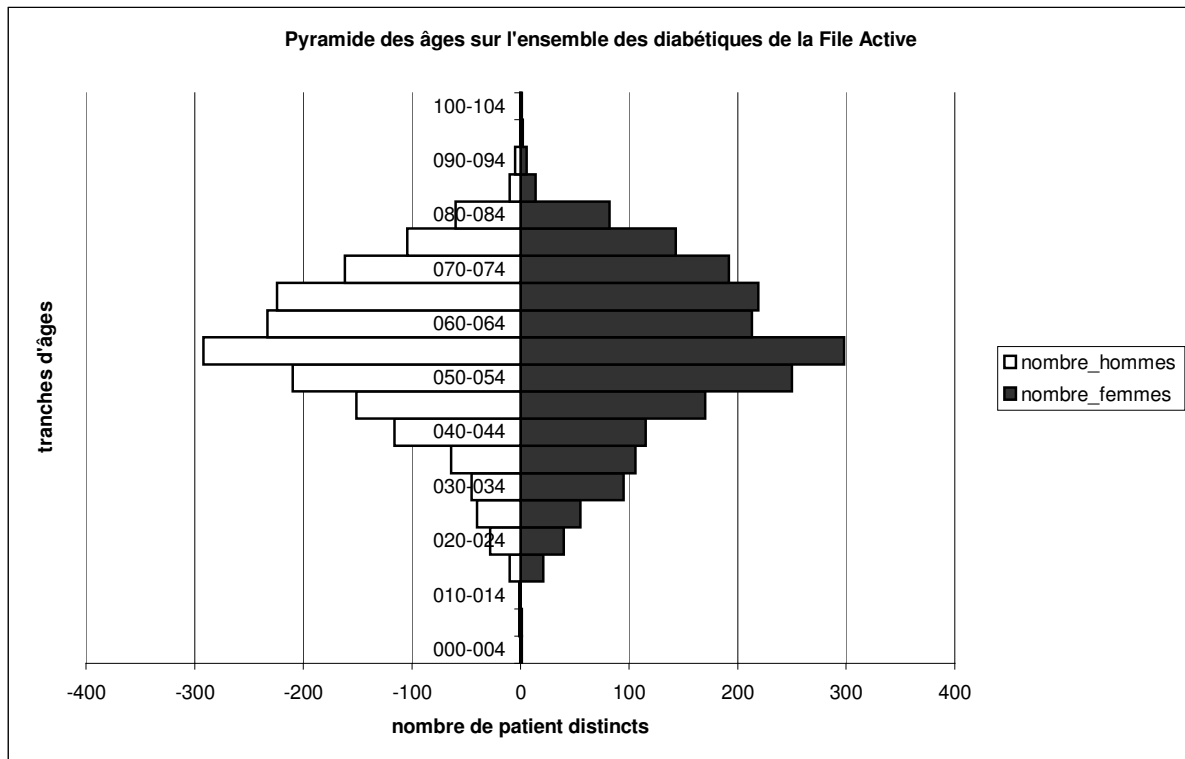
Nous remarquons :

- Une proportion hommes/femmes en faveur des hommes uniquement dans le diabète de type 1
- L'âge moyen est nettement inférieur dans le diabète de type 1 et le diabète gestationnel.



**Figure 32 - Patients diabétiques de la file active, représentation de l'âge moyen et de la répartition hommes/femmes**

Sur la Figure 33 nous avons représenté la pyramide des âges des patients diabétiques de la file active.



**Figure 33 - Pyramide des âges des patients diabétiques de la file active**

### ***b- Diabète type 1 dans la file active***

La Figure 34 représente la répartition hommes / femmes sous forme d'un camembert des diabétiques de type 1 de la file active.

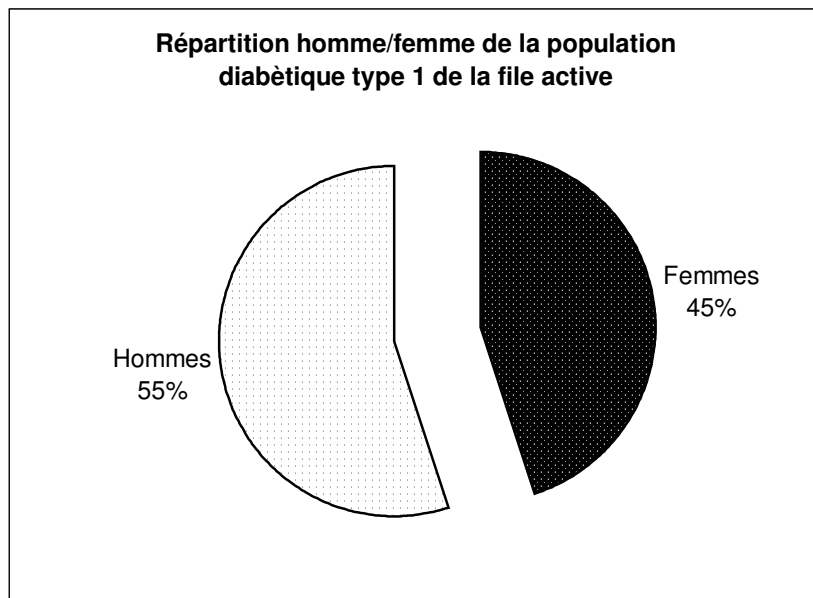


Figure 34 - répartition hommes/femmes des diabétiques de type 1 de la file active.

Sur la Figure 35 qui représente la répartition des patients diabétiques de type 1 selon leur âge et leur sexe, on remarque une nette répartition en faveur des tranches d'âges jeunes.

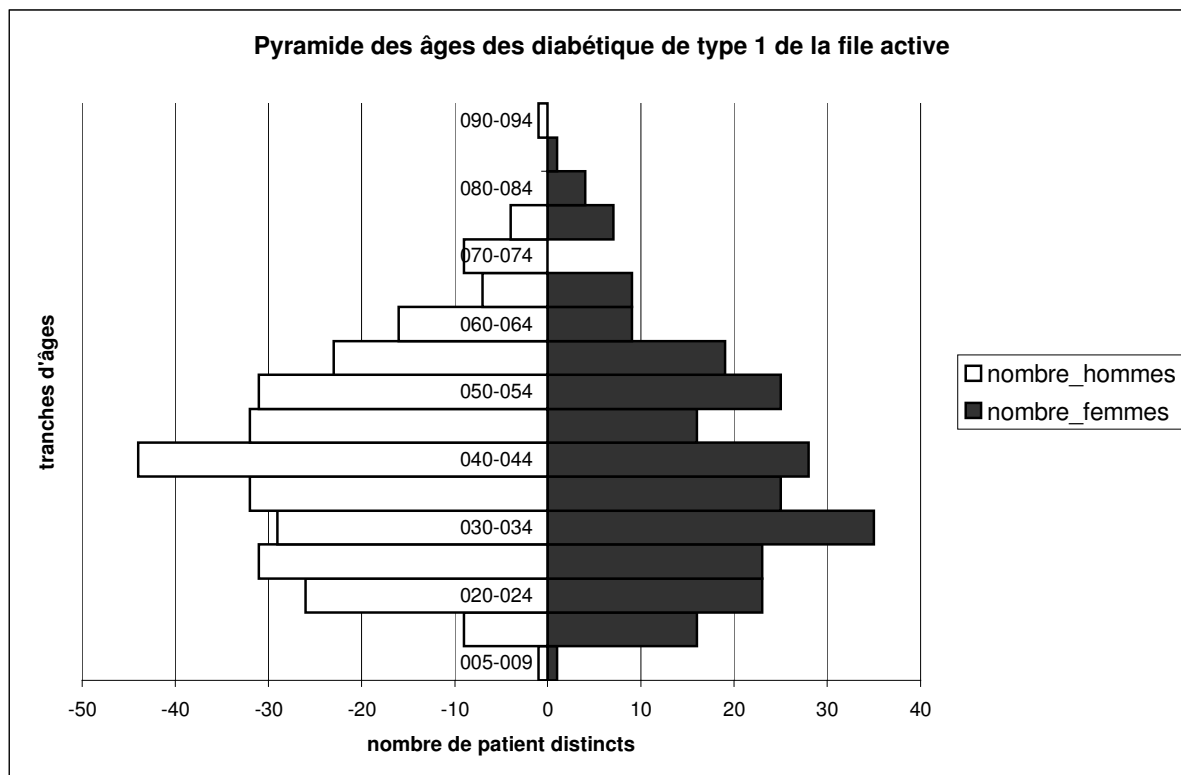


Figure 35 - Pyramide des âges des diabétiques type 1 de la file active

### c- Diabète type 2 dans la file active

La Figure 36 représente la répartition hommes / femmes sous forme d'un camembert des diabétiques de type 2 de la file active.

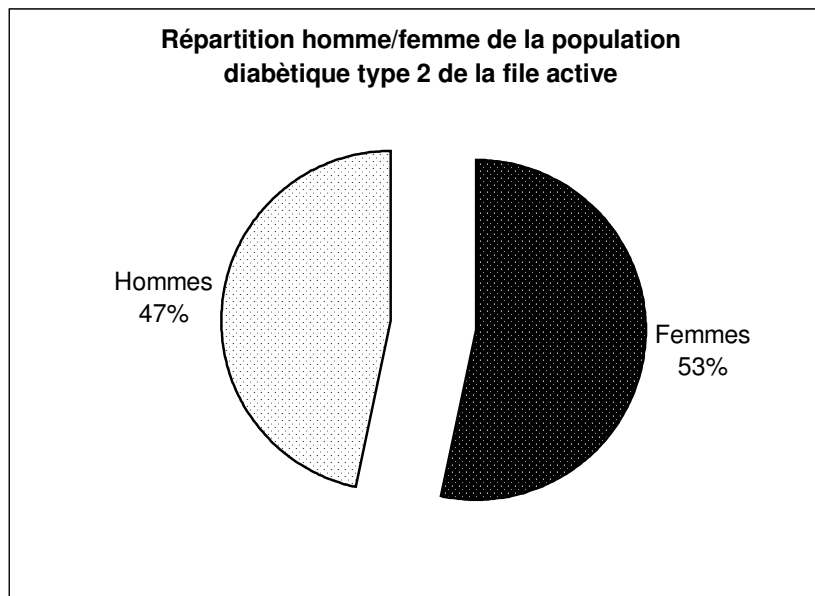


Figure 36 - répartition hommes/femmes des diabétiques type 2 de la file active.

Sur la Figure 37, nous représentons la répartition des patients diabétiques de type 2 selon leur âge et leur sexe. Nous remarquons une nette répartition en faveur des tranches d'âges de plus de 50 ans.

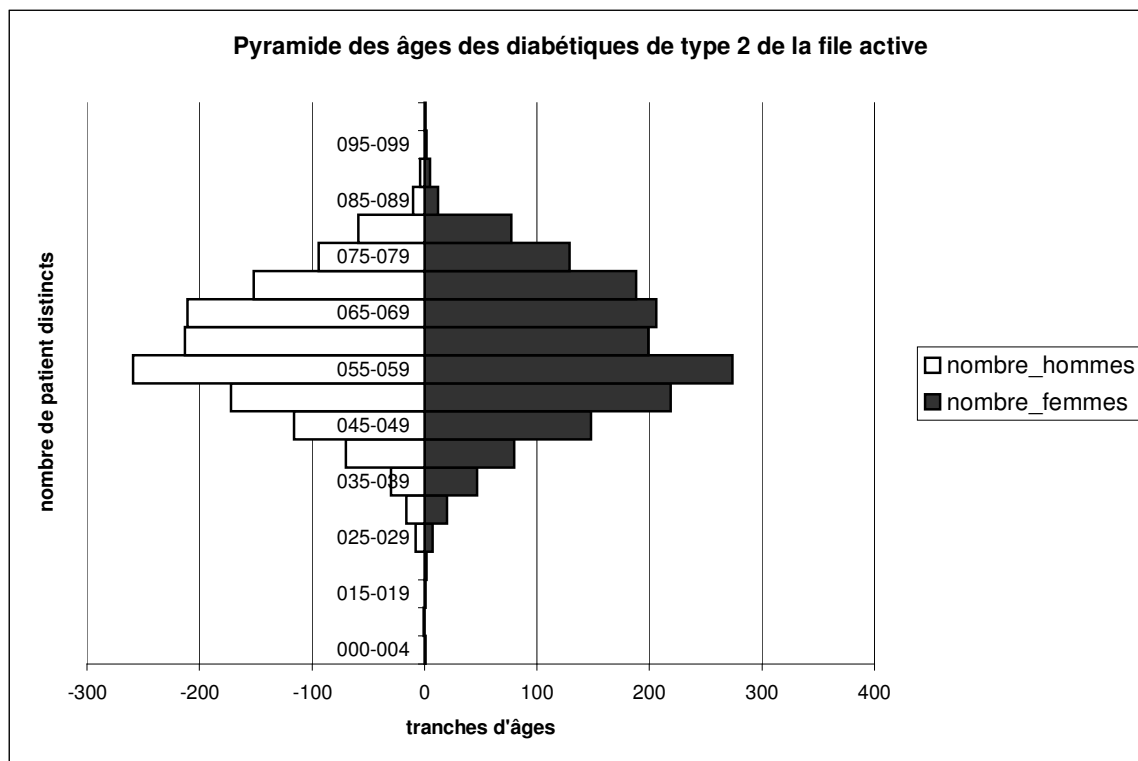


Figure 37 - Pyramide des âges des diabétiques de type 2 de la file active

#### ***d- Diabète gestationnel dans la file active***

Il apparaît évident que 100% des diabètes gestationnels sont des femmes. Nous rappelons qu'il s'agit uniquement des patientes venues à partir de 2004. Auparavant, ces patientes étaient perdues dans la catégorie « diabètes autres ».

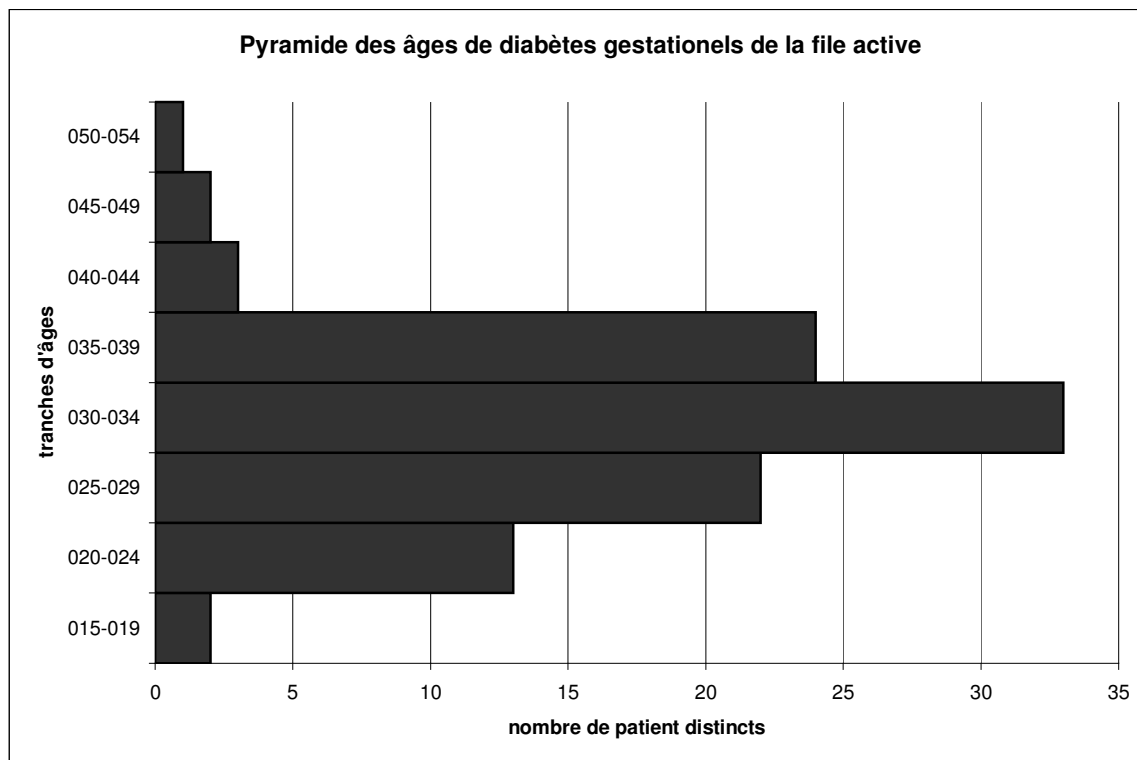


Figure 38 - Pyramide des âges des patientes présentant un diabète gestationnel

#### ***e- Hyperglycémie modérée à jeun (HMJ) dans la file active***

Le nombre de patients restant dans la catégorie HMJ est trop faible pour justifier une représentation graphique. Effectivement, la plupart des patients en HMJ deviennent ensuite des diabétiques de type 2.

### ***7. Diabétiques de la file active – complications, morbidités associées***

#### ***a- Complications ophtalmologiques***

Les complications ophtalmologiques sont déterminées par le seul champ *complications*. Il s'agit donc d'une évaluation basse car elle ne peut tenir compte des oublis d'encodage.

La Figure 39 représente les patients porteurs d'une complication ophtalmologique à côté du nombre de patients total par type de diabète.

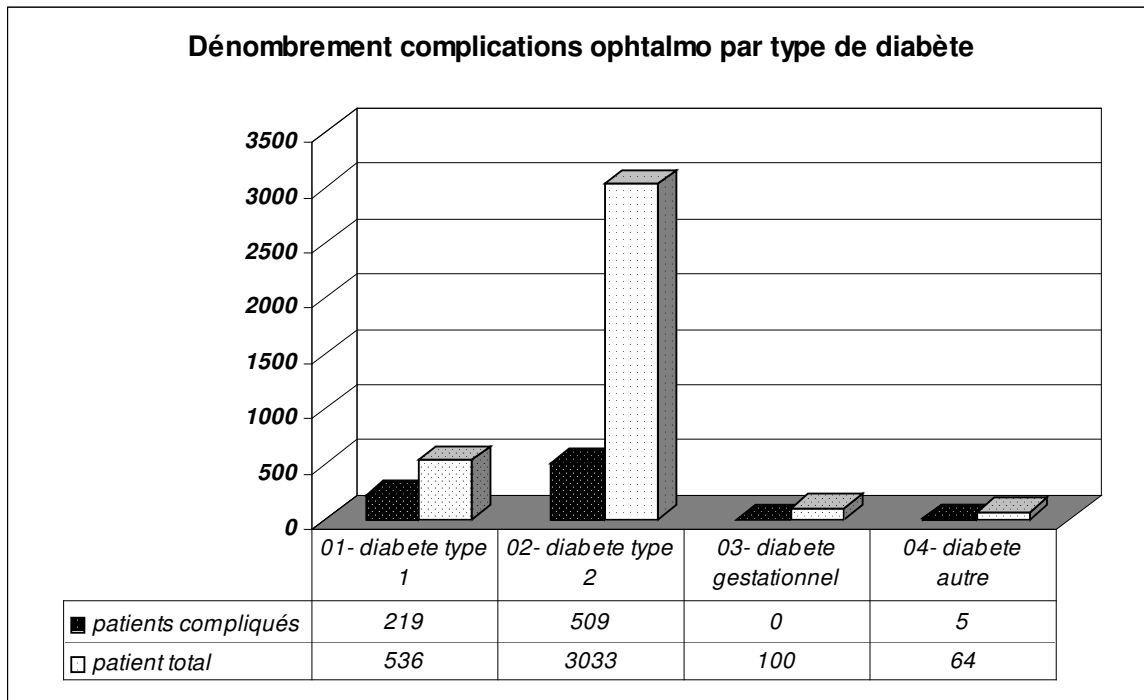


Figure 39 - dénombrement des complications ophtalmologiques

**b- Complications rénales**

La Figure 40 représente les patients dont le champ *diagnostic* mentionnait des complications rénales.

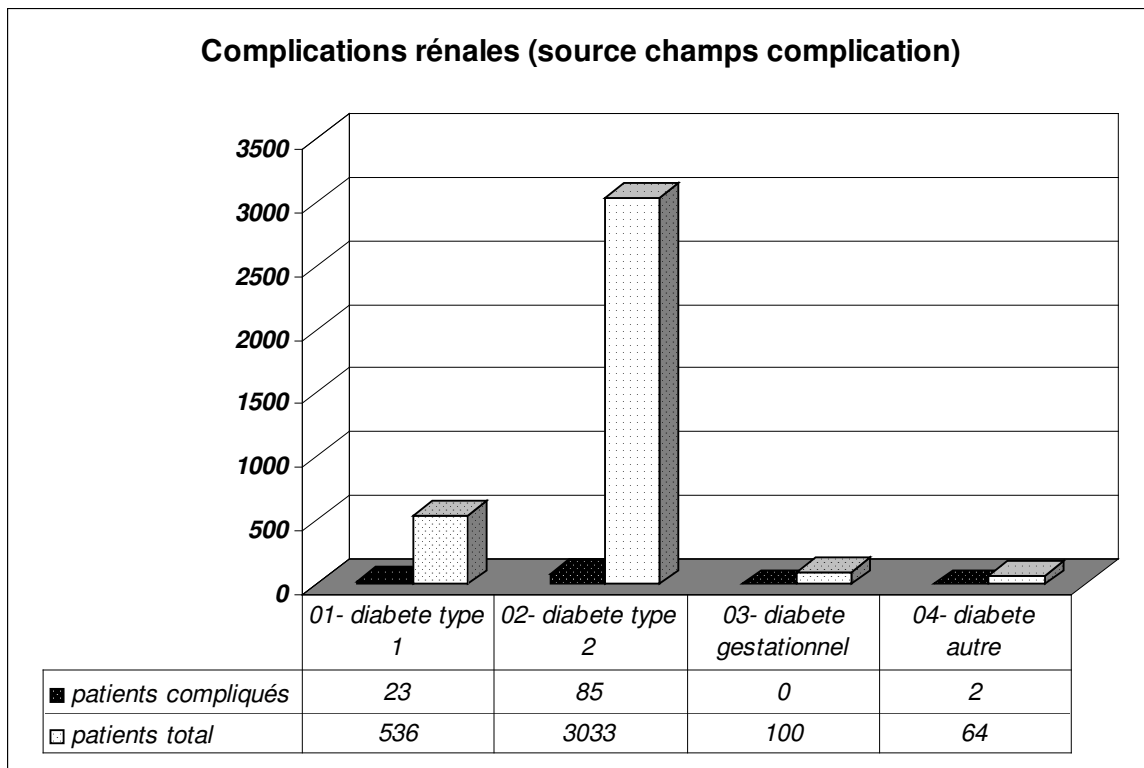


Figure 40 - Dénombrement des complications rénales sur le champs complications.



La Figure 41 représente les complications rénales basées sur les champs *albuminurie*. Nous avons déterminé la micro albuminurie et la macro albuminurie selon les recommandations de la HAS [43]

On définit ainsi trois plages :

- la normo albuminurie : < 20 mg/l (soit < 30 mg/24 h)
- la micro albuminurie : 20 à 200 mg/l (soit 30 à 300 mg/24 h)
- la macro albuminurie. : > 200 mg/l (soit > 300 mg/24 h)

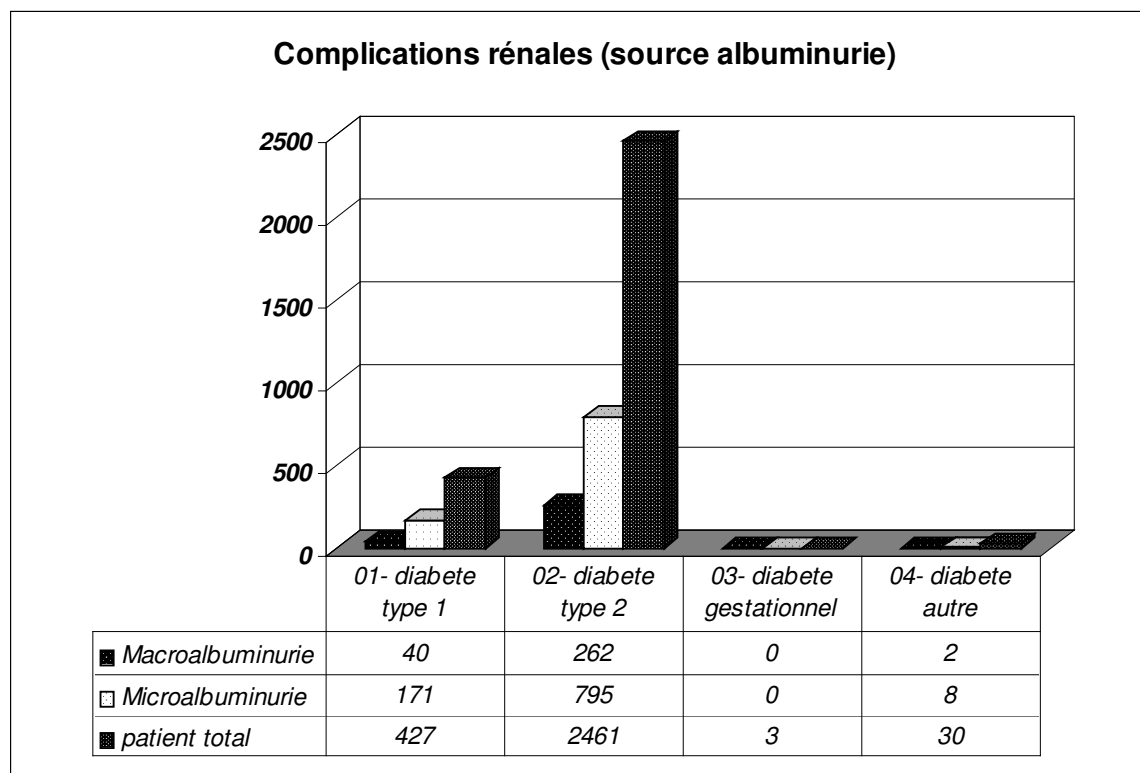


Figure 41 - Dénombrement des complications rénales sur le champs albuminurie

La Figure 42 représente les complications rénales sur le champ *clearance de la créatinine*. Nous avons considéré en nous basant sur les recommandations de la HAS qu'un patient présentait une insuffisance rénale modérée si sa clearance était inférieure à 60 ml/min. [43]

La clearance de la créatinine a été calculée par la formule de Cockcroft [43] :

– avec la créatinémie exprimée en mg/l :

chez l'homme :

$$\text{DFG (ml/min)} = [(140 - \text{âge}) \times \text{poids} / 7,2 \times \text{créatinémie en mg/l}],$$

chez la femme :

$$\text{DFG (ml/min)} = [(140 - \text{âge}) \times \text{poids} / 7,2 \times \text{créatinémie en mg/l}] \times 0,85$$

– avec la créatinémie exprimée en  $\mu\text{mol/l}$

$$\text{DFG (ml/min)} = [(140 - \text{âge}) \times \text{poids} / \text{créatinémie en } \mu\text{mol/l}] \times k,$$

Avec  $k = 1,23$  pour les hommes,  $1,04$  pour les femmes, poids en kg, âge en années.

La base ne nous permettait pas d'avoir sur un même événement le poids et la créatinine lui correspondant. De plus la créatinine et le poids sont variables dans le temps. Nous avons donc choisi d'appliquer une stratégie de manière à minimiser les faux positifs. Nous avons donc choisi de prendre pour référence le poids le plus élevé (ce qui a tendance à maximiser la clearance de la créatinine) et de prendre la créatinine la plus élevée. Par ailleurs comme nous le mentionnons dans la partie matérielle les poids aberrants le sont par excès. Il est bien évident qu'une vision dynamique de la clearance aurait présenté un intérêt certain malheureusement cette étude rétrospective ne peut techniquement le réaliser.

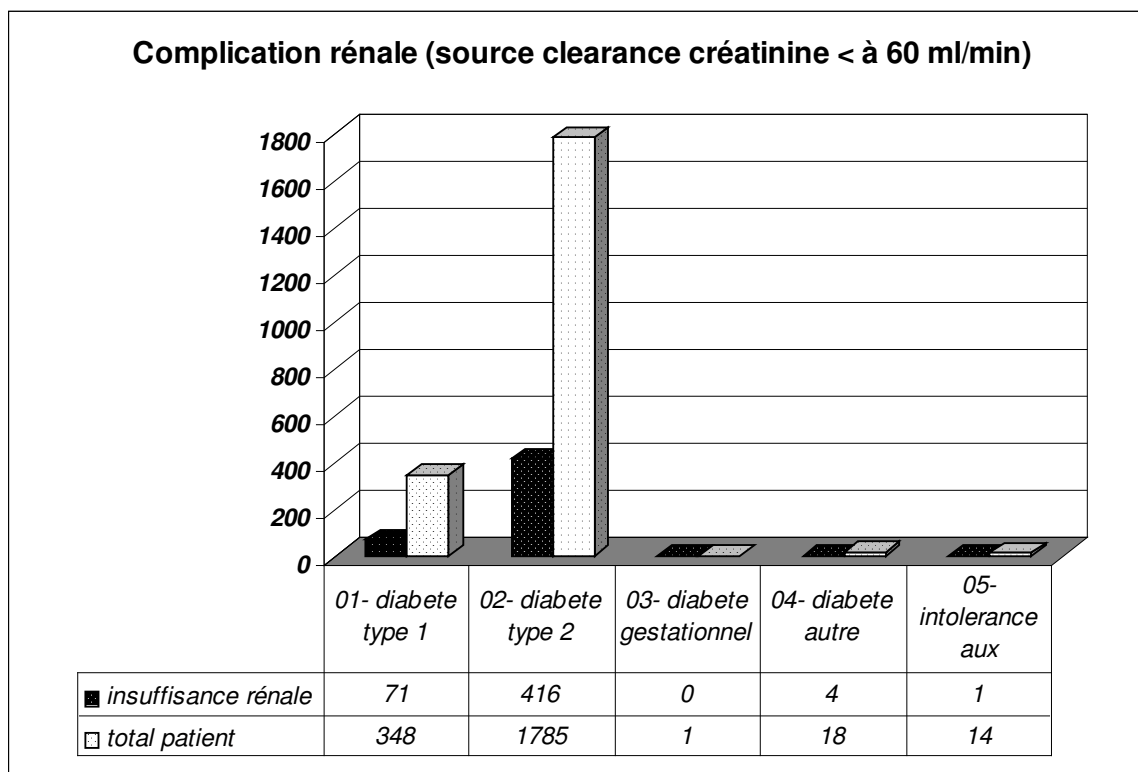
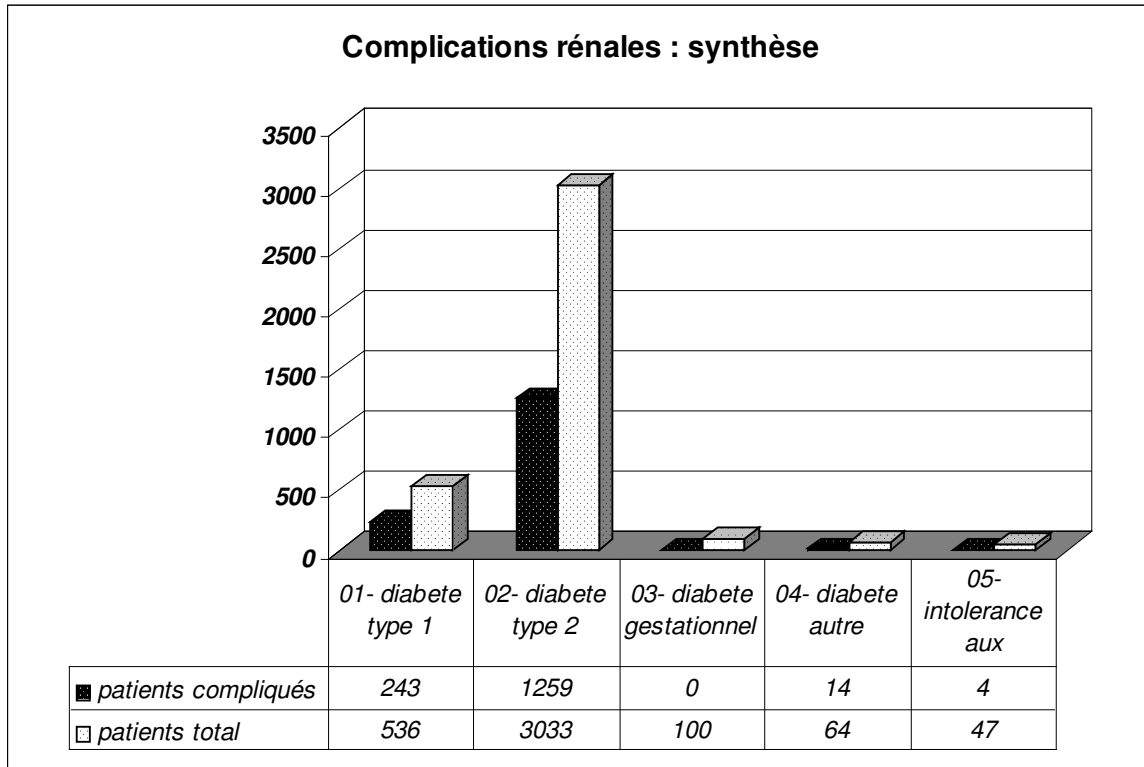


Figure 42 - Complications rénales sur le champ clearance de la créatinine.

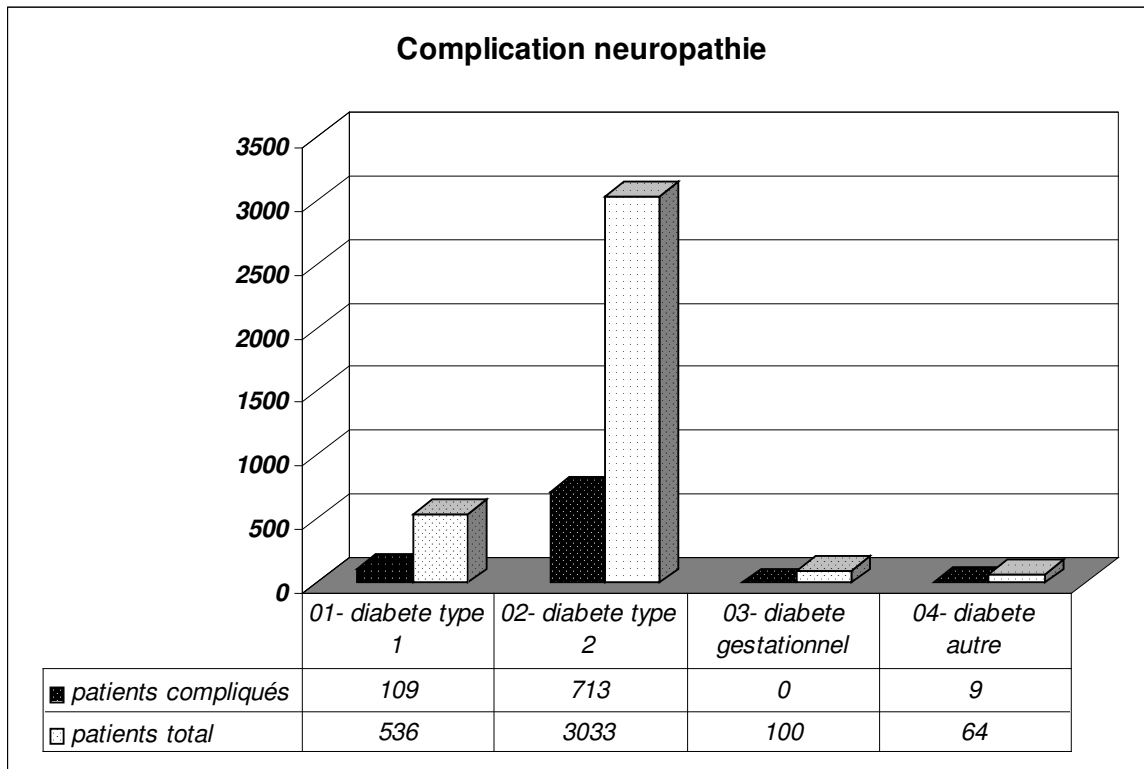
Afin de réaliser une synthèse pour déterminer l'ensemble des complications rénales, nous avons réalisé un cumul des complications notées en comptabilisant tous les patients ayant au moins un champ complication rénale positif (champs complication, champs albuminurie, champs clearance de la créatinine). Ce résultat est présenté dans la Figure 43.



**Figure 43 - Complications rénales synthèse.**

***c- Complications neurologiques***

La Figure 44 représente la répartition des complications neuropathies dénombrées sur le champ *complications*.



**Figure 44 - dénombrement des complications neuropathies**

#### ***d- Complications trophiques***

La Figure 45 représente la répartition des complications du pied dénombrées sur le champ *complications* ou vues en consultation de pied diabétique. Nous avons pu incrémenter le champ complication car il n'y a qu'un médecin spécialisé dans les consultations de pied diabétique, et ces consultations sont individualisables.

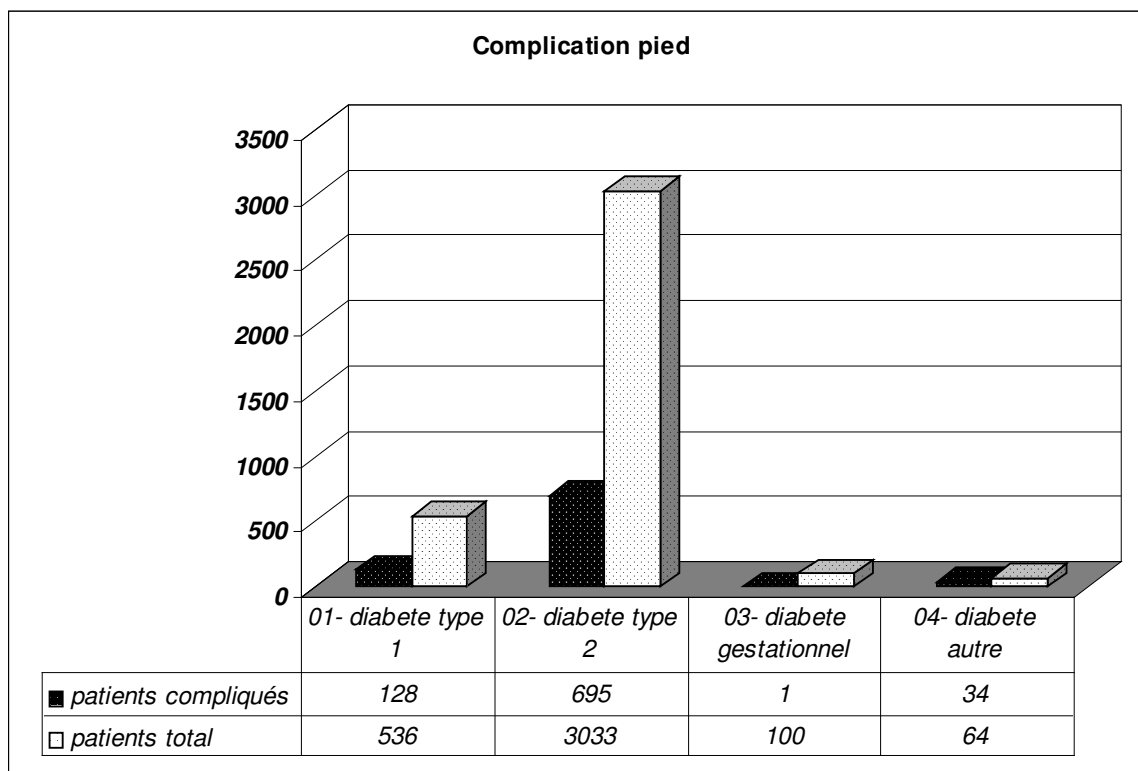


Figure 45 - dénombrement des complications de pied

La Figure 46 représente la répartition des complications d'artérite dénombrées sur le champ *complications* uniquement.

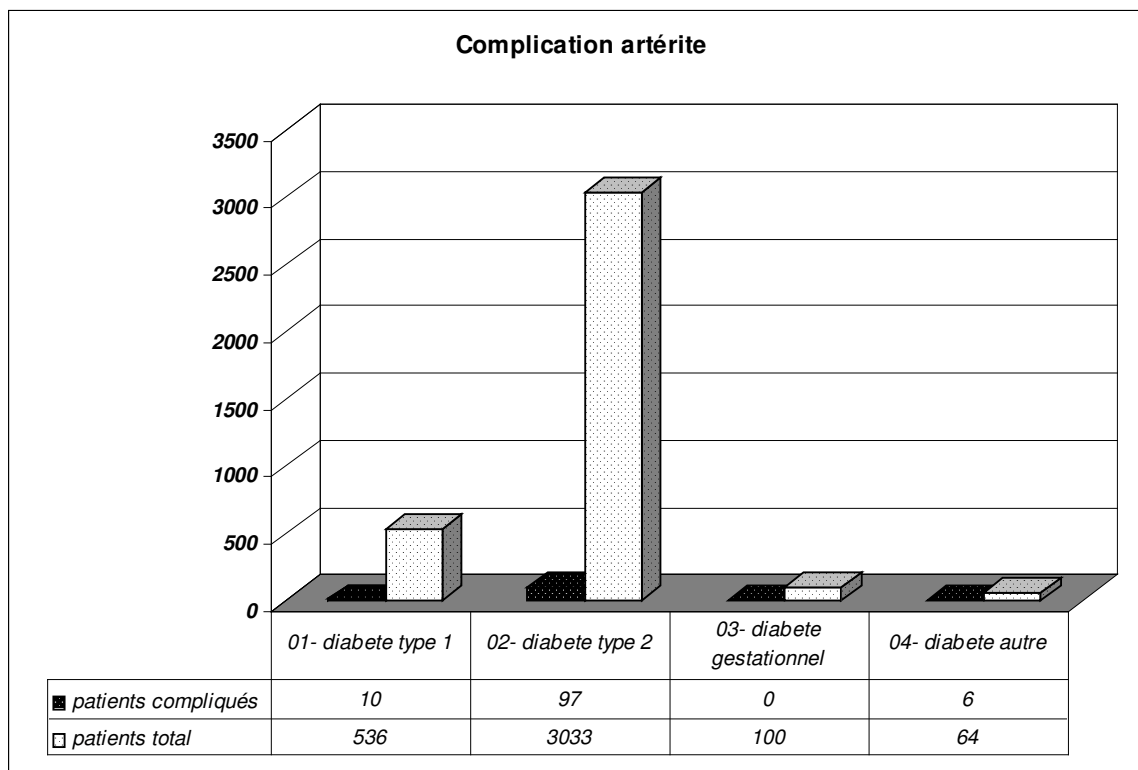


Figure 46 - dénombrement des complications d'artérite

### e- Dyslipidémie

La Figure 47 représente la répartition des complications de dyslipidémie dénombrées sur le champ *complications*.

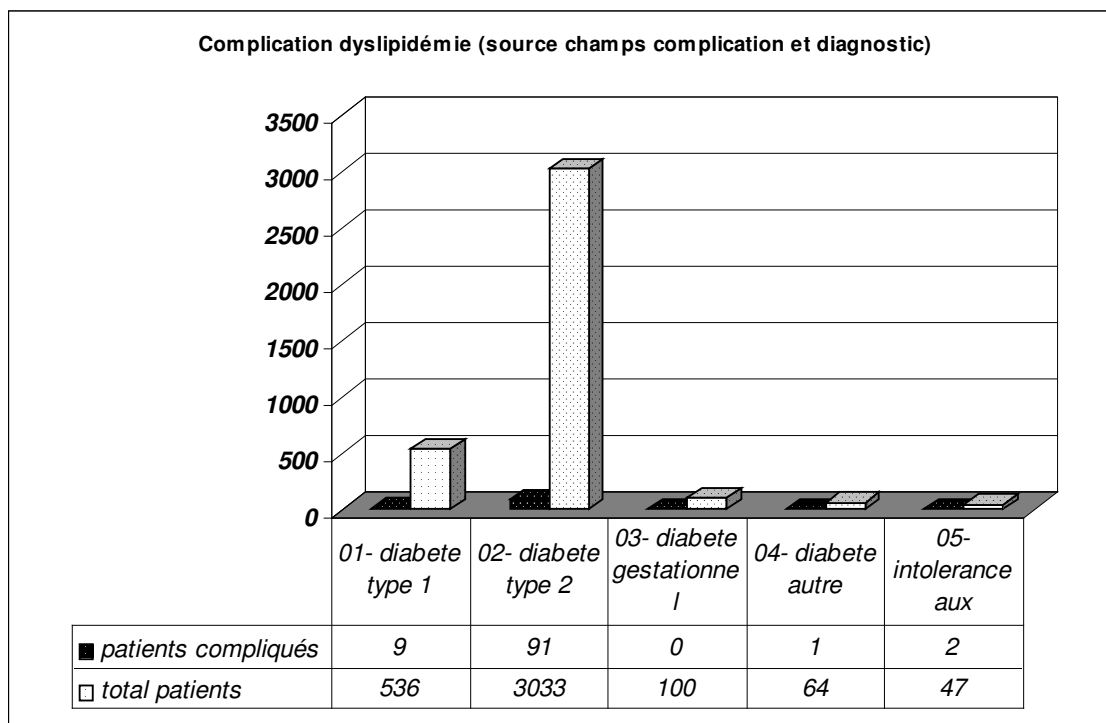
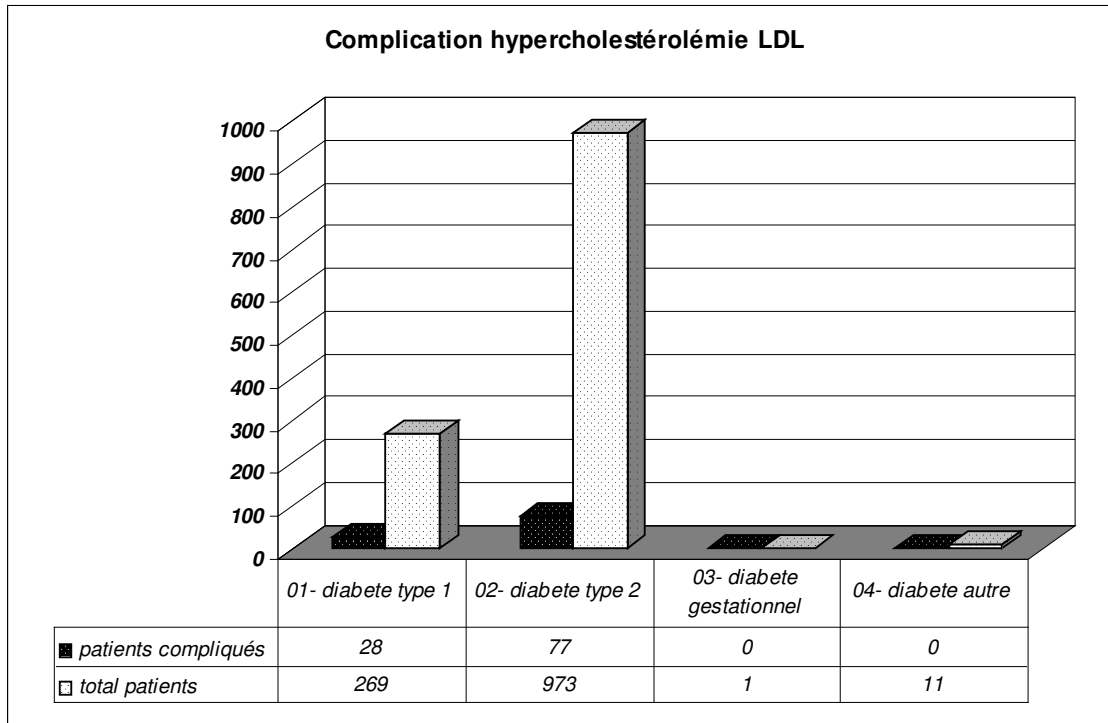


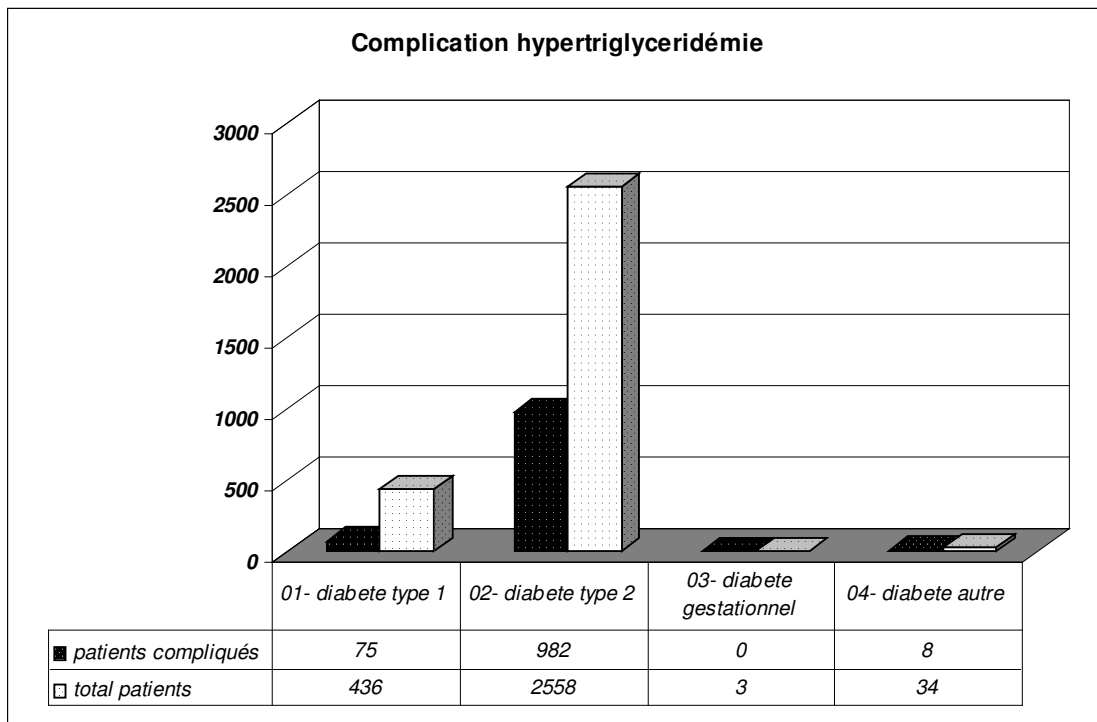
Figure 47 - dénombrement des complications de dyslipidémie sur le champ complications

La Figure 48 représente la complication d'hypercholestérolémie sur le champ LDL, où nous avons considéré comme complication toute valeur du LDL > 2g/l, sans tenir compte des facteurs de risque car nous ne cherchons pas une cholestérolémie cible.



**Figure 48 - dénombrement des complications hypercholestérolémie sur le champ LDL.**

La Figure 48 représente la complication d'hypertriglycéridémie sur le champ *triglycérides*, où nous avons considéré comme complication toute valeur des triglycérides > 2g/l, sans tenir compte des facteurs de risque.



**Figure 49 - dénombrements des complications hypertriglycéridémie sur le champ triglycérides.**

Afin de réaliser une synthèse pour déterminer l'ensemble des complications dyslipidémie, nous avons réalisé un cumul des complications notées en comptabilisant tous les patients ayant au moins un champ complication dyslipidémie positif (champs complication dyslipidémie; champs complication LDL ou champs complication triglycérides). Ce résultat est présenté dans la Figure 50.

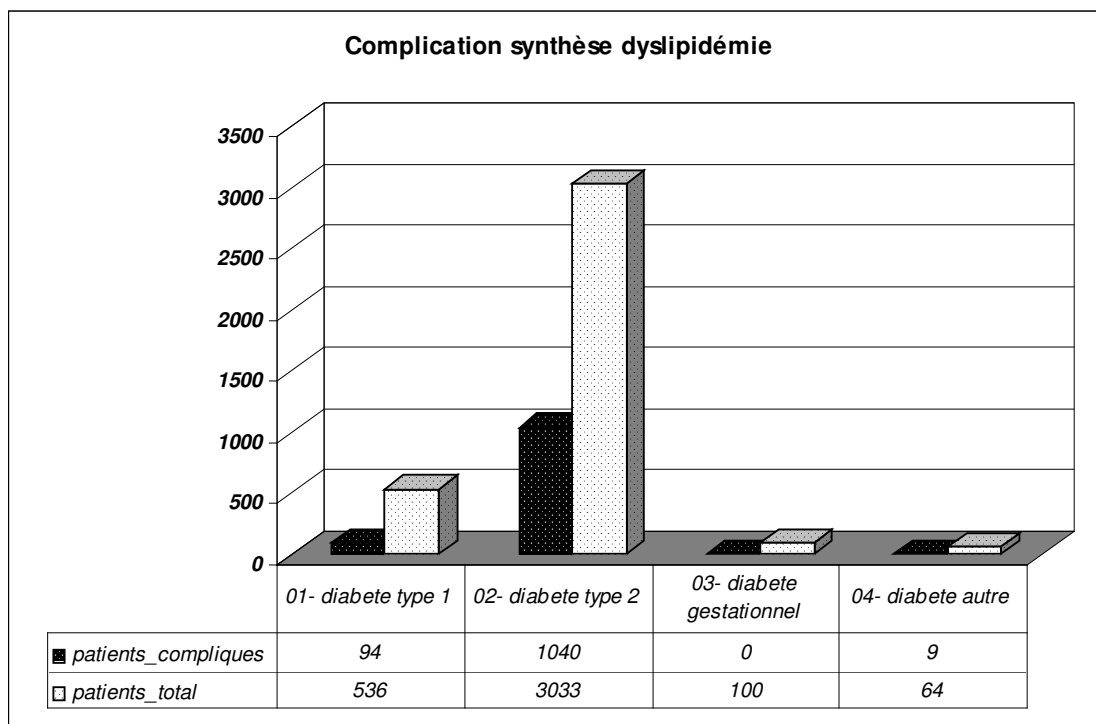


Figure 50 - synthèse des dénombrements des complications de dyslipidémie

### **f- Complications cardiaques**

La Figure 51 représente la répartition des complications de coronarite ou infarctus du myocarde dénombrées sur le champ *complications*.



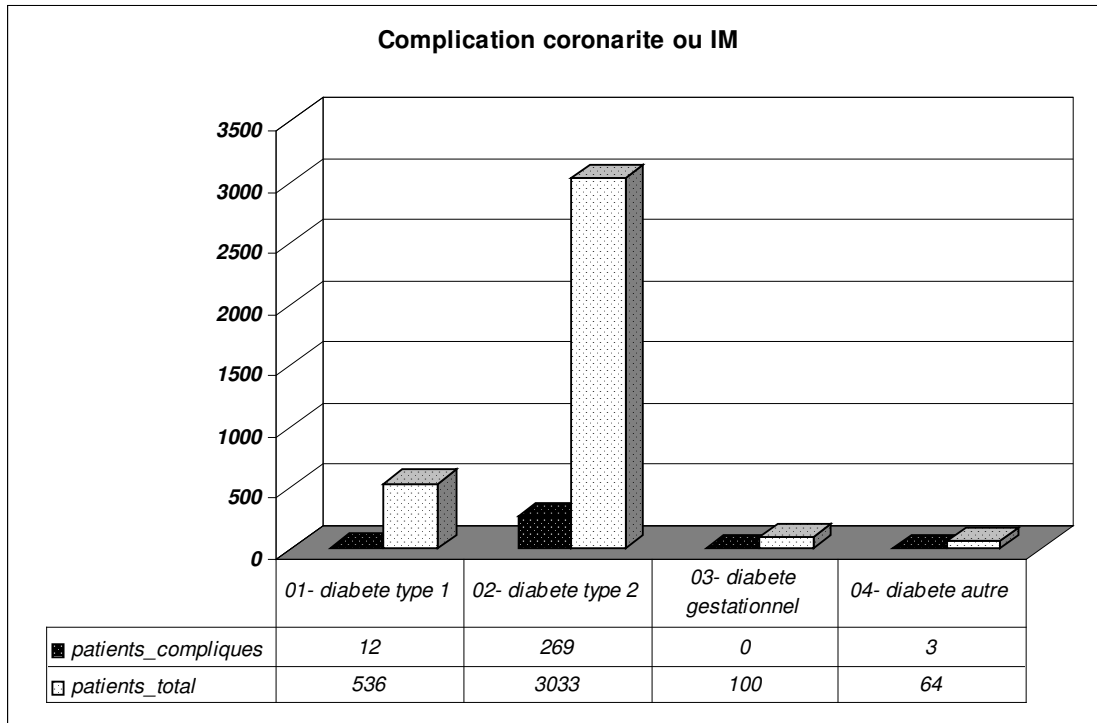


Figure 51 - dénombrements des complications coronarite ou infarctus du myocarde par type de diabète

### **g- Accident vasculaire cérébral**

La Figure 52 représente le dénombrement des complications d'AVC sur le champ complications.

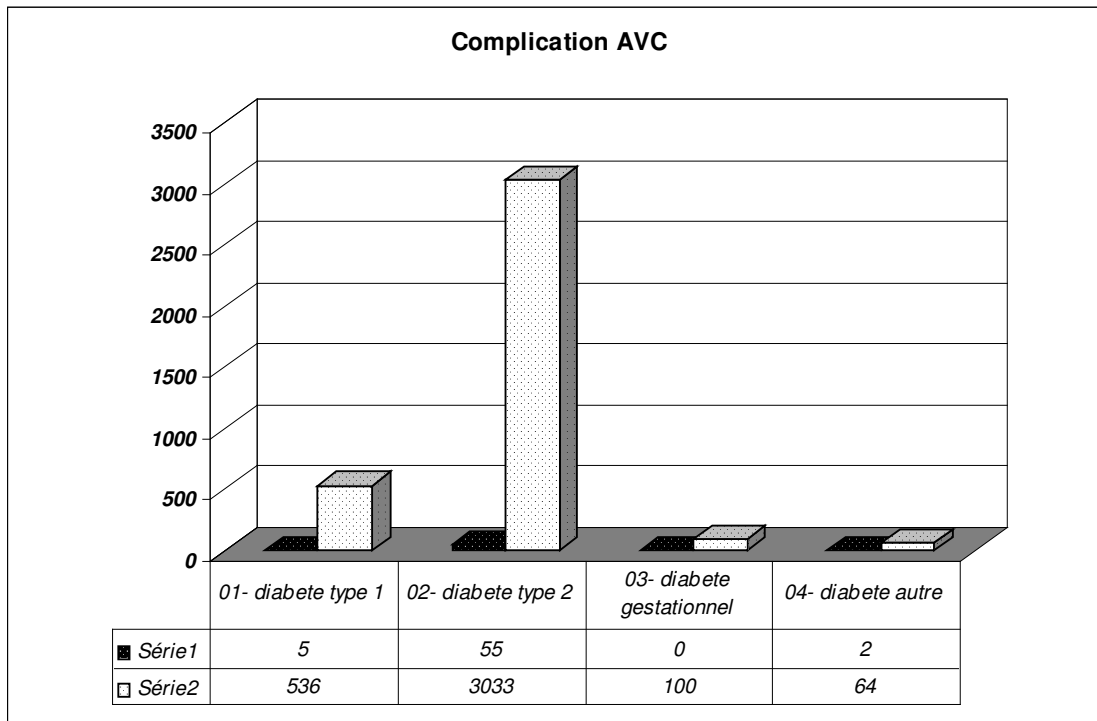
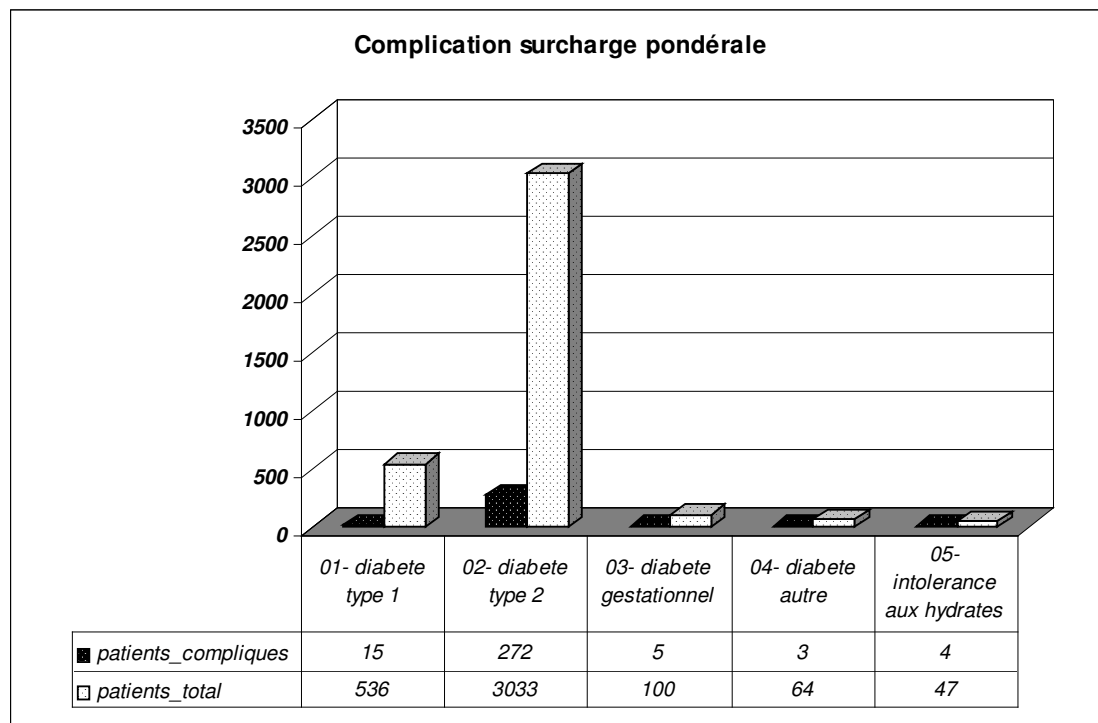


Figure 52 - dénombrement des complications d'AVC par type de diabète

### ***h- Surcharge pondérale***

La Figure 53 représente le dénombrement des complications de surcharge pondérale sur le champ *complications*. Il n'a malheureusement pas été possible de prendre en compte l'IMC car le poids et la taille n'ont pas été bien renseignés comme nous l'avons expliqué dans la partie matériel. Ce résultat minimise donc les surcharges pondérales.



**Figure 53 - dénombrement des complications surcharge pondérale par type de diabète**

### ***i- Hypertension artérielle***

La Figure 54 représente le dénombrement des patients désignés comme hypertendus dans le champ *HTA*.

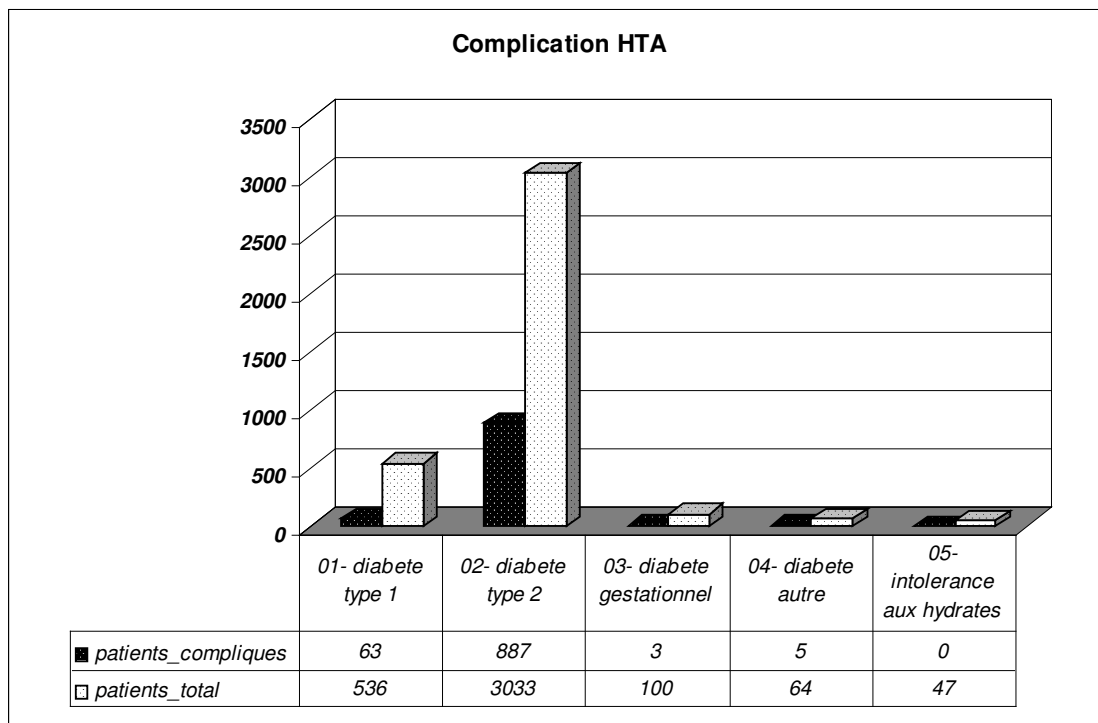


Figure 54 - dénombrement des complications d'HTA par type de diabète.

**j- BILAN, toutes complications**

La Figure 55 représente le dénombrement des patients ayant au moins une complication.

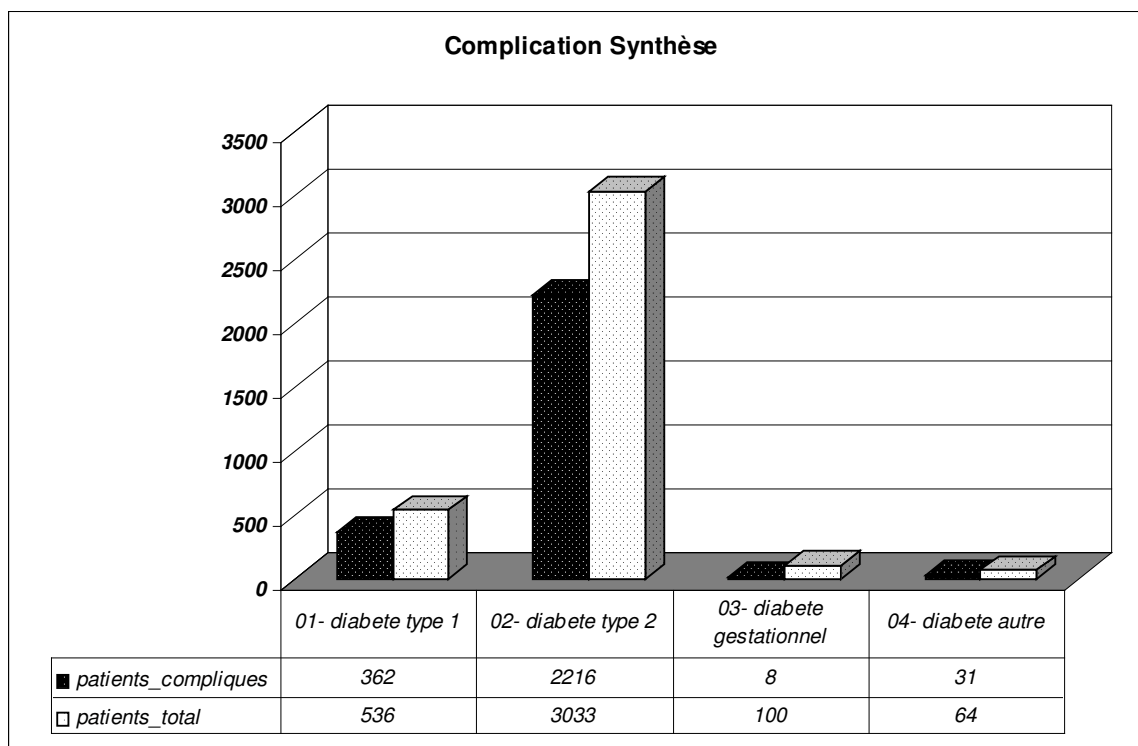


Figure 55 - synthèse des patients ayant une complication par type de diabète

## 8. Diabétiques de la file active - mode de traitement

Nous ne détaillerons pas le traitement des patients diabétiques. En effet, l'exhaustivité des champs de traitement n'est pas satisfaisante. De ce fait, nous devrions réagrèger les différentes valeurs à l'échelle du patient. Or le traitement d'un patient évolue avec d'une part sa maladie, d'autre part les progrès thérapeutiques. Ainsi, il paraît aberrant de considérer la somme des traitements de l'historique du patient. Enfin, en admettant que la base fût parfaite, nous aurions dû choisir une date pour chaque patient, et ce choix pose un problème méthodologique.

## 9. Diabétiques et HMJ - évolution du recrutement

### a- Evolution de la première HbA1C

#### 1) Tous diabétiques et HMJ

Nous représentons la valeur de l'HbA1c lors du recrutement du patient : la moyenne par année (Figure 56) et la distribution complète (Figure 57). On note une forte diminution de la première HbA1C entre 1992 et 1994, sur fond de diminution globale.

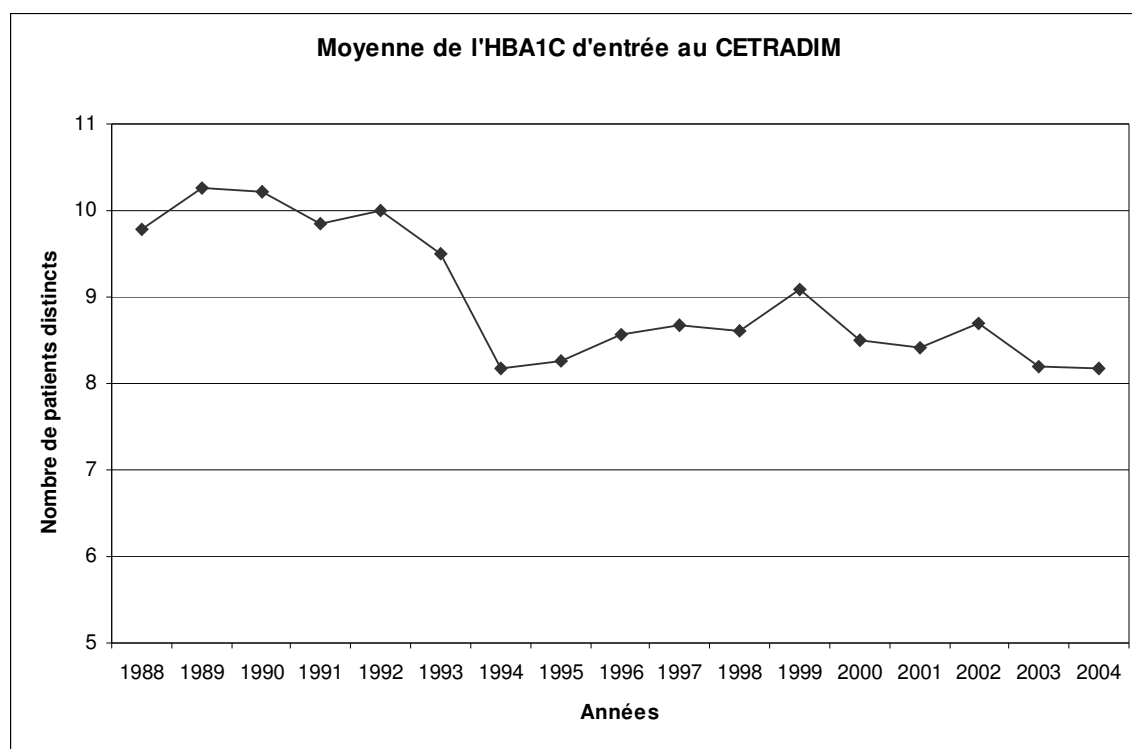


Figure 56 - Evolution de la première HbA1C en fonction du temps

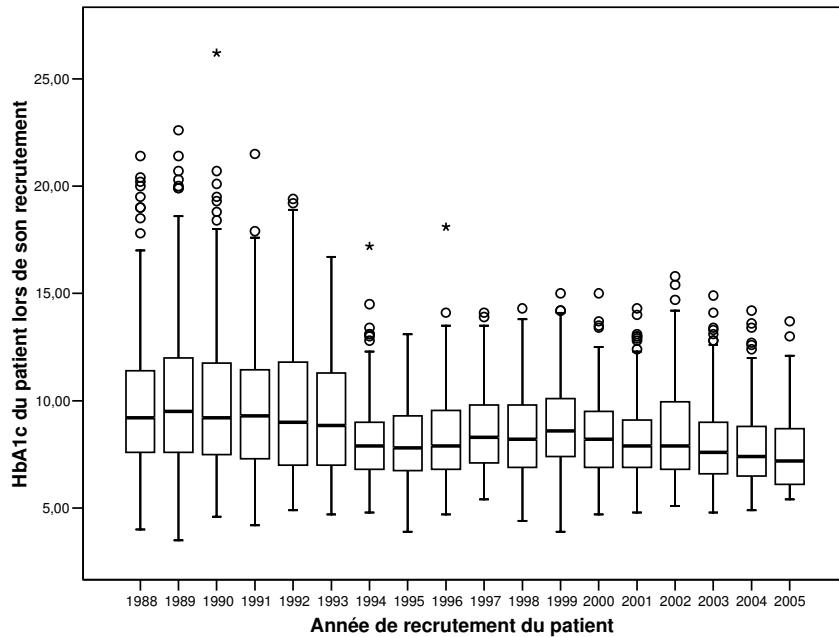


Figure 57 - box plots - distribution des HbA1C lors du recrutement en fonction de l'année de recrutement chez l'ensemble des diabétiques

Nous représentons ci-après (Figure 58) le nuage de points correspondant à ces mesures. Ce graphique est assorti d'une droite d'ajustement linéaire qui suggère une relation linéaire décroissante avec le temps.

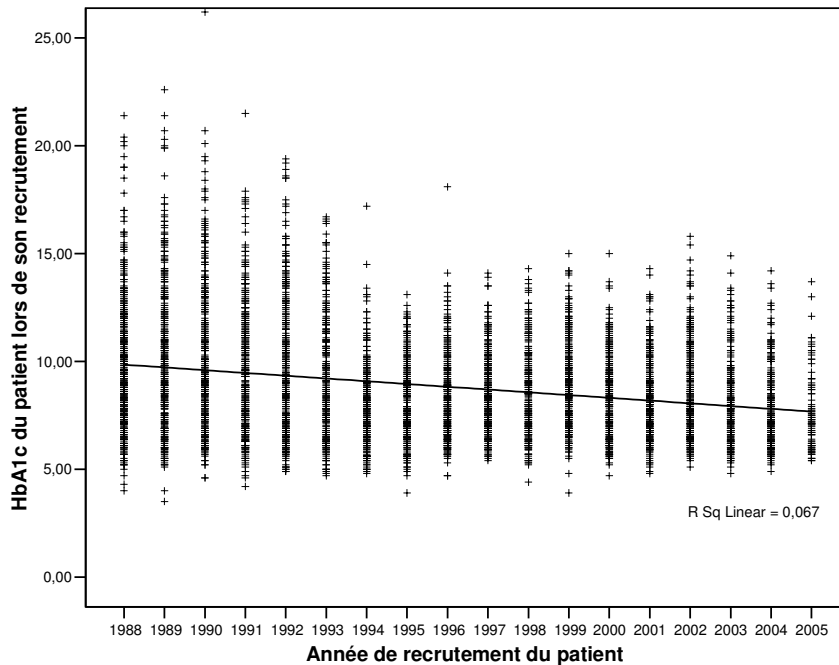


Figure 58 - Nuage de points - HbA1C en fonction de l'année de recrutement sur l'ensemble des diabétiques.

Afin de tester cet effet linéaire décroissant avec le temps, nous réalisons une régression linéaire simple.

**Coefficients(a)**

Model		Unstandardized Coefficients		Standardized Coefficients	t	Sig.	95% Confidence Interval for B	
		B	Std. Error	Beta			Lower Bound	Upper Bound
1	(Constant)	264.650	13.107		20.191	.000	238.954	290.345
	annee	-.128	.007	-.259	-19.514	.000	-.141	-.115

a Dependent Variable: premiere\_hba1c

**ANOVA(b)**

Model		Sum of Squares	df	Mean Square	F	Sig.
1	Regression	2291.444	1	2291.444	380.808	.000(a)
	Residual	31879.776	5298	6.017		
	Total	34171.221	5299			

a Predictors: (Constant), année

b Dependent Variable: premiere\_hba1c

**Tableau 12 - régression linéaire simple - évolution de l'HbA1c lors du recrutement en fonction du temps**

Plus le temps passe, plus les patients sont recrutés avec une HbA1c basse. Ceci témoigne d'un changement des pratiques de l'ensemble des médecins correspondants : les diabétiques sont probablement dépistés plus tôt et sont adressés plus rapidement au CETRADIMN pour une prise en charge thérapeutique. L'année de recrutement a un effet linéaire décroissant sur l'HbA1c lors du recrutement (IC95 de la pente strictement négatif, ANOVA de la régression avec  $p < 0.001$ ).

Nous détaillons ci-après l'évolution en distinguant diabète type 1 et type 2 (Figure 59).

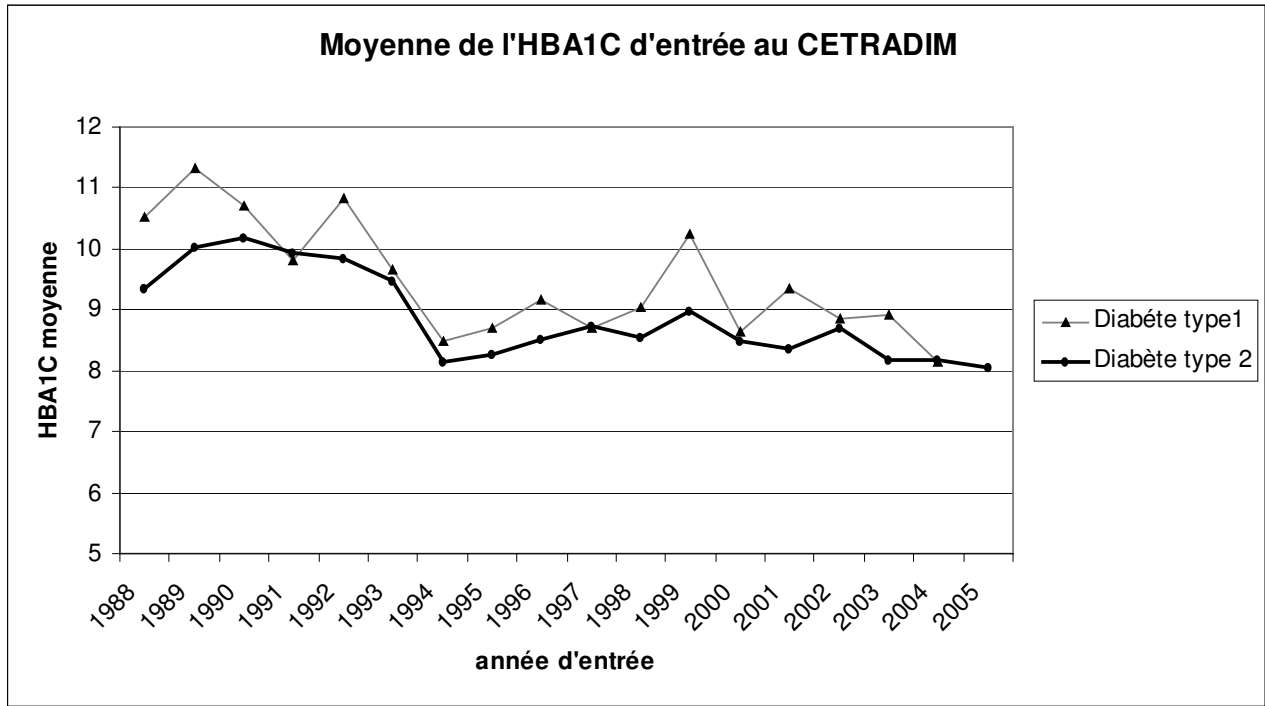


Figure 59 - Evolution de l'HBA1C de recrutement des diabétiques type 1 et type 2 en fonction de l'année de recrutement

### 2) diabètes type 2

Nous représentons l'évolution de l'HbA1c lors du recrutement pour les seuls diabétiques de type 2 (Figure 60).

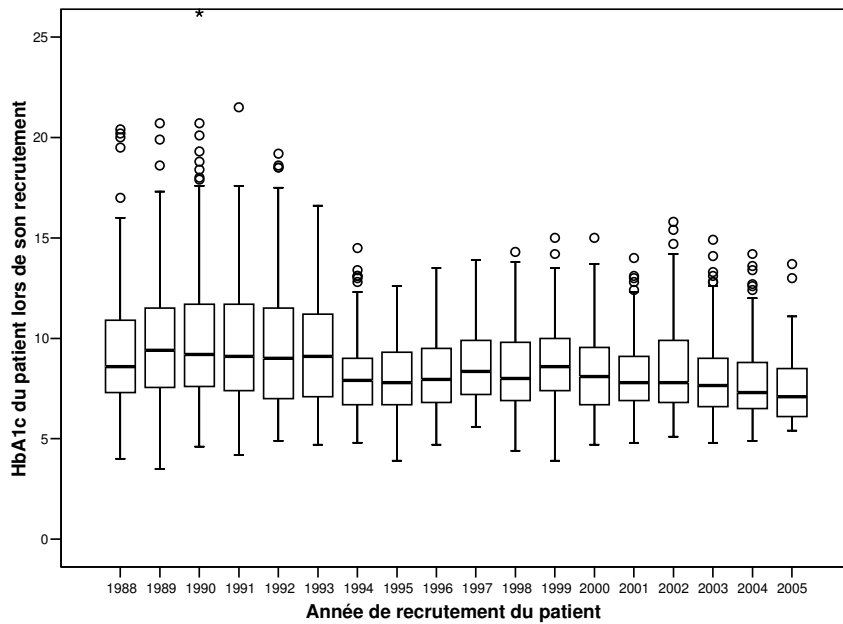


Figure 60 - Box plot - distribution des HbA1C lors du recrutement en fonction de l'année de recrutement chez les diabétiques de type 2

On observe un décrochage en 1994, sur fond de décroissance progressive des HbA1c. Nous représentons ces mêmes mesures sur un nuage de points complétés par une droite d'ajustement linéaire (Figure 61).

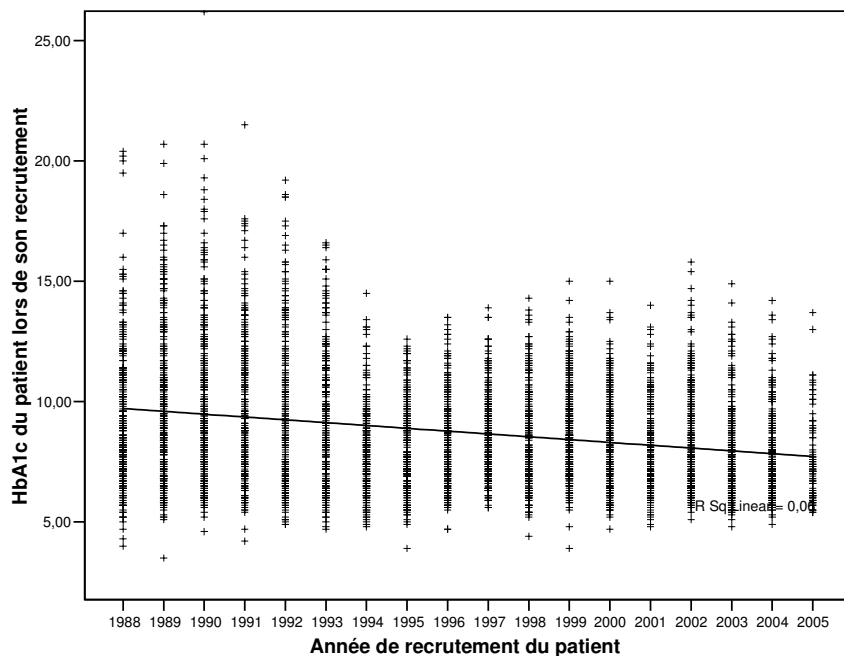


Figure 61 - Nuage de points - HbA1C en fonction de l'année de recrutement sur les diabétiques de type 2.

Comme précédemment, nous réalisons une régression linéaire simple depuis l'année de recrutement (Tableau 13).

Coefficients(a)								
Model		Unstandardized Coefficients		Standardized Coefficients	t	Sig.	95% Confidence Interval for B	
		B	Std. Error	Beta			Lower Bound	Upper Bound
1	(Constant)	242.934	13.879		17.504	.000	215.724	270.144
	annee	-.117	.007	-.245	-16.872	.000	-.131	-.104

a Dependent Variable: premiere\_hba1c

ANOVA(b)						
Model		Sum of Squares	df	Mean Square	F	Sig.
1	Regression	1580.465	1	1580.465	284.673	.000(a)
	Residual	24811.273	4469	5.552		
	Total	26391.737	4470			

a Predictors: (Constant), annee  
b Dependent Variable: first\_hb\_glc

Tableau 13 – Régression simple - Evolution de l'HbA1c des patients lors de leur recrutement en fonction de l'année de recrutement

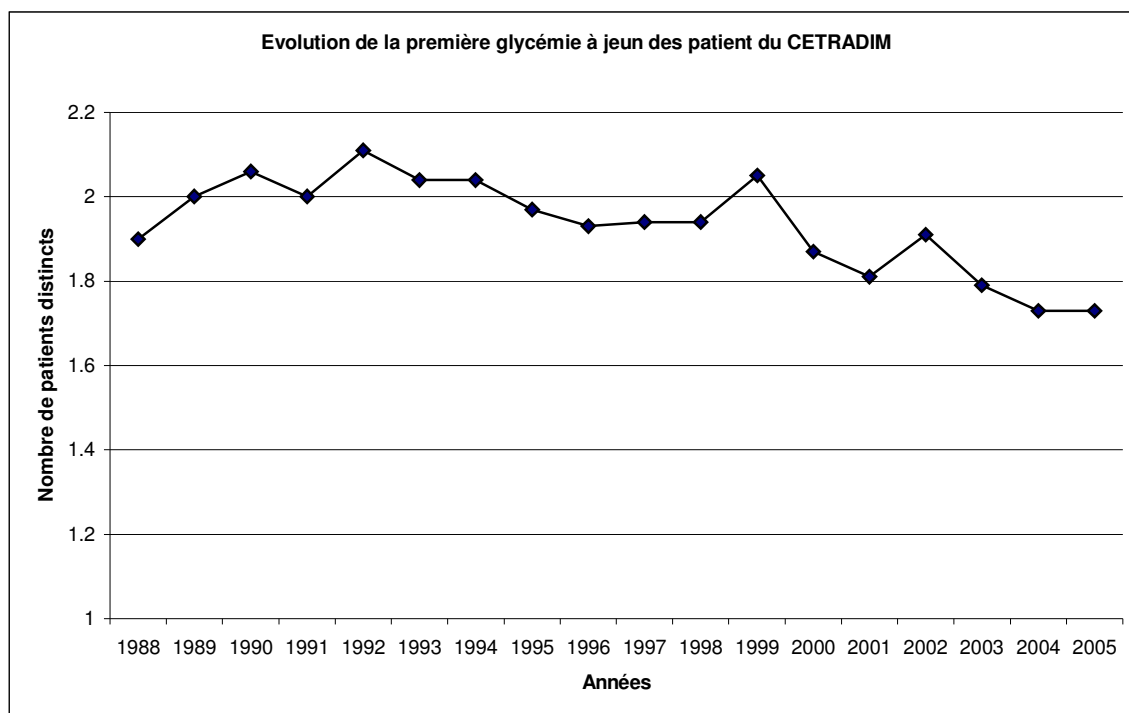


Comme précédemment, nous montrons une décroissance significative au fil des ans (IC95 de la pente strictement négatif, ANOVA de la régression  $p < 0.001$ ).

### ***b- Evolution de la première glycémie à jeun***

#### ***1) Tous patients diabétiques et HMJ***

Nous représentons sur la Figure 62 l'évolution de la moyenne de la glycémie à jeun des patients diabétiques (tous confondus) mesurée lors de leur recrutement, année après année. On remarque une tendance générale à la baisse.



**Figure 62 - Evolution de la première glycémie à jeun en fonction de l'année de recrutement chez l'ensemble des diabétiques**

Ces mesures sont ensuite représentées à l'aide de box plots (Figure 63) et d'un nuage de points assorti d'un ajustement linéaire (Figure 73).

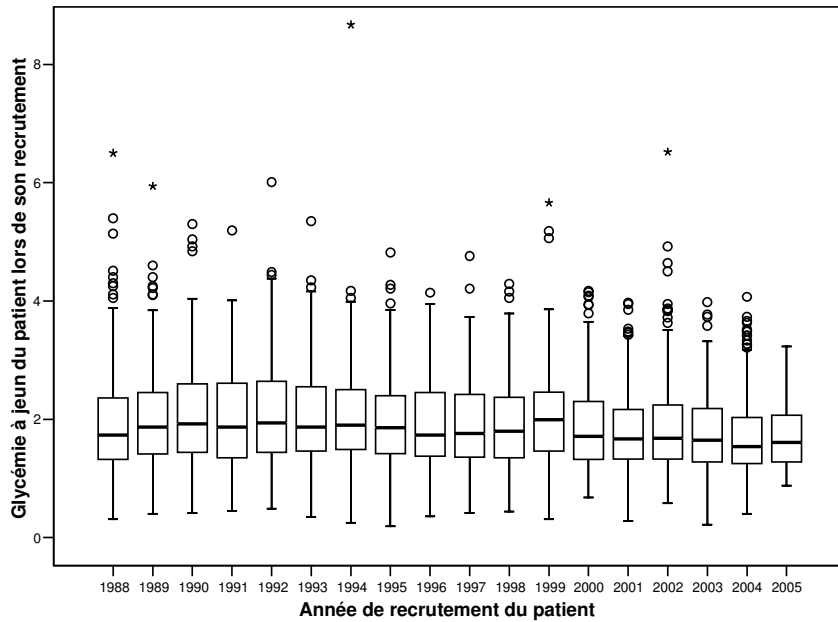


Figure 63 – Box plots - distribution des premières glycémies à jeun selon l'année de recrutement, tous diabétiques

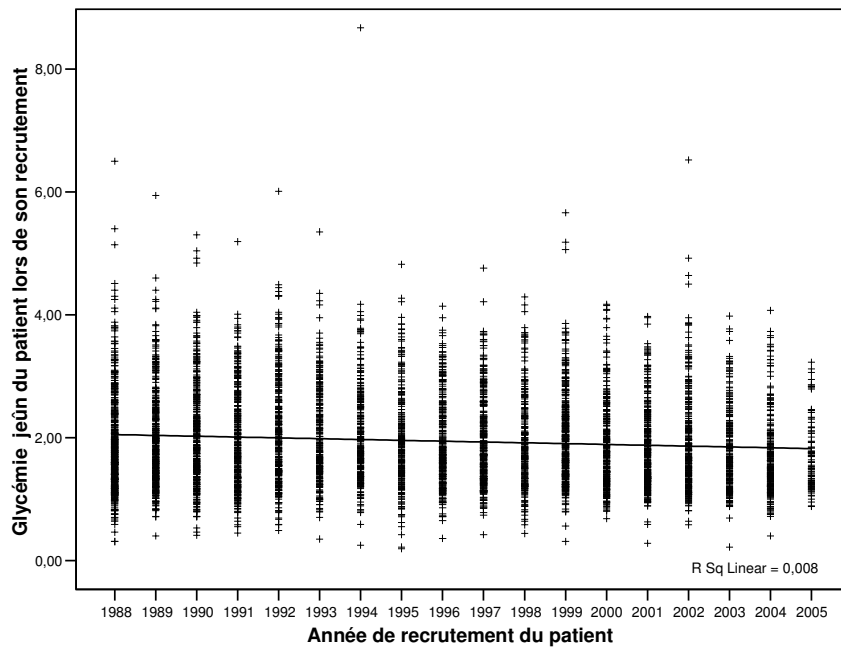


Figure 64 : nuage de points - distribution des premières glycémie à jeun selon l'année de recrutement, tous diabétiques

Cet ajustement suggère une relation linéaire décroissante entre la première glycémie à jeun et l'année de recrutement. Nous la testons à l'aide d'une régression linéaire simple.

**Coefficients(a)**

Model		Unstandardized Coefficients		Standardized Coefficients	t	Sig.	95% Confidence Interval for B	
		B	Std. Error	Beta			Lower Bound	Upper Bound
1	(Constant)	28.696	4.115		6.974	.000	20.630	36.762
	annee	-.013	.002	-.089	-6.501	.000	-.017	-.009

a Dependent Variable: premiere\_glycemie\_a\_jeun

**ANOVA(b)**

Model		Sum of Squares	df	Mean Square	F	Sig.
1	Regression	25.185	1	25.185	42.263	.000(a)
	Residual	3180.956	5338	.596		
	Total	3206.141	5339			

a Predictors: (Constant), annee

b Dependent Variable: premiere\_glycemie\_a\_jeun

**Tableau 14 - régression linéaire simple - évolution de la première glycémie à jeun selon l'année de recrutement, tous diabétiques**

La régression linéaire simple montre que la glycémie à jeun des patients lors du recrutement diminue significativement au fil des ans (IC95 de la pente négatif, ANOVA de la régression  $p < 0.001$ ).

Bien que la glycémie à jeun ne soit pas un marqueur fiable de l'équilibre glycémique, par sa forte variabilité intra-journée, cette évolution témoigne de l'évolution des pratiques : les diabétiques sont dépistés et pris en charge plus tôt dans l'évolution de leur maladie.

Pour des raisons physiopathologiques, la glycémie est encore plus variable chez les patients en diabète de type 1, c'est la raison pour laquelle nous ne les détaillerons pas plus bas (Figure 74).

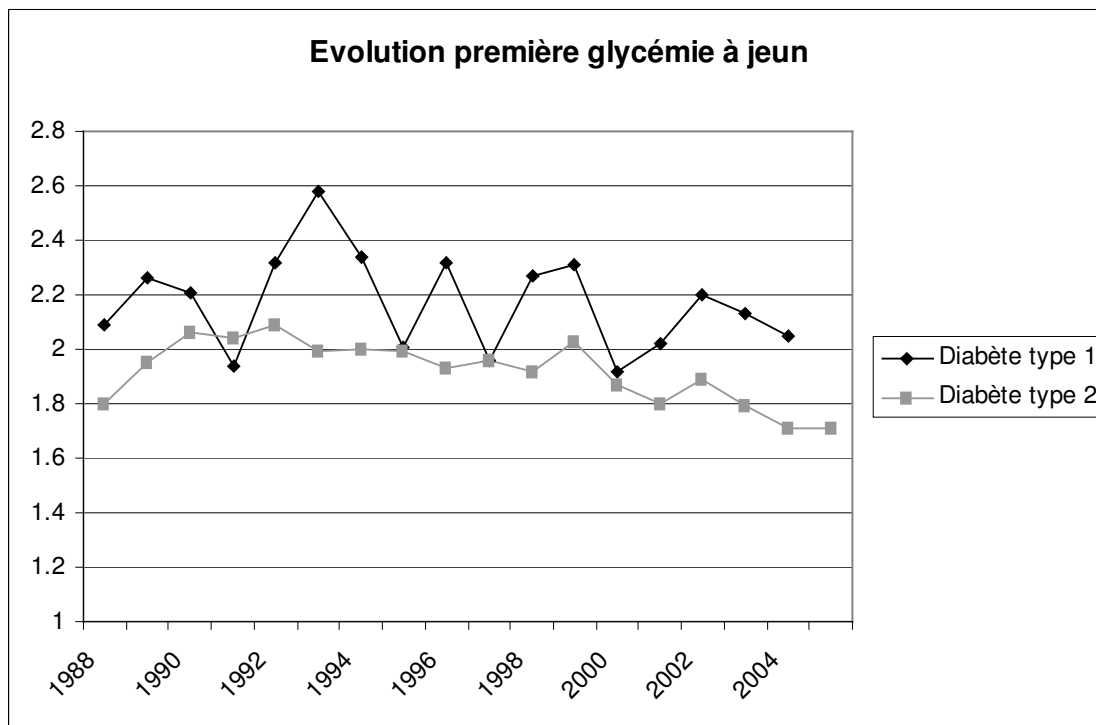


Figure 65 - Evolution de la première glycémie à jeun en fonction de l'année de recrutement des diabétiques de type1 et de type2.

## 2) patients diabétiques type 2

Nous représentons sur la Figure 75 et la Figure 67 l'évolution des distributions de la glycémie à jeun des patients diabétiques de type 2 mesurée lors de leur recrutement, année après année. On remarque une tendance générale à la baisse. L'ajustement linéaire suggère une relation linéaire décroissante.

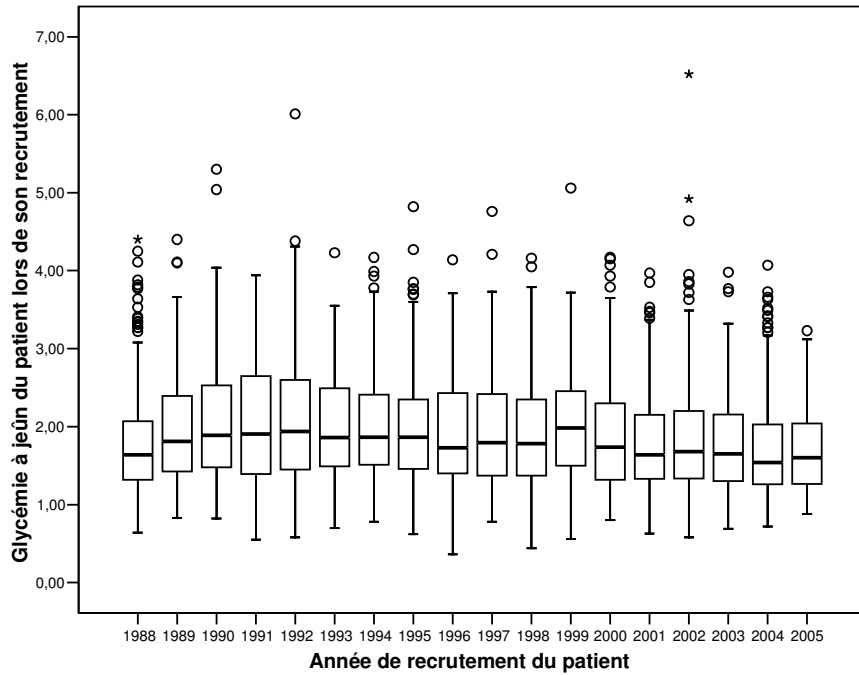


Figure 66 - Box Plots - distribution des première glycémies à jeun selon l'année de recrutement, diabétiques de type2

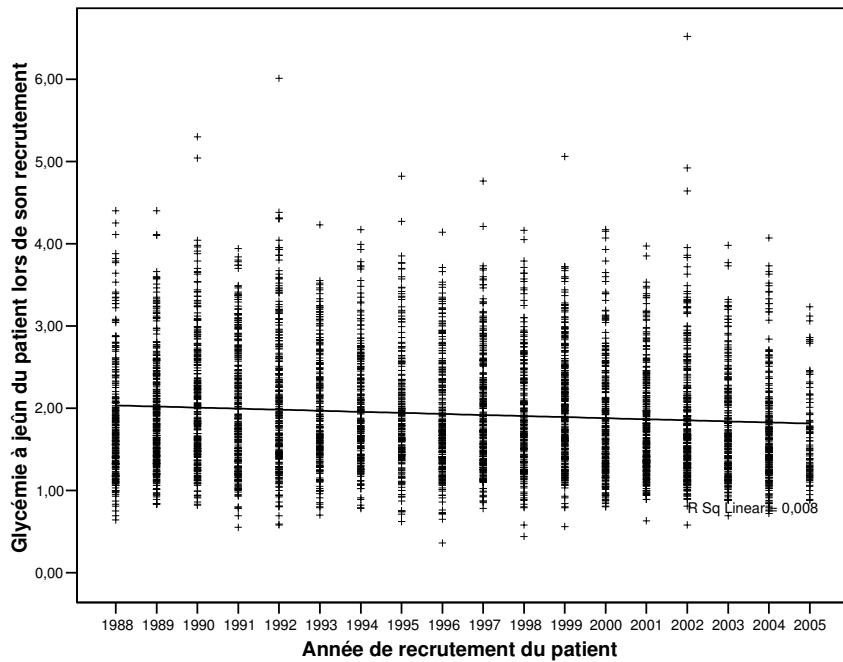


Figure 67 - Nuage de points - distribution des première glycémies à jeun selon l'année de recrutement, diabétiques de type2

Nous testons cette hypothèse à l'aide d'une régression linéaire simple.

**Coefficients(a)**

Model		Unstandardized Coefficients		Standardized Coefficients	t	Sig.	95% Confidence Interval for B	
		B	Std. Error	Beta			Lower Bound	Upper Bound
1	(Constant)	27.662	4.199		6.588	.000	19.430	35.895
	annee	-.013	.002	-.091	-6.128	.000	-.017	-.009

a Dependent Variable: premiere\_glycemie\_a\_jeun

**ANOVA(b)**

Model		Sum of Squares	df	Mean Square	F	Sig.
1	Regression	19.166	1	19.166	37.557	.000(a)
	Residual	2295.430	4498	.510		
	Total	2314.596	4499			

a Predictors: (Constant), annee

b Dependent Variable: premiere\_glycemie\_a\_jeun

**Tableau 15 - régression linéaire simple - évolution de la première glycémie à jeun selon l'année de recrutement, diabétiques type 2**

La régression linéaire montre que la glycémie à jeun lors du recrutement diminue significativement au fil des ans (IC95 de la pente strictement négatif, ANOVA de la régression  $p < 0.001$ ). Comme mentionnée plus haut, cette évolution semble découler directement d'une modification des pratiques de dépistage et de prise en charge des patients diabétiques de type 2.

***c- Parmi les diabétiques type 1, évolution de la proportion de patients vierges***

La Figure 68 présente, année par année, la proportion de patients diabétiques de type 1 se présentant au CERADIMN sans aucun traitement par insuline. Cette figure met en avant une proportion croissante de diabétiques de type 1 adressés au CETRADIMN avant prise en charge thérapeutique.

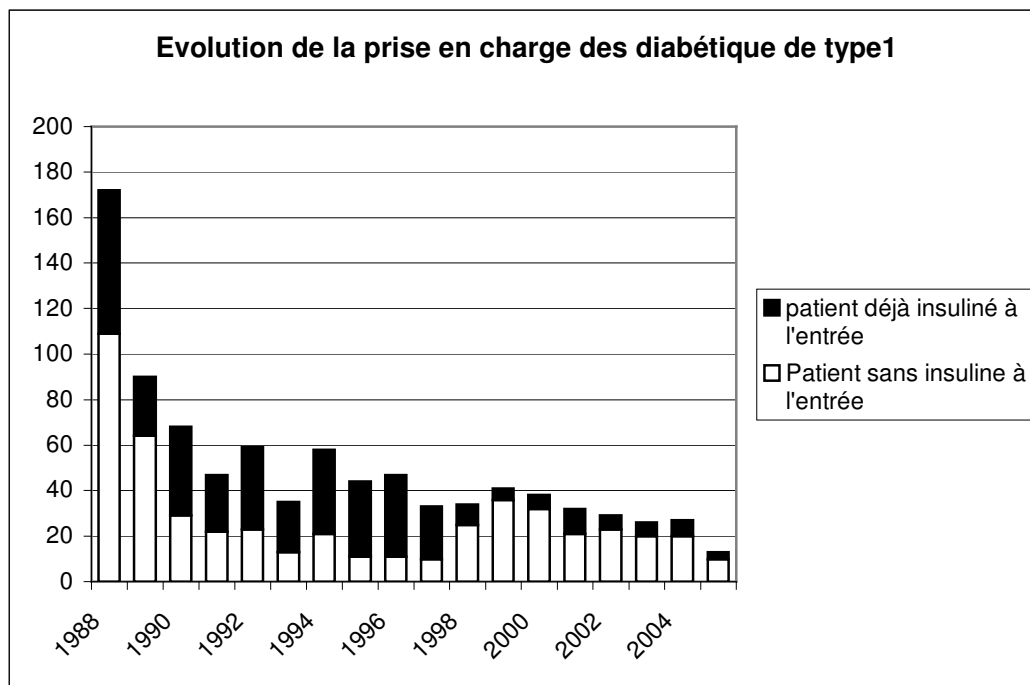


Figure 68 - Evolution de la proportion des patients diabétiques de type 1 insulinés et non insulinés à l'entrée.

## **B. Validité du dosage de l'hémoglobine glycosylée**

### **1. Problématique**

En préalable à la conception et au test d'indicateurs d'utilité fondés sur la mesure de l'hémoglobine glycosylée / hémoglobine glyquée / HbA1c, il importe de s'assurer de la stabilité de la méthode de dosage.

La méthode de dosage utilisée a changé au cours du recueil des données que nous étudions :

- jusqu'au 01/01/1994 : dosage de l'hémoglobine glyquée par chromatographie par échange d'ions sur micro colonnes Eurobio
- du 01/01/1994 au 01/04/2004 : dosage de l'HbA1c par chromatographie basse pression sur l'automate Glycollab (Instrumentation Laboratory)
- à partir du 01/04/2004 : dosage immunologique de l'HbA1c par la technique Tina-quant sur Hitachi (Roche)

Pour chacune des deux dates charnières, nous isolons les résultats obtenus dans les 50 jours précédents (« avant ») et dans les 50 jours suivants (« après ») et réalisons une comparaison. Le nombre de 50 est choisi arbitrairement, ni trop faible (afin d'éviter les biais de sélection et d'avoir suffisamment de mesures) ni trop important (afin d'éviter les changements de pratiques médicales et de recrutement).

## 2. Avant et après le 01/01/1994

Le groupe « avant » est constitué des 134 mesures réalisées dans les 50 jours précédant le 01/04/2005, le groupe « après » est constitué des 142 mesures réalisées dans les 50 jours suivant le 01/04/2004.

La distribution des mesures réalisées est la suivante :

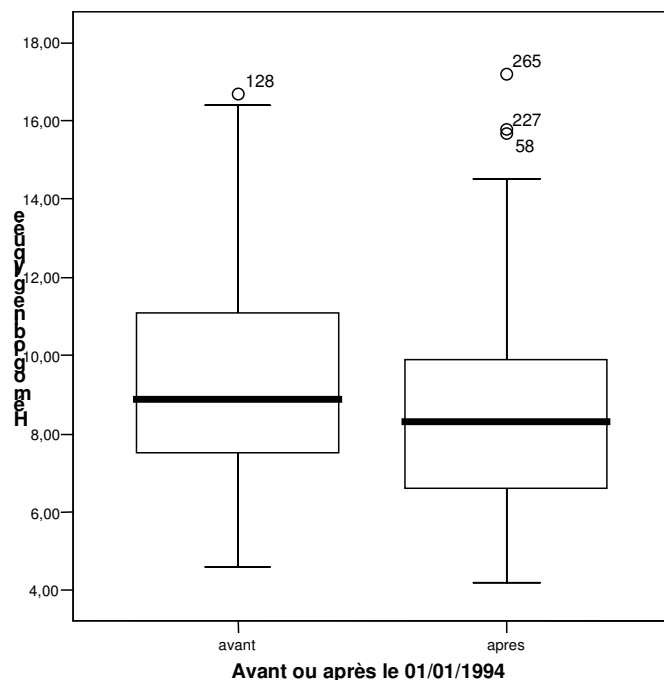


Figure 69 - Box plots - distribution des hémoglobines glyquées, 50 jours avant et après le 01/01/1994

Nous réalisons un test de Student sur échantillons indépendants afin de comparer les moyennes :

Group Statistics					
	groupe	N	Mean	Std. Deviation	Std. Error Mean
hb_glc	avant	134	9.3687	2.48494	.21467
	après	142	8.7035	2.50408	.21014

Independent Samples Test										
		Levene's Test for Equality of Variances		t-test for Equality of Means						
		F	Sig.	t	df	Sig. (2-tailed)	Mean Difference	Std. Error Difference	95% Confidence Interval	
									Lower	Upper
hb_glc	Equal variances assumed	.005	.943	2.214	274	.028	.66514	.30047	.07362	1.25665
	Equal variances not assumed			2.214	273.3	.028	.66514	.30040	.07375	1.25653

Tableau 16 – test de Student - comparaison des moyennes des hémoglobines glyquées 50 jours avant et après le 01/01/1994



Le test de Levene ne montre pas de différence significative entre les variances des deux groupes ( $p=0.943$ ). Le test de Student montre une différence significative entre les moyennes des deux groupes ( $p=0.028$ ).

Il faut donc considérer que le changement de méthode de dosage survenu en 1994 modifie les résultats obtenus.

### 3. Avant et après le 01/04/2004

Le groupe « avant » est constitué des 203 mesures réalisées dans les 50 jours précédant le 01/04/2005, le groupe « après » est constitué des 164 mesures réalisées dans les 50 jours suivant le 01/04/2004.

La distribution des mesures réalisées est la suivante :

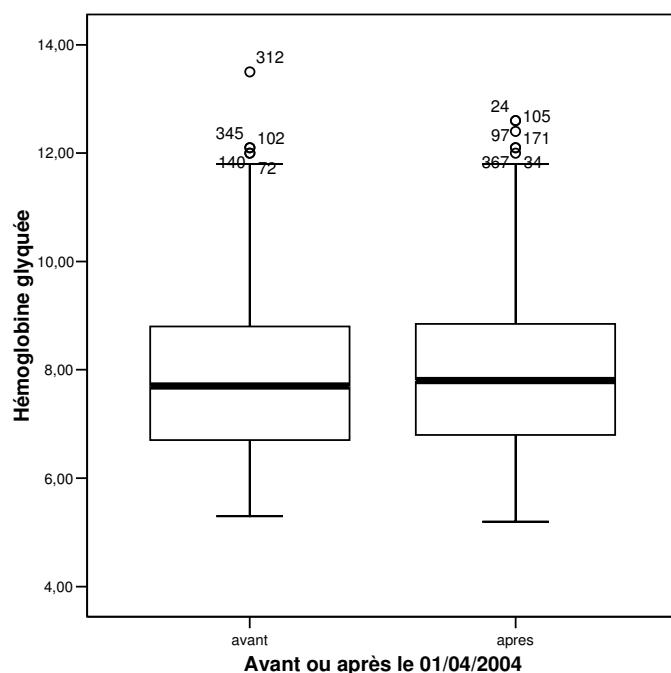


Figure 70 - Box plots - distribution des hémoglobines glyquées, 50 jours avant et après le 01/04/2004

Nous réalisons un test de Student sur échantillons indépendants afin de comparer les moyennes :

Group Statistics										
	groupe	N	Mean	Std. Deviation	Std. Error Mean					
hb_glc	avant	203	7.9596	1.58100	.11096					
	apres	164	7.9866	1.59056	.12420					

Independent Samples Test										
		Levene's Test for Equality of Variances		t-test for Equality of Means						
		F	Sig.	t	df	Sig. (2-tailed)	Mean Difference	Std. Error Difference	95% Confidence Interval	
									Lower	Upper
hb_glc	Equal variances assumed	.012	.911	-.162	365	.871	-.02698	.16644	-.35429	.30033
	Equal variances not assumed			-.162	348.1	.871	-.02698	.16655	-.35455	.30059

Tableau 17 – test de Student - comparaison des moyennes des hémoglobines glyquées avant et après le 01/04/2004

Le test de Levene ne montre pas de différence significative entre les variances des deux groupes ( $p=0.911$ ). Le test de Student ne montre pas de différence significative entre les moyennes des deux groupes ( $p=0.871$ ).

Compte tenu du contexte biologique, nous considérerons que le changement de méthode de dosage survenu en 2004 ne modifie pas les résultats obtenus.

### ***C. Indicateurs d'utilité***

#### ***1. Diabétiques de type 2 - suivi sur un an (recommandations ANAES-HAS)***

Le Tableau 18, résume par année l'évolution du taux de réalisation d'examens selon les recommandations de la HAS. La dernière ligne permet de comparer par rapport aux résultats de l'étude ENTRED sur la France effectuée en 2001.

Il est à noter que les résultats viennent en complément du suivi déjà effectué par les médecins traitants des patients. Ces bilans sont effectués en tenant compte des bilans déjà effectués. Il ne saurait en aucun cas s'agir de la prise en charge globale comme c'est le cas de l'étude Entred.

année	3 dosage d'HBA1C	Dosage des lipides	Créatininémie	Micro albuminurie	Consultation ophtalmologie	ECG
1989	12.42	73.5	78.88	48.65	79.71	79.71
1990	42.71	74.96	78.39	40.65	78.73	78.73
1991	37.77	68.34	74.45	57.84	74.76	74.76
1992	33.48	68.09	75.04	54.04	75.89	75.89
1993	33.15	68.69	72.78	56.14	73.91	73.91
1994	32.5	66.5	71.25	56	71.75	71.75
1995	30.22	63.47	70.27	57.28	71.84	71.84
1996	28.72	63.96	69.03	52.59	69.59	69.59
1997	26.74	63.93	72.13	59.44	73.37	73.37
1998	30.71	61.32	69.08	56.11	79.49	79.49
1999	28.81	55.19	67.26	50.89	78.67	78.67
2000	25.89	61.26	67.2	57.09	76.77	76.77
2001	26.29	65.27	69.86	63.31	80.26	80.26
2002	23.99	66.07	69.31	64.62	76.33	76.33
2003	30.35	67.57	69.34	66.02	77.27	77.27
2004	33.89	63.59	66.11	60.43	73.74	73.74
Entred 2001	29.9	63.4	71.3	17.7	43	29.8

**Tableau 18 - Suivi des diabétiques et adéquation aux recommandations de la HAS**

La Figure 71 résume le Tableau 18 sous forme de radars. Le premier radar représente les résultats de l'étude ENTRED. Ensuite un trait gris servira de comparaison sur les autres radars exposant les résultats tirés de notre base année après année.

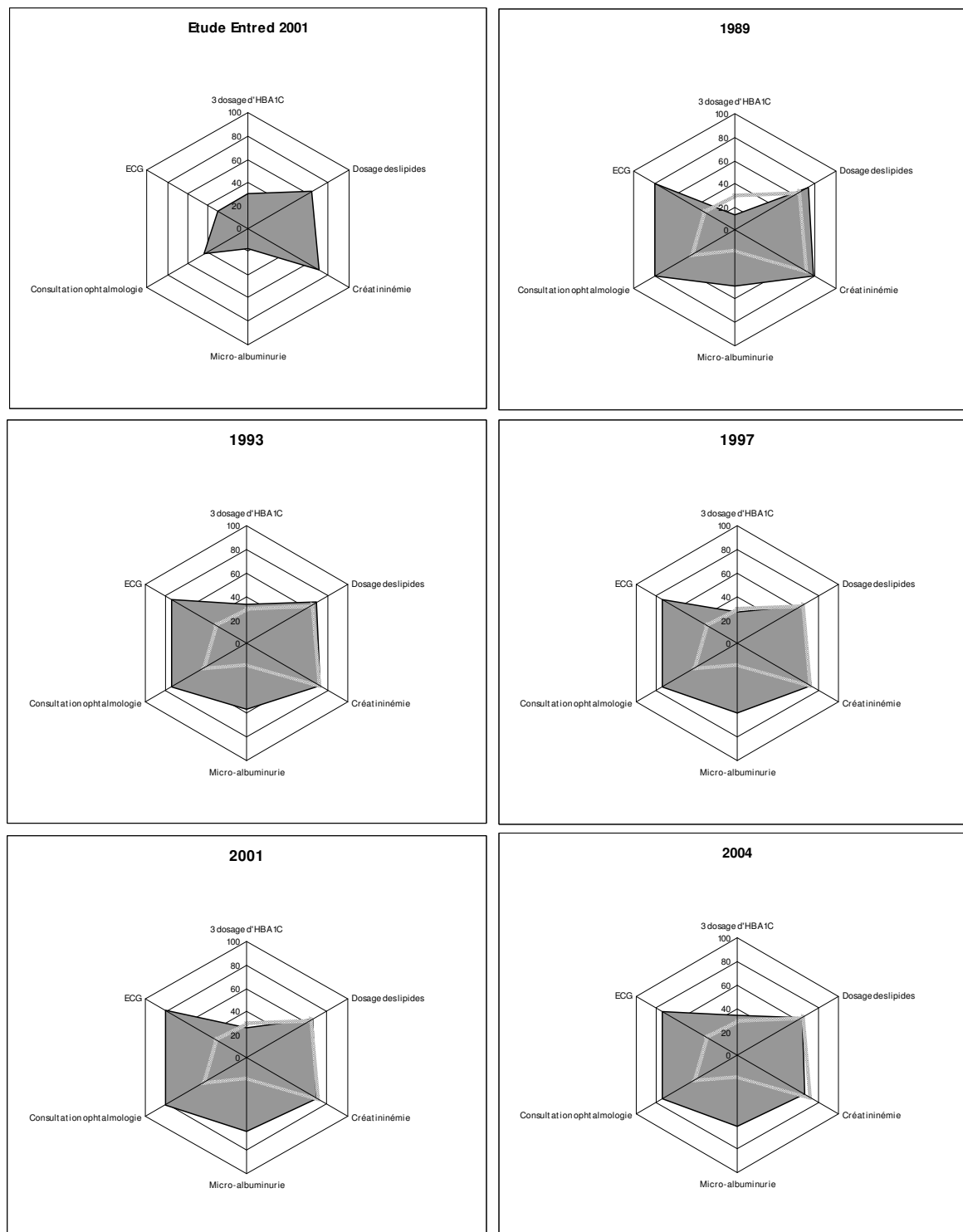


Figure 71 - Comparatif du suivi des recommandations HAS selon les années par rapport à l'étude ENTRED 2001

## 2. "Education" du médecin traitant

Nous définissons le numéro d'ordre ainsi : soit un médecin traitant M, qui adresse plusieurs patients au CETRADIMN. Ces patients sont classés par l'ordre chronologique de leur première venue au CETRADIMN, P1, P2, P3... Cet indice

constitue alors une variable que nous avons intitulée « numéro d'ordre », propre à chaque patient, supérieur ou égal à 1.

Nous allons tester ci-après si ce numéro d'ordre participe statistiquement à la formation de certaines variables biologiques.

La mesure porte sur les 2230 patients diabétiques de type 2 dont le médecin traitant est renseigné et dont la biologie est renseignée.

### **a- Numéro d'ordre et première hémoglobine glycosylée (HbA1c)**

Nous représentons ci-après la distribution des valeurs de la première HbA1c selon le numéro d'ordre du patient, à l'aide de box plots.

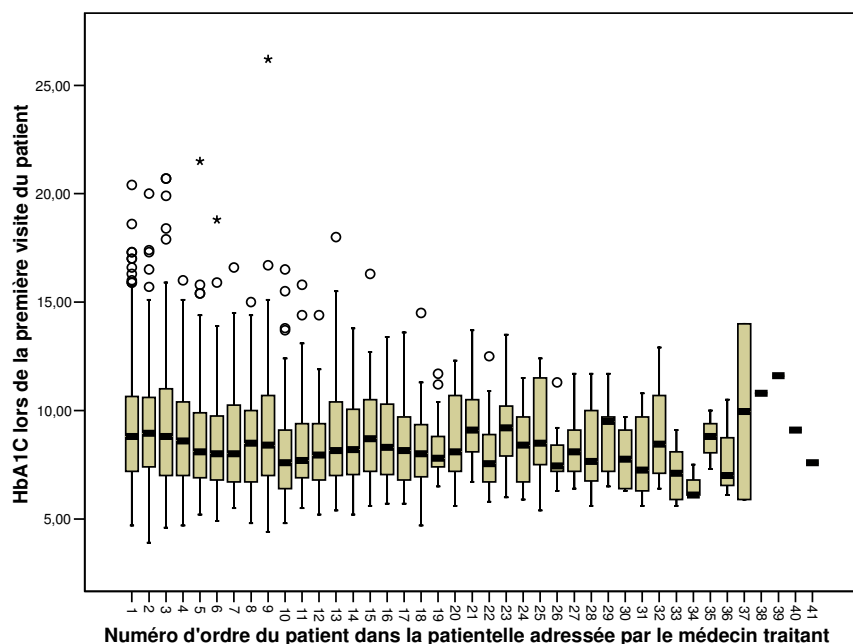


Figure 72 - Box plots - première HbA1c du patient en fonction de son numéro d'ordre

La représentation suivante traite des mêmes variables, mais en nuage de points. L'intérêt est double : rappeler l'aspect décroissant des effectifs en fonction du numéro d'ordre, et représenter un ajustement linéaire.

Ce graphique suggère que la première HbA1c du patient pourrait décroître lorsque son numéro d'ordre augmente.

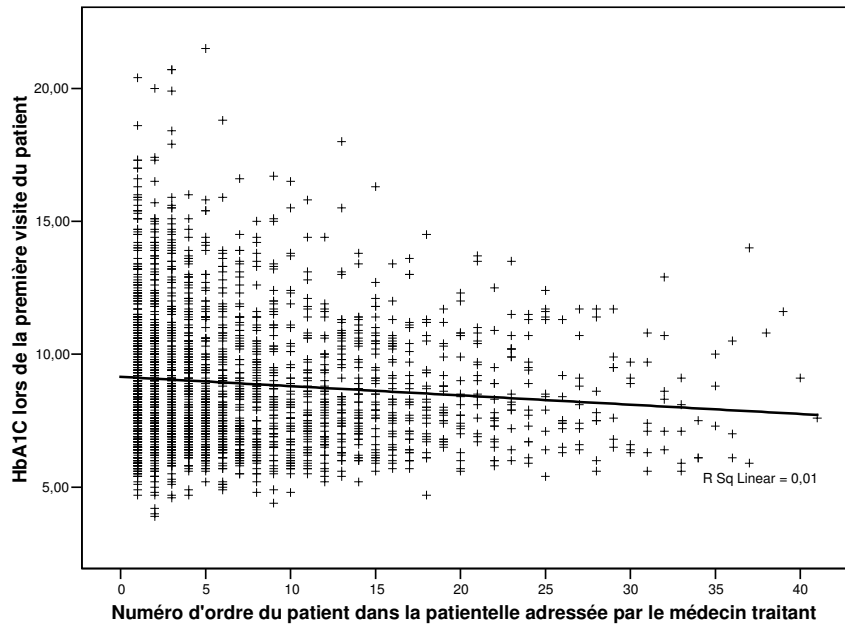


Figure 73 - Nuage de points - première HbA1c du patient en fonction de son numéro d'ordre

Nous réalisons ci-après une régression linéaire simple afin d'évaluer cet effet.

Coefficients(a)							
	Unstandardized Coefficients		Standardized Coefficients Beta	t	Sig.	95% Confidence Interval for B	
	B	Std. Error				Lower Bound	Upper Bound
(Constant)	9.148	.076		120.206	.000	8.999	9.297
numero_ordre	-.035	.007	-.099	-4.705	.000	-.049	-.020

a Dependent Variable: premiere\_hba1c

ANOVA(b)					
	Sum of Squares	df	Mean Square	F	Sig.
Regression	140.079	1	140.079	22.141	.000(a)
Residual	14096.030	2228	6.327		
Total	14236.110	2229			

a Predictors: (Constant), numero\_ordre  
b Dependent Variable: premiere\_hba1c

Tableau 19 - régression linéaire simple - première HbA1c du patient en fonction de son numéro d'ordre

Cette régression montre que le numéro d'ordre participe dans une relation linéaire à la formation de la première HbA1c du patient (ANOVA de la régression :  $p < 0.001$ ). La pente de cette relation est significativement négative (intervalle de confiance à 95% [-0.049 ; -0.020],  $p < 0.001$ ).

***b- Numéro d'ordre, année de recrutement et première hémoglobine glycosylée (HbA1c)***

Nous représentons ci-dessous la distribution des années de recrutement des patients en fonction de leur numéro d'ordre, à l'aide de Box Plots.

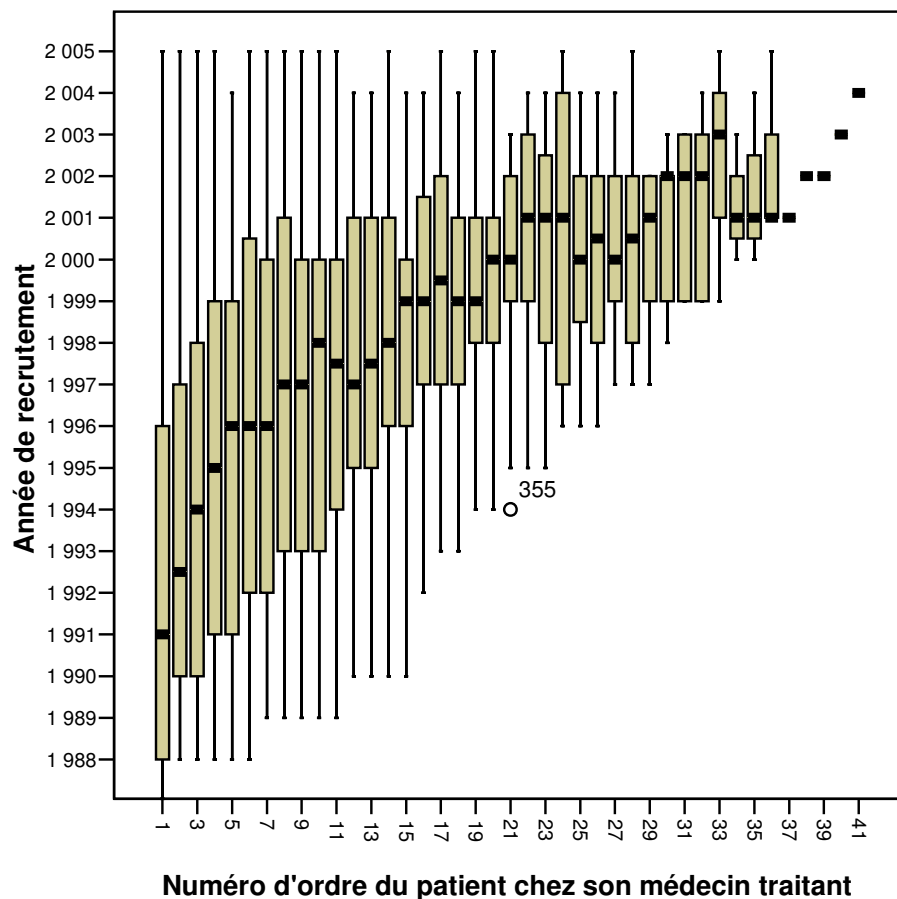


Figure 74 - Box Plots - année de recrutement du patient en fonction de son numéro d'ordre

Cette distribution suggère que l'année de recrutement pourrait être un facteur de confusion. Nous réalisons alors une régression linéaire multiple afin d'évaluer les effets concomitants et indépendants de l'année de recrutement du patient et de son numéro d'ordre sur sa première hémoglobine glycosylée mesurée au CETRADIMN.

**ANOVA(c)**

	Sum of Squares	df	Mean Square	F	Sig.
Regression	947.097	2	473.548	79.358	.000(b)
Residual	13289.013	2227	5.967		
Total	14236.110	2229			

b Predictors: (Constant), numero\_ordre, annee

c Dependent Variable: first\_hb\_glc

**Coefficients(a)**

	Unstandardized Coefficients		Standardized Coefficients Beta	t	Sig.	95% Confidence Interval for B	
	B	Std. Error				Lower Bound	Upper Bound
(Constant)	279.191	23.221		12.023	.000	233.654	324.728
numero_ordre	.007	.008	.020	.892	.372	-.009	.023
annee	-.135	.012	-.266	-11.629	.000	-.158	-.113

a Dependent Variable: first\_hb\_glc

Tableau 20 - régression linéaire multiple - première HbA1c du patient en fonction de l'année et du numéro d'ordre

La régression montre que le modèle linéaire multiple participe significativement à la constitution de la première HbA1c du patient (ANOVA de la régression,  $p < 0.001$ ). En revanche, si l'effet décroissant de l'année de recrutement est prouvé ( $p < 0.001$  et IC95 de la pente strictement inférieur à zéro), le numéro d'ordre n'intervient pas significativement ( $p = 0.372$  et IC95 de la pente encadrant zéro).

**c- Numéro d'ordre et première glycémie à jeun**

Nous représentons ci-après la distribution des valeurs de la première glycémie à jeun du patient selon son numéro d'ordre, à l'aide de box plots.

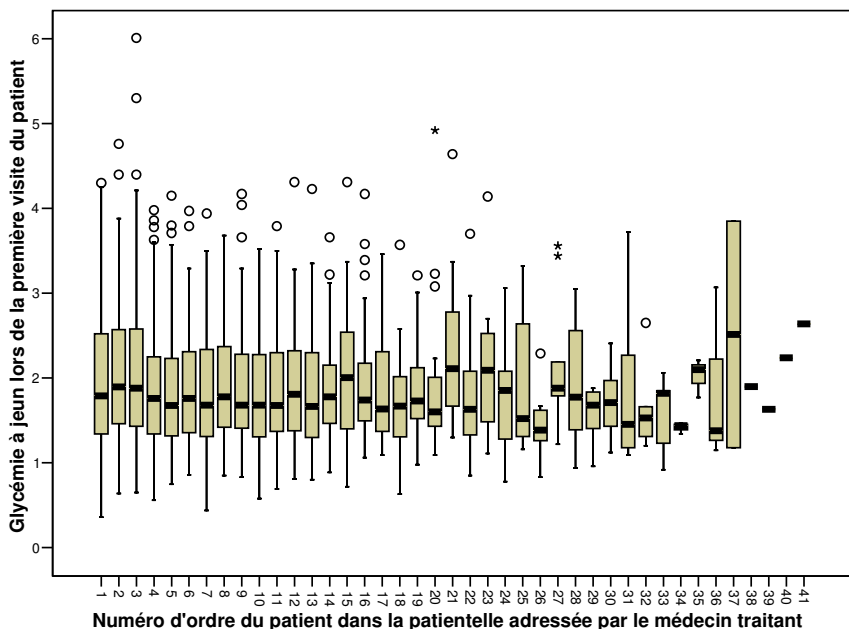


Figure 75 - Box Plots - première glycémie à jeun du patient en fonction de son numéro d'ordre

La représentation suivante traite des mêmes variables, mais en nuage de points. L'intérêt est double : rappeler l'aspect décroissant des effectifs en fonction du numéro d'ordre, et représenter un ajustement linéaire.

Ce graphique suggère que la première glycémie à jeun du patient pourrait décroître lorsque son numéro d'ordre augmente.



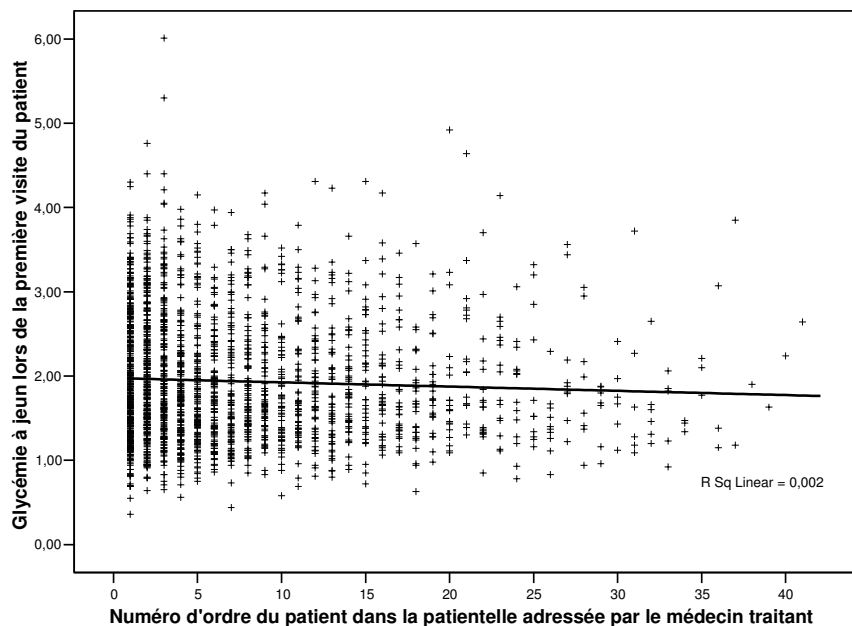


Figure 76 - nuage de points - première glycémie à jeun du patient en fonction de son numéro d'ordre

Nous réalisons ci-après une régression linéaire simple afin d'évaluer cet effet.

Coefficients(a)							
	Unstandardized Coefficients		Standardized Coefficients Beta	t	Sig.	95% Confidence Interval for B	
	B	Std. Error				Lower Bound	Upper Bound
(Constant)	1.975	.022		88.809	.000	1.931	2.018
numero_ordre	-.005	.002	-.049	-2.320	.020	-.009	-.001

a Dependent Variable: premiere\_glycemie\_a\_jeun

ANOVA(b)					
	Sum of Squares	df	Mean Square	F	Sig.
Regression	2.908	1	2.908	5.383	.020(a)
Residual	1203.530	2228	.540		
Total	1206.438	2229			

a Predictors: (Constant), numero\_ordre  
 b Dependent Variable: premiere\_glycemie\_a\_jeun

Tableau 21 - régression linéaire simple - première glycémie à jeun du patient en fonction de son numéro d'ordre

La régression montre un effet significativement décroissant du numéro d'ordre sur la première glycémie à jeun. Nous ne poursuivons toutefois pas nos investigations car :

- La glycémie à jeun ne témoigne pas avec fiabilité de l'équilibre glycémique
- Nous retrouvons, comme précédemment, que l'année de recrutement est un facteur de confusion.

### 3. Evolution des Diabétiques suivis

#### a- Nuage de points

Sur ces représentations, nous représentons pour tous les patients tous les duplets {valeur de l'HbA1c ; date de la mesure en jours}, en recalant tous les patients à zéro pour leur première visite au CETRADIMN.

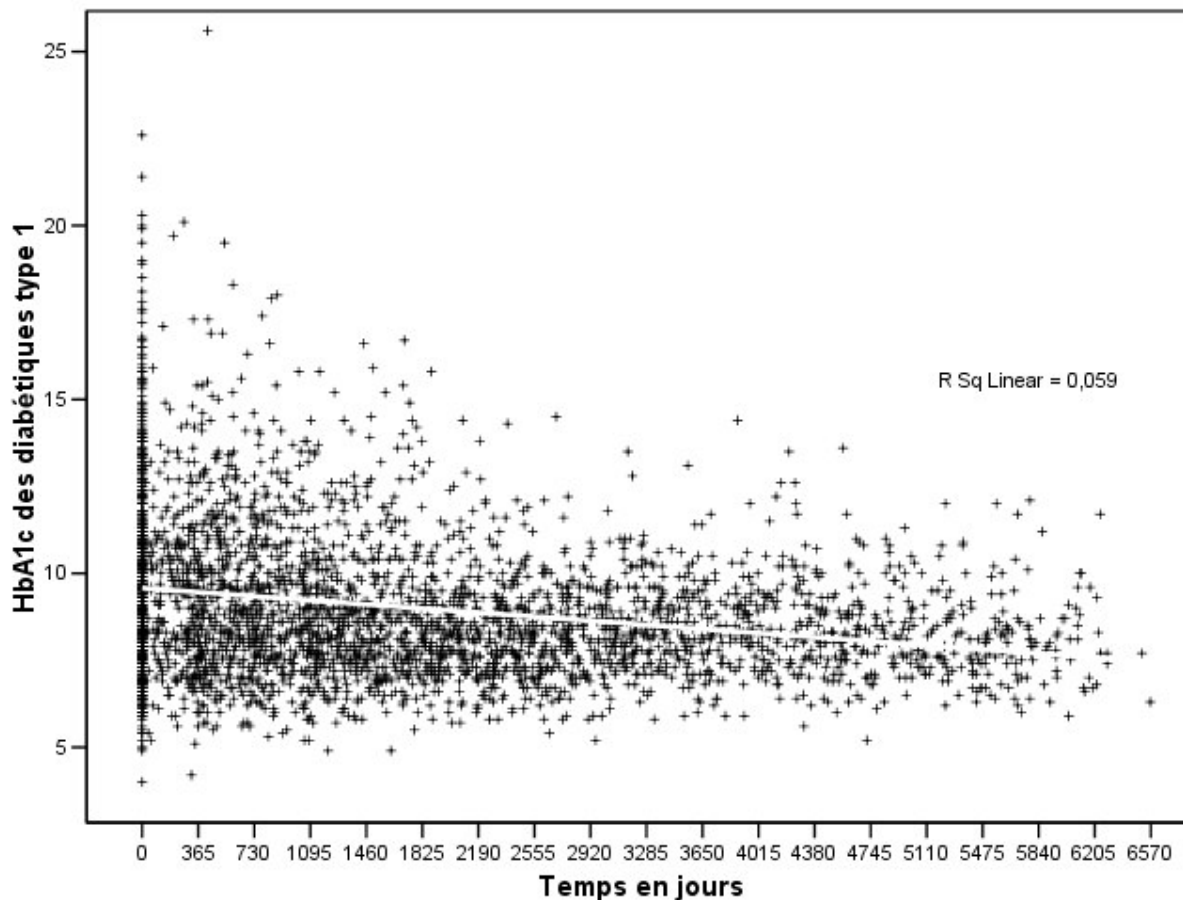


Figure 77 - nuage de points - HbA1c des diabétiques de type 1 au fil de la prise en charge

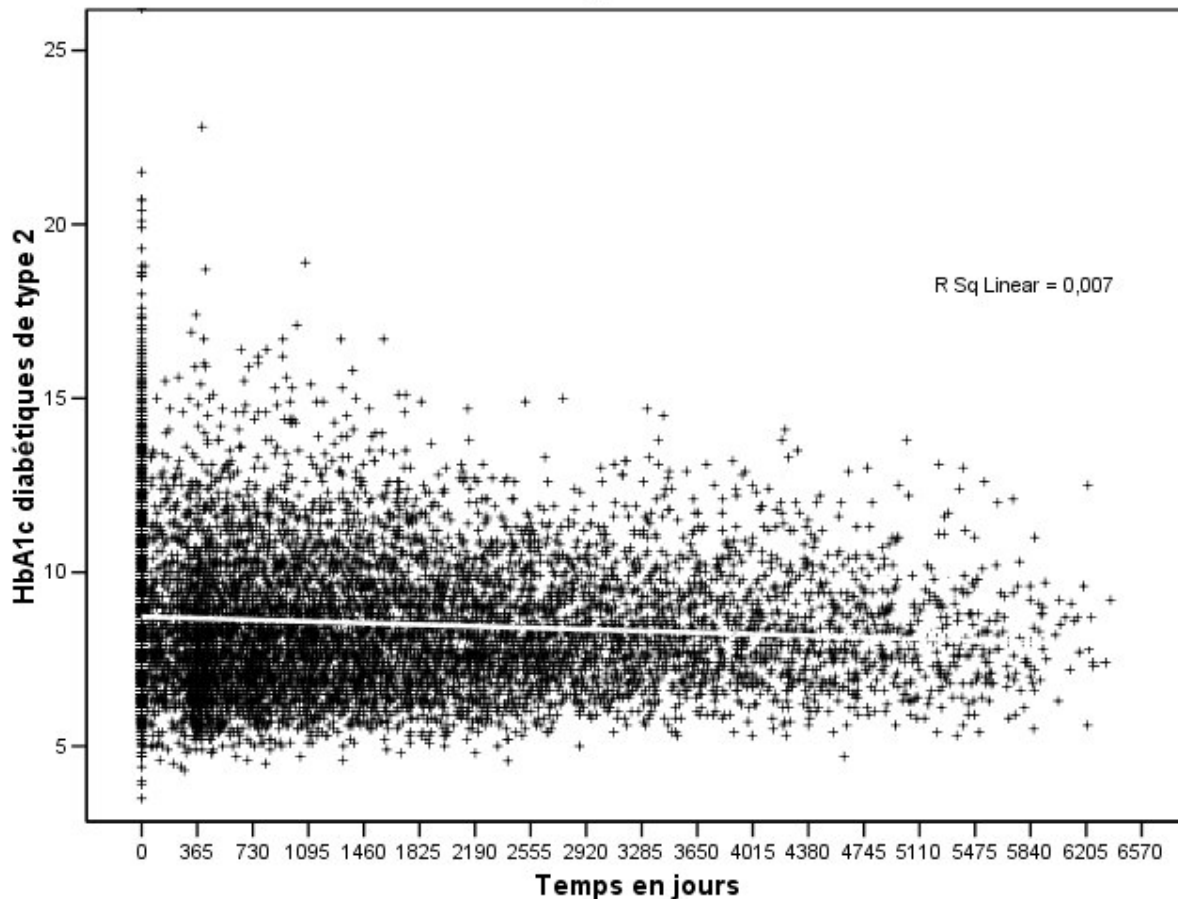


Figure 78 - nuage de points - HbA1c des diabétiques de type 2 au fil de la prise en charge

Ces nuages de points sont ici assortis d'un ajustement linéaire, représenté par une ligne blanche. Les ajustements cubiques et quadratiques sont dans les deux cas moins bons. L'ajustement linéaire pourrait suggérer une décroissance de l'HbA1c, cependant nous ne l'interpréterons pas en raison de nombreux biais.

***b- Tendence moyenne (semestrialisation)***

Une des difficultés de traitement provient du caractère aléatoire et asynchrone des dates de consultation. Afin de poursuivre notre exploration, ces dates de mesure sont arrondies au numéro du semestre par excès, et non au jour exact. Pour les patients diabétiques type 1 puis type 2, nous représentons successivement la distribution des valeurs en fonction du semestre de la mesure sous forme de Box Plots (Figure 79 – box plots - distribution de l'HbA1c des diabétiques type 1 au fil des semestres , Figure 81 – box plots - distribution de l'HbA1c des diabétiques type 2 au fil des semestres) puis les courbes représentant, pour chaque semestre, le nombre de patients, la moyenne des HbA1c, la moyenne diminuée de l'écart type, la moyenne augmentée de l'écart type (Figure 80 - courbe des moyennes - HbA1c des diabétiques type 1 au fil des semestres , Figure 82 - courbe des moyennes - HbA1c des diabétiques type 2 au fil des semestres).

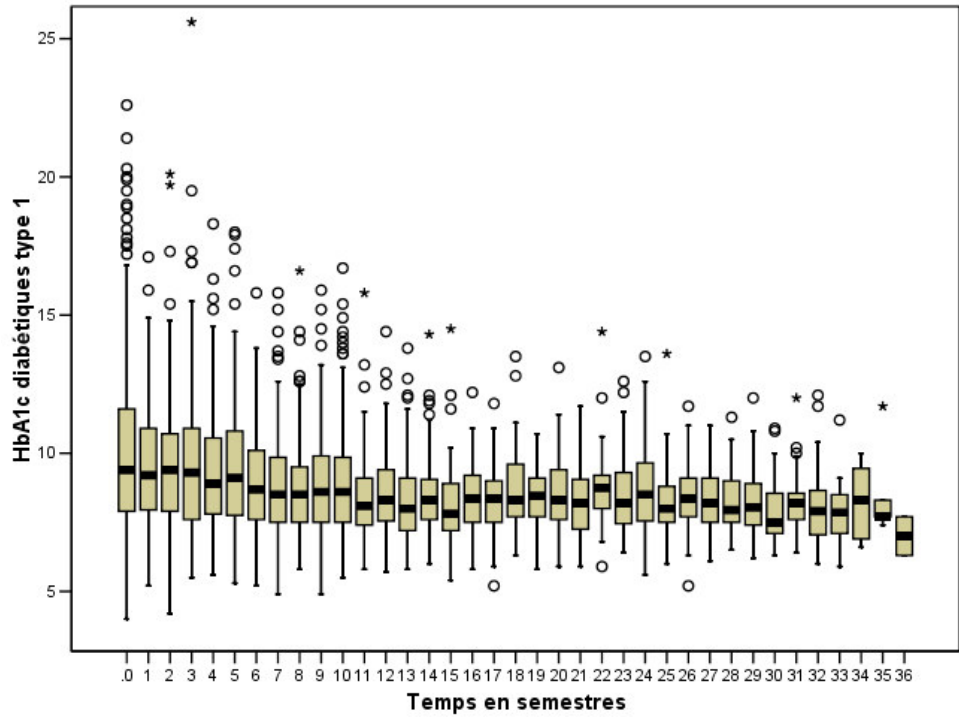


Figure 79 – box plots - distribution de l'HbA1c des diabétiques type 1 au fil des semestres

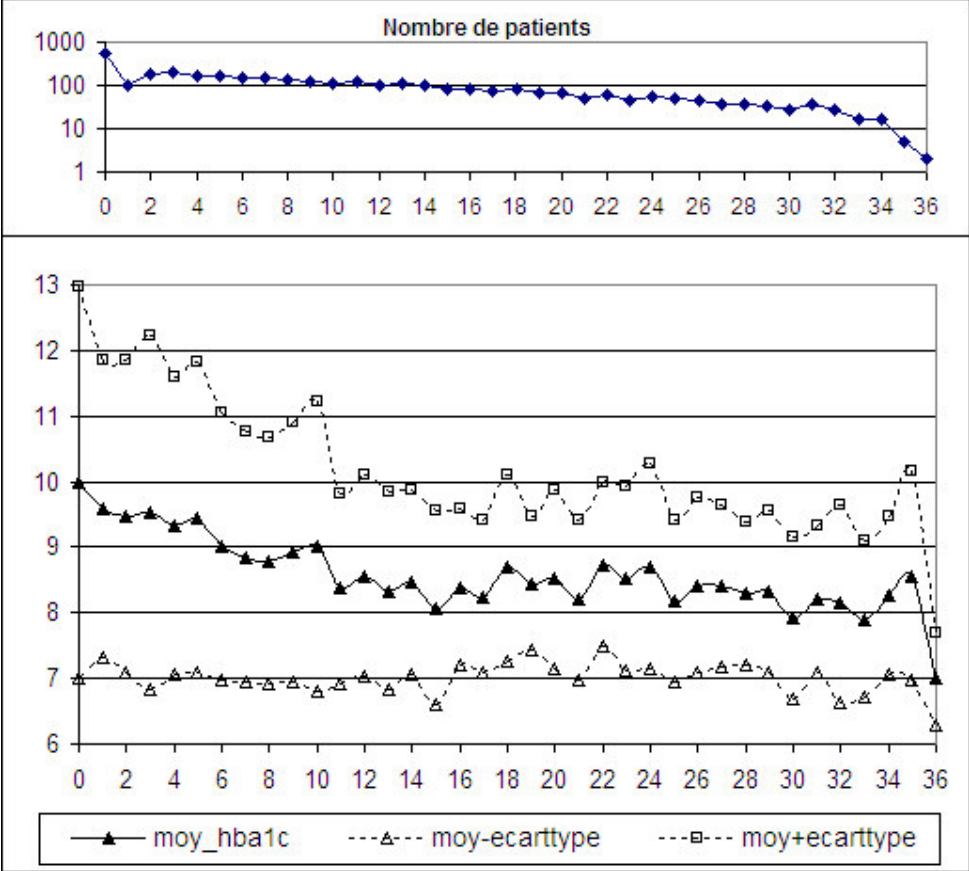


Figure 80 - courbe des moyennes - HbA1c des diabétiques type 1 au fil des semestres

Ces graphiques suggèrent une décroissance des HbA1c des diabétiques de type 1 durant les 5 premières années de prise en charge, puis une évolution variable.

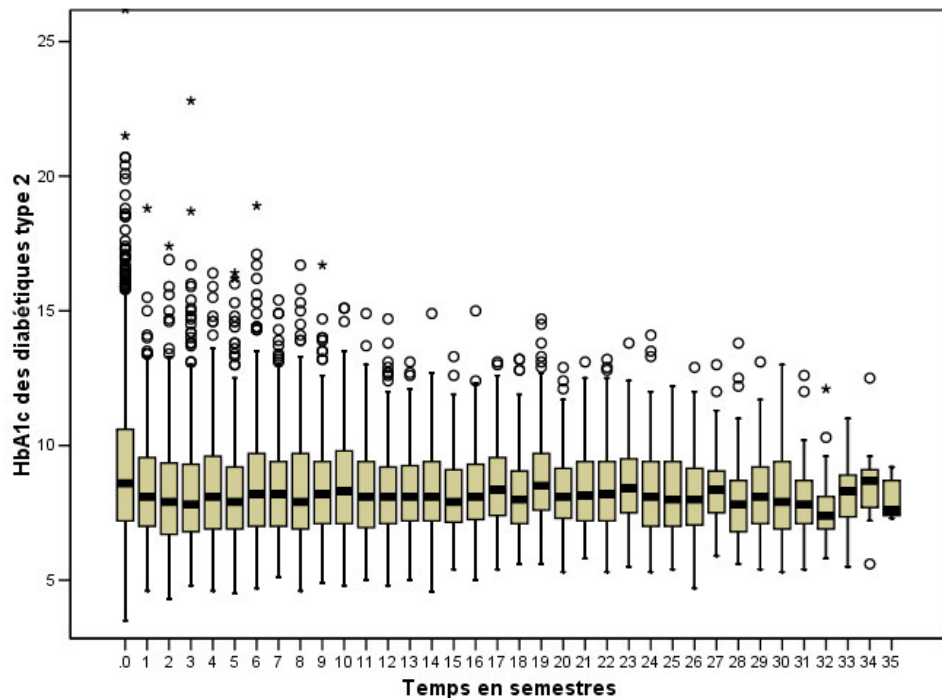


Figure 81 – box plots - distribution de l'HbA1c des diabétiques type 2 au fil des semestres

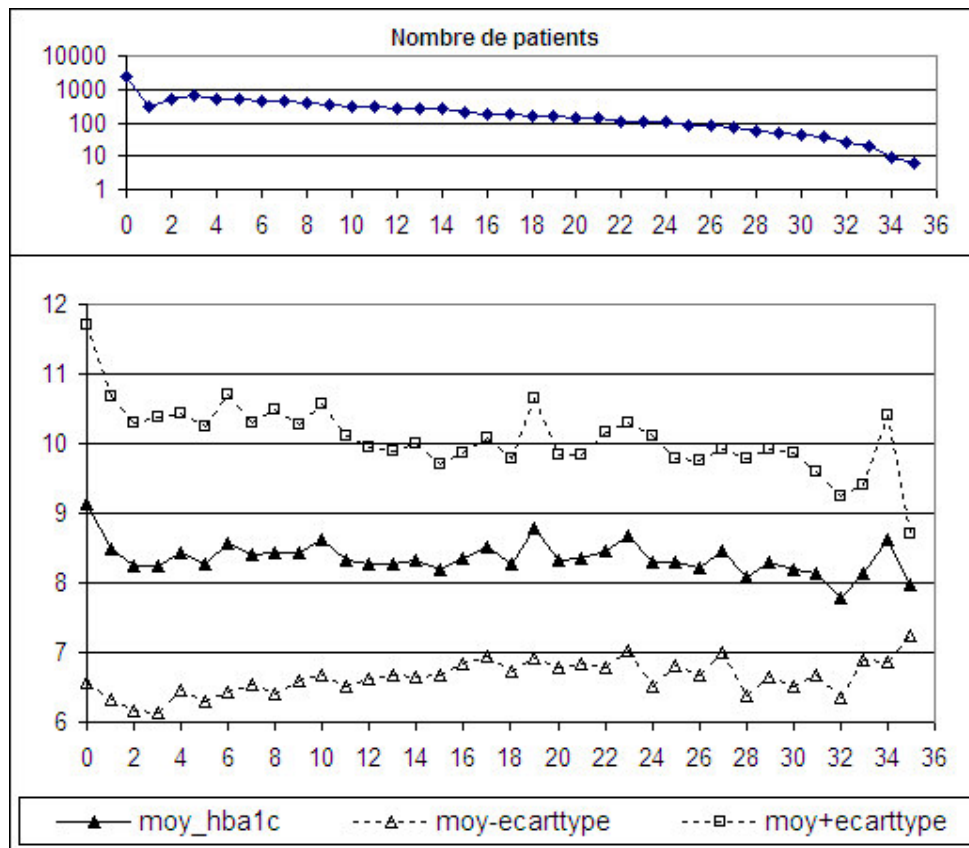


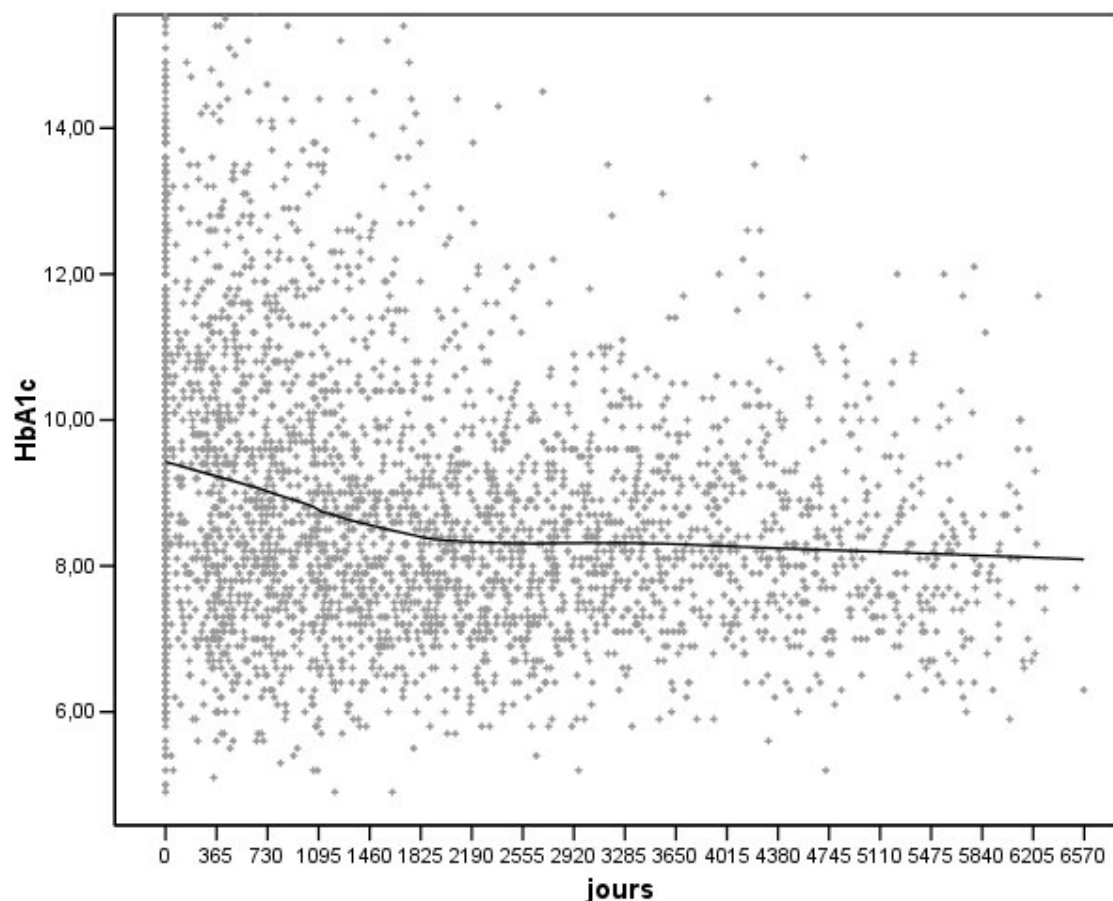
Figure 82 - courbe des moyennes - HbA1c des diabétiques type 2 au fil des semestres

Ces graphiques suggèrent une forte décroissance des HbA1c des diabétiques de type 2 durant la première année de prise en charge, puis une stabilisation avec rétrécissement de la variance.

### ***c- Tendance moyenne (lissage)***

Dans le même ordre d'idée que précédemment, il est possible de réaliser un lissage par moyenne mobile [44] par la méthode d'Epanechnikov. Cette méthode est développée dans la discussion.

Nous présentons dans un premier temps les HbA1c des diabétiques de type 1.



**Figure 83 - lissage Epanechnikov - HbA1c des diabétiques type 1 au fil du suivi**

Nous présentons ci-après un agrandissement de la courbure maximale de la courbe.

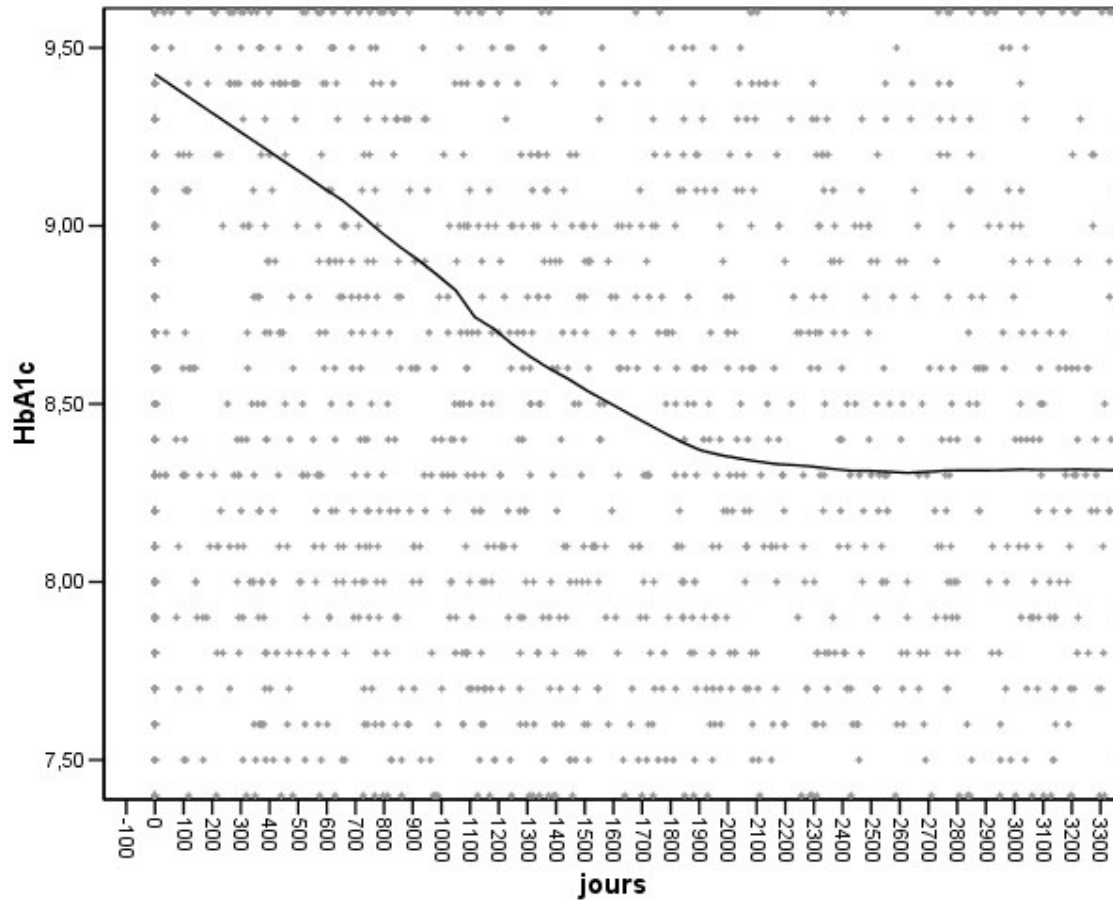


Figure 84 - lissage Epanechnikov - HbA1c des diabétiques type 1 au fil du suivi - ZOOM

Cette courbe lissée suggère une évolution linéaire biphasique des HbA1c des patients diabétiques de type 1 :

- une première phase de forte décroissance de l'HbA1c, de 0 à 2000 jours
- une deuxième phase de discrète décroissance de l'HbA1c, après 2000 jours

Nous présentons dans un deuxième temps les HbA1c des diabétiques de type 2.

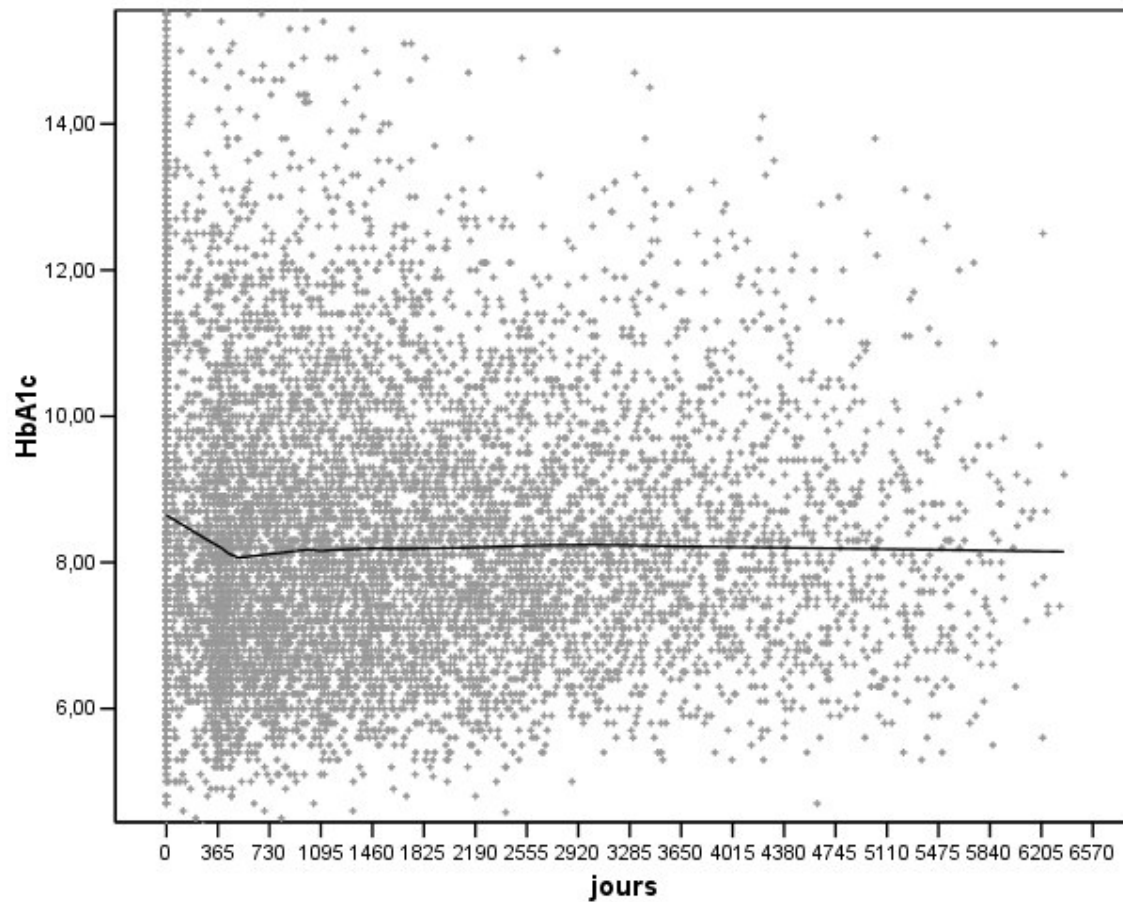


Figure 85 - lissage Epanechnikov - HbA1c des diabétiques type 2 au fil du suivi

Nous présentons ci-après un agrandissement de la courbure maximale de la courbe.



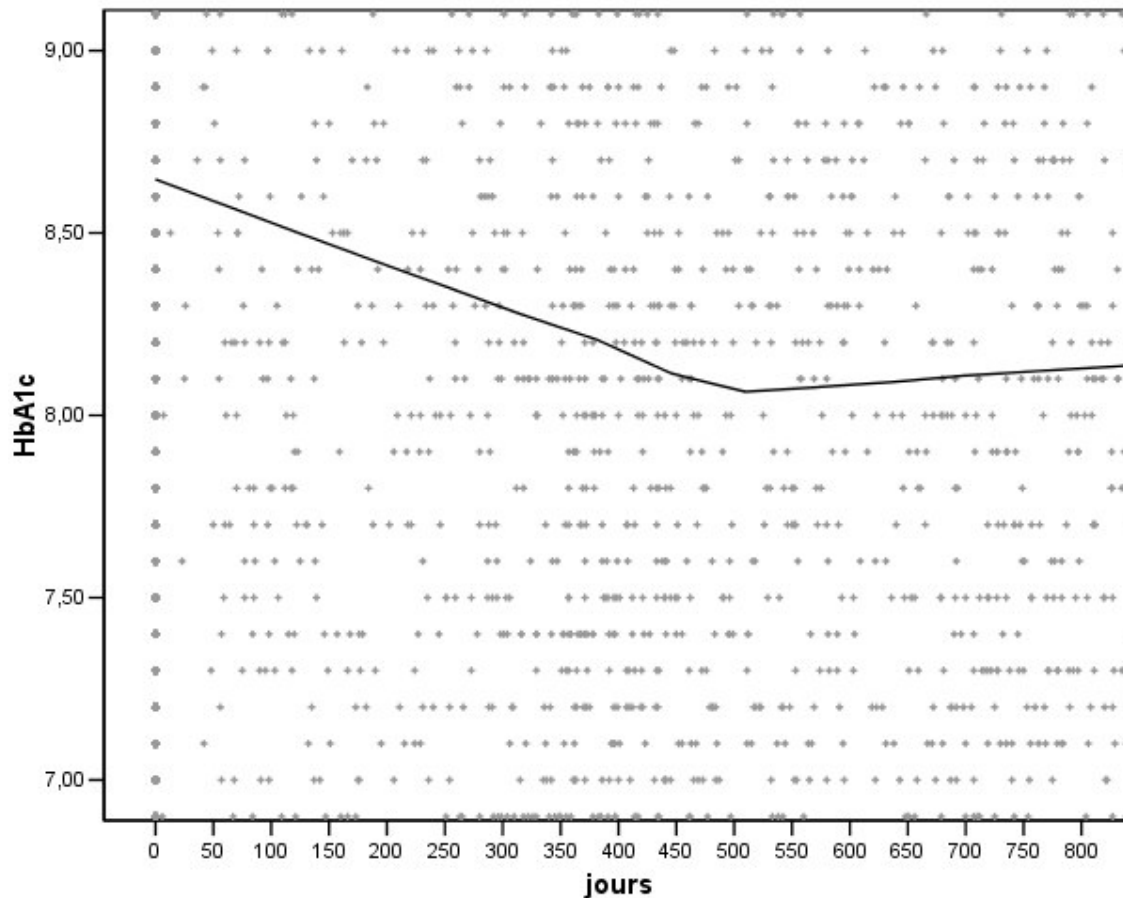


Figure 86 - lissage Epanechnikov - HbA1c des diabétiques type 2 au fil du suivi - ZOOM

Cette courbe lissée suggère une évolution linéaire biphasique des HbA1c des patients diabétiques de type 2 :

- une première phase de forte décroissance de l'HbA1c, de 0 à 500 jours
- une deuxième phase de discrète croissance de l'HbA1c, après 500 jours

#### ***d- Analyses harmoniques***

Le principe des analyses harmoniques est développé dans la Discussion. Nous réalisons ces analyses harmoniques à l'aide du logiciel « analyse harmonique » de Cristian Preda, maître de conférence à l'université Lille 2 [45].

Le but des analyses harmoniques est de décrire les profils d'évolution des HbA1c des patients en fonction du temps. L'analyse doit être menée sur des patients dont la durée de suivi est identique. Arbitrairement, nous sélectionnons tous les patients suivis au moins 8 ans, nous tronquons ce suivi à 10 ans.

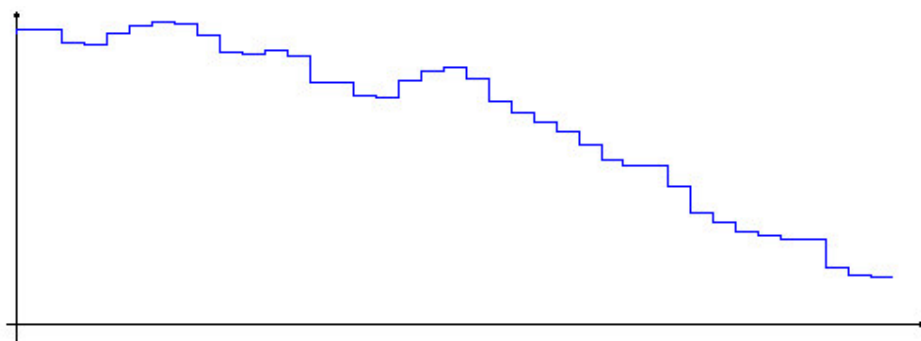
(Pour des raisons techniques, nous devons répéter à exactement 10 ans la dernière HbA1c mesurée avant échéance des 10 ans. Nous réalisons cela en programmant un script en PHP5. A défaut, le programme extrapole au-delà de la dernière mesure par la valeur zéro. Il en résulte que la plus grande source de variation devient le retour à zéro, et son asynchronisme provoque une entrée en

résonance des composantes principales, certes très esthétique, mais totalement ininterprétable.)

Nous demandons un découpage équidistant en 40 variables, qui correspondent donc à 1 trimestre.

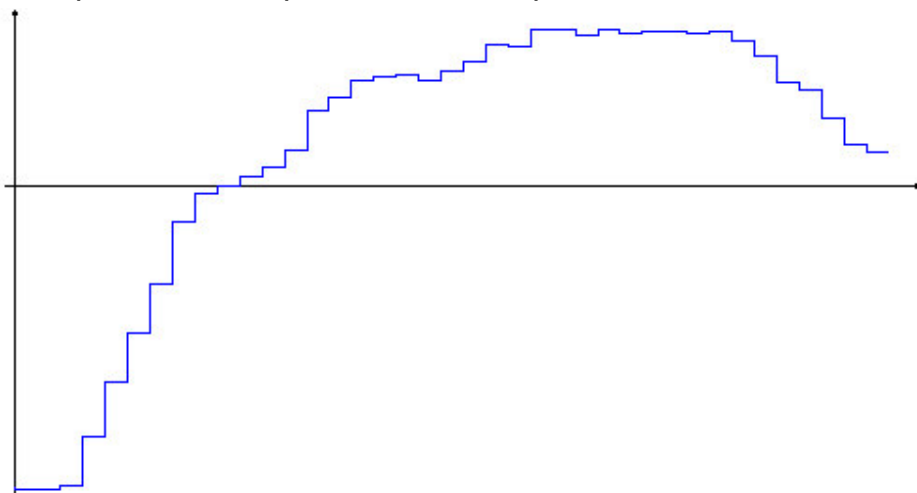
### **Les résultats chez les 191 diabétiques de type 1 étudiés sont les suivants :**

La première composante principale explique 61% de la variance totale. Elle est fortement positive chez les patients dont l'HbA1c diminue tout au long des 10 années.



**Figure 87 - première composante principale - diabétiques type 1**

La deuxième composante principale explique 14% de la variance totale. Elle est fortement négative chez les patients dont l'HbA1c diminue fortement pendant 5 ans, puis diminue plus lentement, puis remonte enfin au bout de 8 ans.



**Figure 88 - deuxième composante principale - diabétiques type 1**

Les deux premières composantes principales forment le premier plan factoriel, qui résume à lui seul 75.6% de la variance totale. Le diagramme suivant est peu utile à l'interprétation. Il montre la participation des variables trimestrielles dans la constitution des deux premières composantes principales.

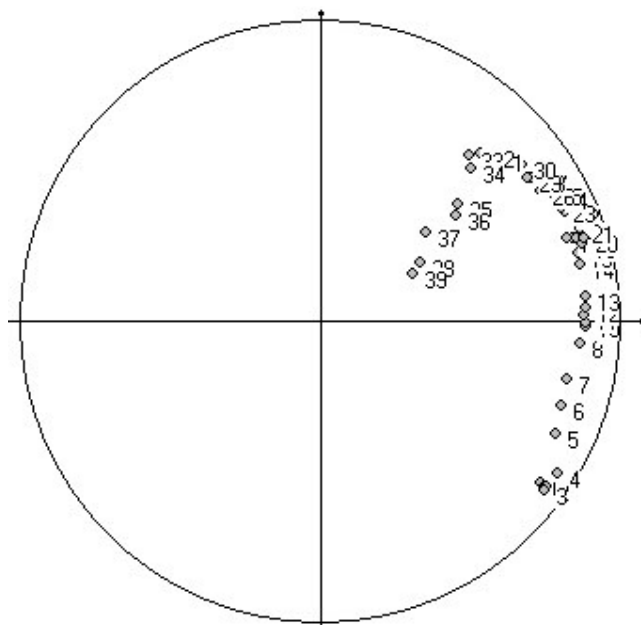


Figure 89 - contribution des variables trimestrielles au premier plan factoriel - diabétiques type 1

Enfin, le nuage de points suivants montre la répartition des patients sur le premier plan factoriel. Nous rappelons également comment interpréter ce plan.

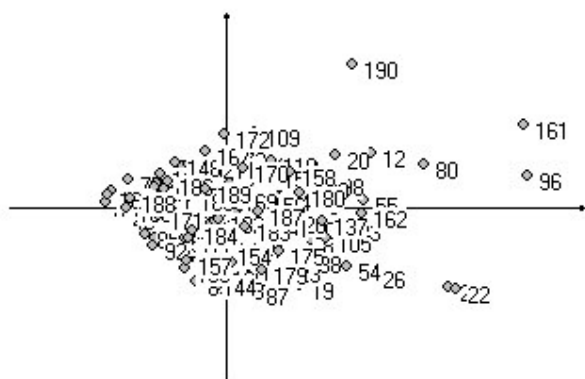


Figure 90 - positionnement des patients diabétiques type 1 sur le premier plan factoriel

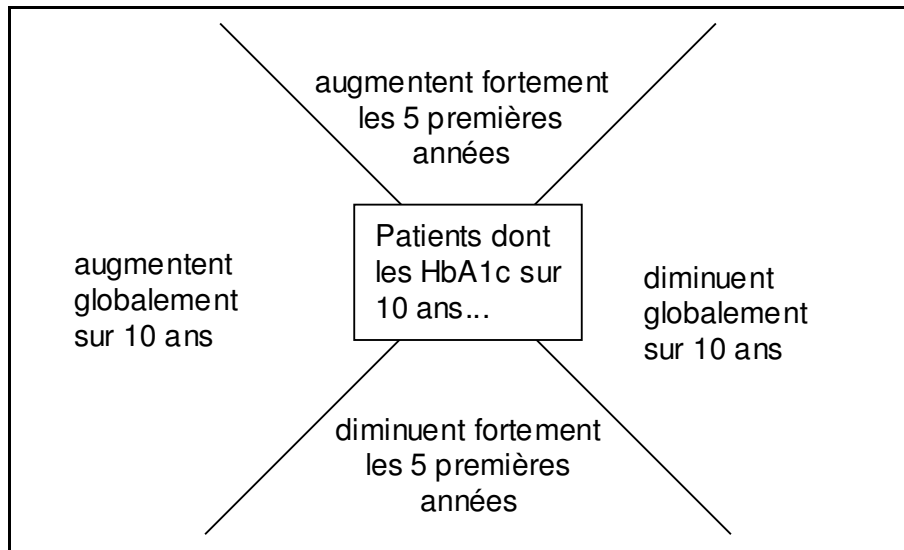


Figure 91 - interprétation du premier plan factoriel - diabétiques type 1

**Les résultats chez les 471 diabétiques de type 2 étudiés sont les suivants :**

La première composante principale explique 63% de la variance totale. Elle est fortement positive chez les patients dont l'HbA1c diminue de la 5<sup>e</sup> à la 10<sup>e</sup> année.

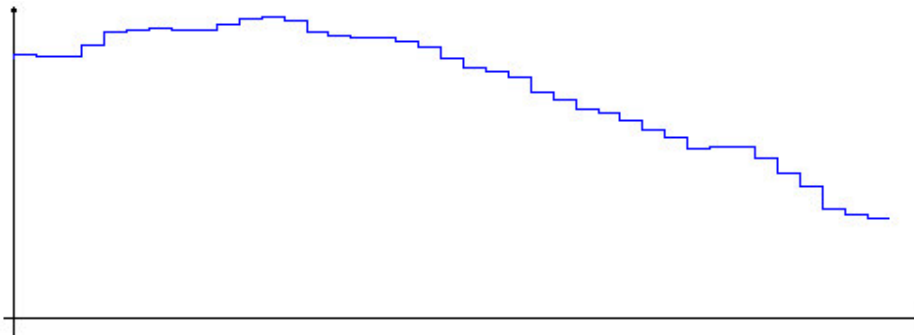
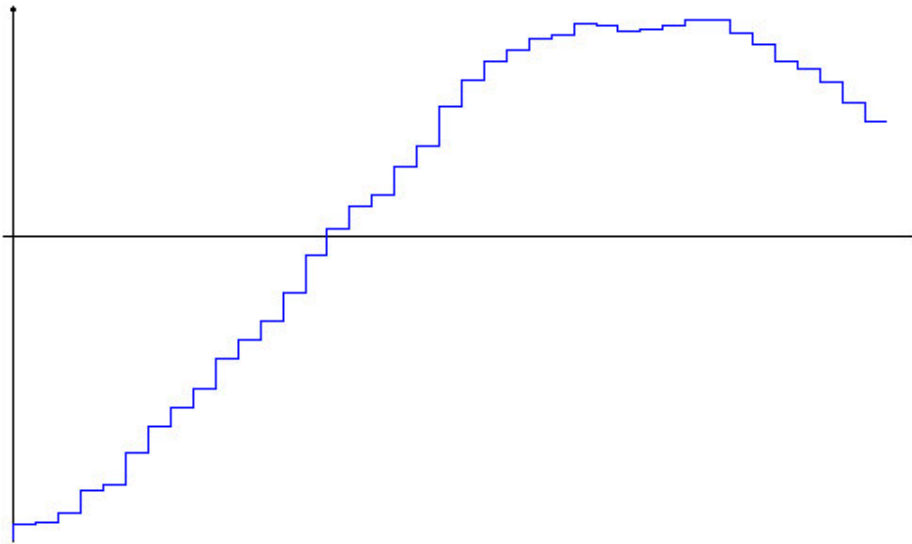


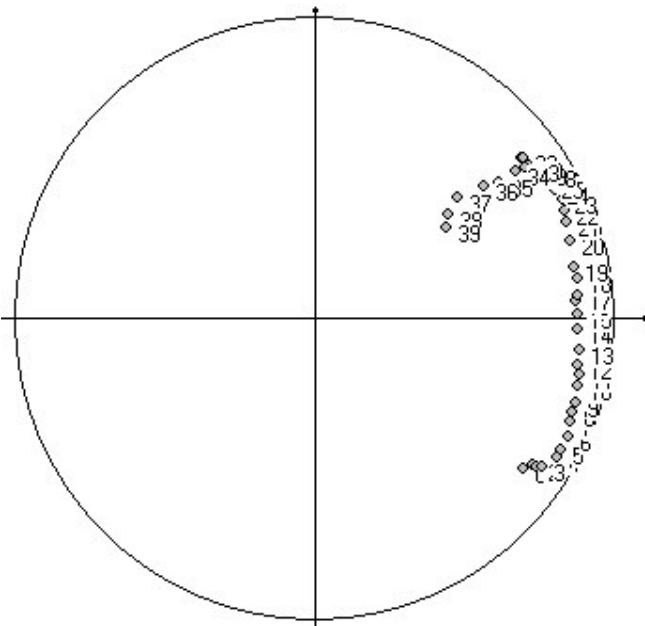
Figure 92 - première composante principale - diabétiques type 2

La deuxième composante principale explique 14% de la variance totale. Elle est fortement négative chez les patients dont l'HbA1c diminue fortement pendant 5 ans, puis se stabilise enfin au bout de 8 ans.



**Figure 93 - deuxième composante principale - diabétiques type 2**

Les deux premières composantes principales forment le premier plan factoriel, qui résume à lui seul 76.66% de la variance totale. Le diagramme suivant est peu utile à l'interprétation. Il montre la participation des variables trimestrielles dans la constitution des deux premières composantes principales.



**Figure 94 - contribution des variables trimestrielles au premier plan factoriel - diabétiques type 2**

Enfin, le nuage de points suivants montre la répartition des patients sur le premier plan factoriel. Nous rappelons également comment interpréter ce plan.



Figure 95 - positionnement des patients diabétiques type 2 sur le premier plan factoriel

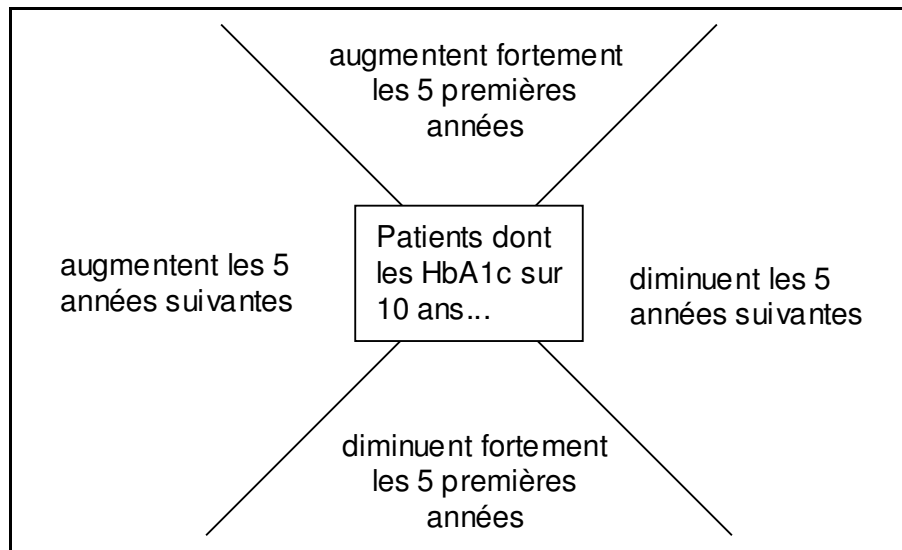


Figure 96 - interprétation du premier plan factoriel - diabétiques type 2

### e- Analyses de données longitudinales

Nous proposons et testons le modèle de régression linéaire multiple, ci-après. La conception et l'interprétation de ce modèle seront expliquées dans la discussion. Sont intégrés tous les patients diabétiques de type 1 ou 2 ayant au moins 1 mesure de l'HbA1c.

#### Modèle testé :

- Régression linéaire multiple :

$$Y = a1.t0 + a2.x2 + a3.x3 + a4.x4 + b + \epsilon$$

#### Constante :

- @seuil = 2000 jours chez les diabétiques type 1
- @seuil = 500 jours chez les diabétiques type 2

#### Variables :

- y = HbA1c mesurée au temps t

- $t_0$  = date première mesure de l'HbA1c de ce patient
- $t$  = date de la mesure de l'HbA1c
- $x_2 = SI( t-t_0 \leq @seuil , (t-t_0) , 0)$
- $x_3 = SI( t-t_0 \leq @seuil , 1 , 0)$
- $x_4 = SI( t-t_0 \geq @seuil , (t-t_0) , 0)$
- $x_5 = SI( t-t_0 \geq @seuil , 1 , 0)$
- $\epsilon$  = erreur d'ajustement au modèle linéaire

**Coefficients :**

- $a_1$  = pente générale liée à l'année d'inclusion
- $a_2$  = pente du premier segment
- $a_3$  = correction d'ordonnée à l'origine du premier segment
- $a_4$  = pente du deuxième segment
- $b$  = constante

**Tableau 22 - Evolution des HbA1c des patients diabétiques au fil du suivi : proposition de modèle**

**Résultat pour les 9817 diabétiques de type 1 :**

**ANOVA(b)**

Model		Sum of Squares	df	Mean Square	F	Sig.
1	Regression	2083.904	4	520.976	127.667	.000(a)
	Residual	14086.664	3452	4.081		
	Total	16170.568	3456			

a Predictors: (Constant), A4, A1, A2, A3

b Dependent Variable: hb\_glc

**Coefficients(a)**

Model		Unstandardized Coefficients		Standardized Coefficients		t	Sig.	95% Confidence Interval for B	
		B	Std. Error	Beta				Lower Bound	Upper Bound
1	(Constant)	9.204	.189			48.723	.000	8.834	9.575
	A1	-.136	.009	-.243		-14.704	.000	-.154	-.118
	A2	-.289	.026	-.219		-11.233	.000	-.339	-.238
	A3	1.290	.196	.289		6.590	.000	.906	1.673
	A4	-.052	.018	-.122		-2.859	.004	-.087	-.016

a Dependent Variable: hb\_glc

**Tableau 23 - Modélisation des HbA1c des diabétiques type 1 : résultat**

Sous l'hypothèse d'un modèle linéaire biphasique, l'ajustement linéaire est de bonne qualité (ANOVA de la régression  $p < 0.001$ ). L'année d'inclusion a un effet linéaire notable ( $p < 0.001$ , IC95=[-0.154 ; -0.118]) : les HbA1c sont de plus en plus basses avec le temps (les patients seraient recrutés plus tôt et mieux traités).

Les HbA1c diminuent de manière importante durant les 2000 premiers jours ( $p < 0.001$ , IC95=[-0.339 ; -0.238]) puis continuent de diminuer modérément les jours suivants ( $p < 0.001$ , IC95=[-0.087 ; -0.016]).

## Résultat pour les diabétiques de type 2 :

**ANOVA(b)**

Model		Sum of Squares	df	Mean Square	F	Sig.
1	Regression	3209.643	4	802.411	198.947	.000(a)
	Residual	39574.687	9812	4.033		
	Total	42784.330	9816			

a Predictors: (Constant), A4, A2, A1, A3  
b Dependent Variable: hb\_glc

**Coefficients(a)**

Model		Unstandardized Coefficients		Standardized Coefficients	t	Sig.	95% Confidence Interval for B	
		B	Std. Error	Beta			Lower Bound	Upper Bound
1	(Constant)	9.330	.062		149.595	.000	9.208	9.453
	A1	-.110	.005	-.244	-23.401	.000	-.119	-.101
	A2	-.820	.071	-.127	-11.532	.000	-.959	-.681
	A3	.570	.063	.133	8.981	.000	.446	.695
	A4	-.061	.007	-.117	-8.115	.000	-.075	-.046

a Dependent Variable: hb\_glc

**Tableau 24 - Modélisation des HbA1c des diabétiques type 2 : résultat**

Sous l'hypothèse d'un modèle linéaire biphasique, l'ajustement linéaire est de bonne qualité (ANOVA de la régression  $p < 0.001$ ). L'année d'inclusion a un effet linéaire notable ( $p < 0.001$ , IC95=[-0.119 ; -0.101]) : les HbA1c sont de plus en plus basses avec le temps (les patients seraient recrutés plus tôt et mieux traités).

Les HbA1c diminuent de manière très importante durant les 500 premiers jours ( $p < 0.001$ , IC95=[-0.959 ; -0.681], soit trois fois plus vite que chez les diabétiques type 1) puis continuent de diminuer plus modérément les jours suivants ( $p < 0.001$ , IC95=[-0.075 ; -0.046]).

## II. *Éléments administratifs*

### A. *Données d'activité*

Le Tableau 25 résume les données administratives de l'hôpital Victor Provo de Roubaix.

Les données ne portent pas toujours sur des années complètes :

- L'année 1990 ne comprend des données que jusqu'au mois de septembre compris (\*).

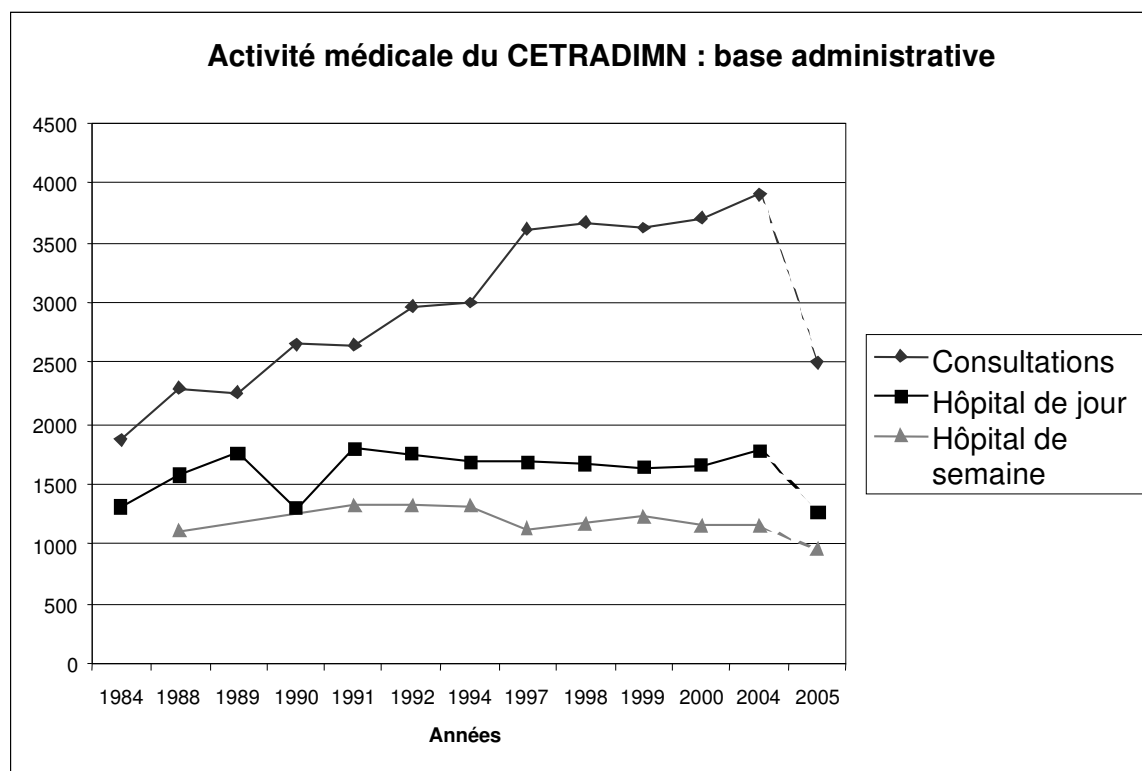


- L'année 2005 s'arrête au mois d'août compris (\*).

Les AMI sont comptabilisés sans notion d'unité jusqu'en 1994 puis, ensuite, différenciés en nombre et en valeur. Les valeurs étant proches de 1, on ne peut déduire l'unité utilisée avant 1994.

	1984	1988	1989	1990*	1991	1992	1994	1997	1998	1999	2000	2004	2005*
cs médicale	1874	2296	2265	2668	2641	2981	3001	3620	3670	3634	3700	3912	2501
cs infirmière	845	2106	2562	2434	3310	3428	3631	4614	4680	5326	5524	5247	3153
hôpital de jour	1302	1578	1757	1299	1789	1750	1682	1680	1661	1642	1654	1782	1254
hôpital de semaine		1102	-	-	1318	1322	1303	1124	1164	1227	1154	1155	959
cs diététicienne	1378	1716	1977	1575	2807	3257	3456	3488	3499	3381	3424	3249	2071
AMI	575	922	480	440	1325	1805	1074	1592	1645	1653	1558	1657	1012
cs psychologue			378	290	398		564	546	503	350	825	1164	461
cs podologie							168	440	470	511	436	607	494

**Tableau 25 - Récapitulatif des données administratives d'activité du CETRADIMN**



**Figure 97 - Evolution de l'activité médicale du CETRADIMN sur la base des données administratives**

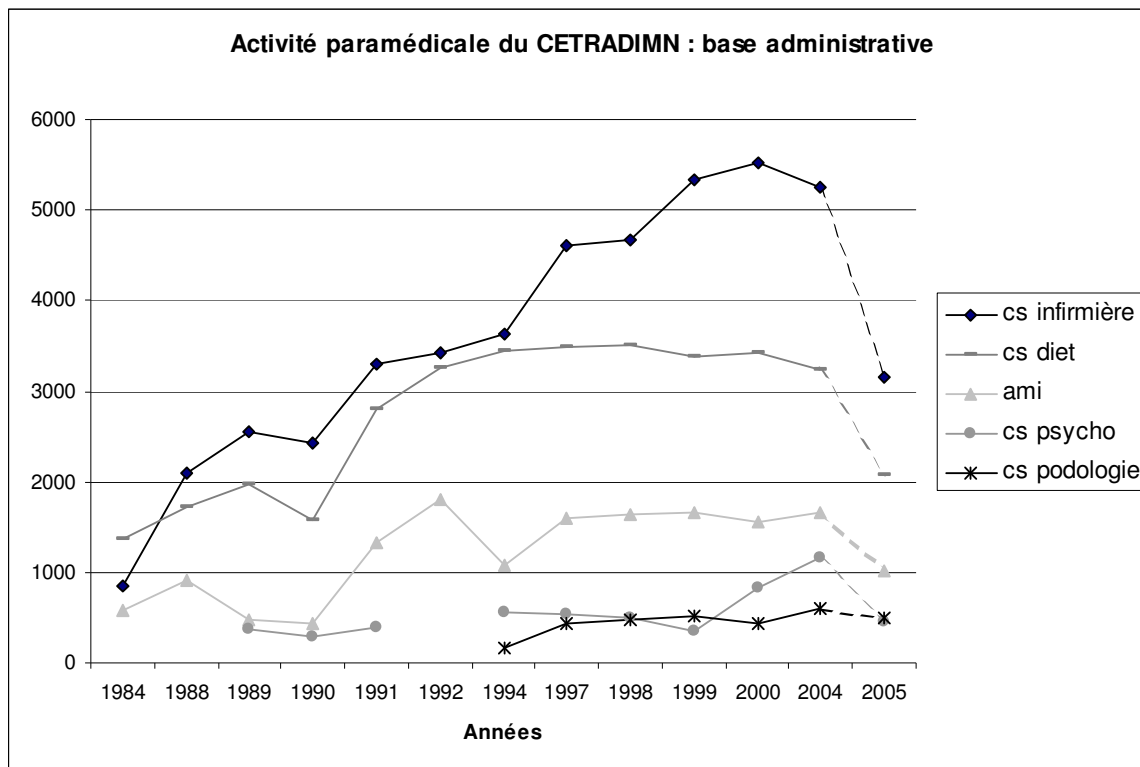
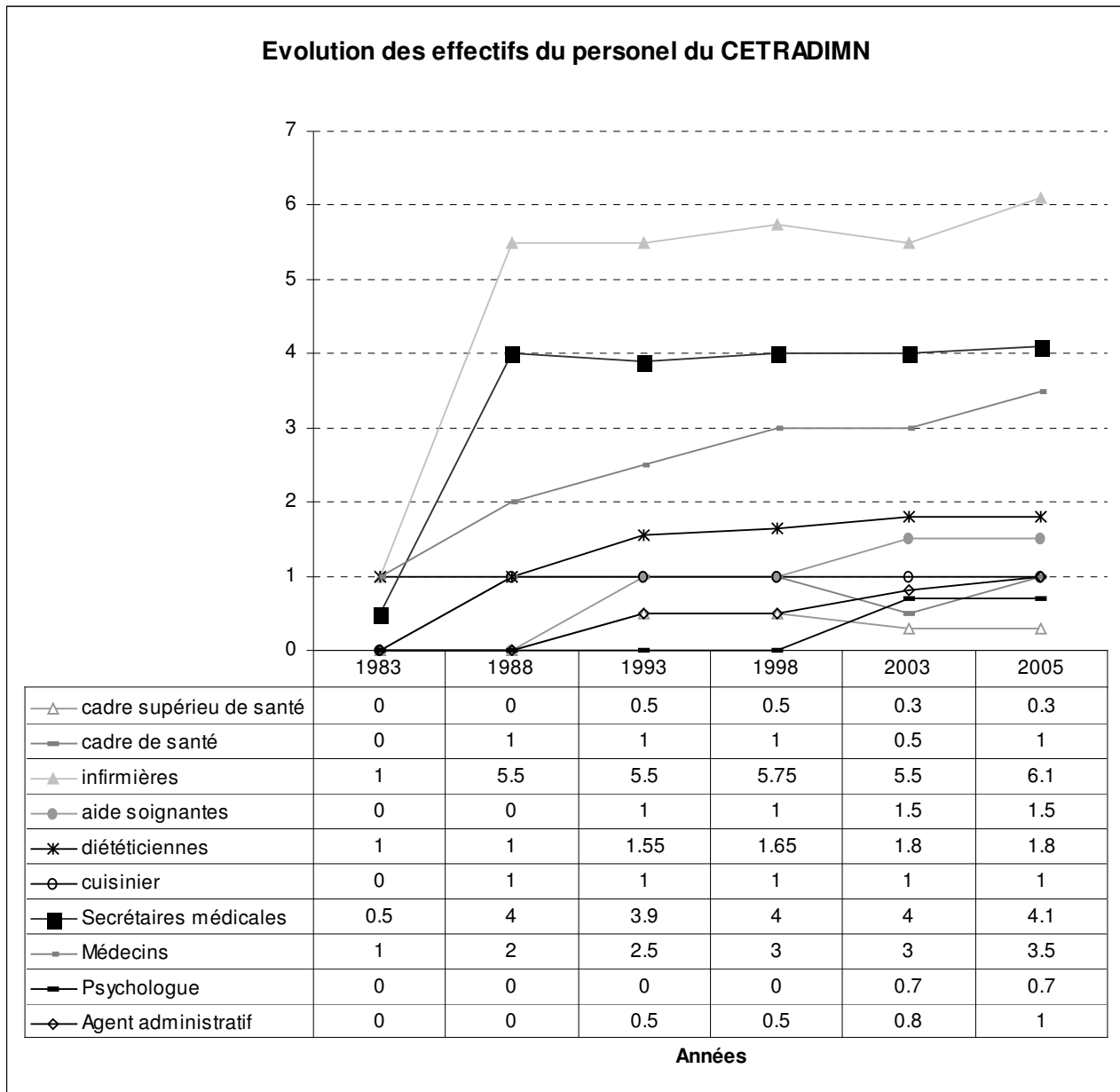


Figure 98 : Evolution de l'activité paramédicale du CETRADIMN sur la base des données administrative

### ***B. Données du personnel***

La Figure 99 représente l'évolution des effectifs du CETRADIMN, on remarquera une relative stabilité du nombre de secrétaires par rapport à un nombre de médecins croissant.



**Figure 99 - Evolution des effectifs médicaux et paramédicaux du CETRADIMN**

### ***C. Temps consacré à l'éducation du patient***

Le tableau ci-dessous (Tableau 26) résume les différentes modalités d'éducation du patient réalisées en groupes, avec le détail pratique quantifié pour chaque type de séquence.

EVALUATION 2004 DE L'EDUCATION THERAPEUTIQUE DU PATIENT AU CETRADIMN

<i>Titre de séquence</i>	<i>Mode de prise en charge</i>	<i>Nombre de séances/an</i>	<i>Nombre de patients par séance</i>	<i>Nombre patients total/an</i>	<i>Qualification des intervenants</i>	<i>Prévisions séances 2005</i>
Semaine classique	Hospitalisation de semaine	29	6 à 7	202	Ide/diet/médecins cuisinier/as	32 semaines
Semaine diététique	Hospitalisation de semaine	7	6 à 7	42	Diet/ide/cuisinier/as	9 semaines
Semaine adaptation de doses	Hospitalisation de semaine	1	7	7	Medecins/ide/as/ cuisinier	1 semaine
Module pied à risque	ambulatoire	2 de 2 jours	6 à 7	13	Medecin/podologue /psycho/ide/as	2 modules
Micro module diabète et risque cardio vasculaire	ambulatoire	2	3 à 4	7	Diet/ide/as	2 modules
Education thérap. patient DNIDabord diététique et culinaire	ambulatoire	2 fois par an de 12 séances	5 à 6	11	Diet/cuisinier	2 modules
Module d'éducation spécifique culture maghrébine	ambulatoire	Non réalisé en 2004				
TCC patients obèses à risque cardiovasculaire		2 fois par an chaque fois comprenant 16 séances	8	16	Psychologue/diet	2 fois 16 séances

Tableau 26 - les séquences proposées en groupe de patients au CETRADIMN

Le cadre supérieur de santé a mesuré en 1981 le temps consacré à l'éducation thérapeutique du patient, sur la base de 6 semaines d'hospitalisation de jour (5 pour diabétiques, 1 pour surpoids). Il s'agit d'une mesure chronométrée au fil de la journée. Le tableau réalisé en mai 2005 est une application directe des mêmes pourcentages, en tenant compte de l'évolution des effectifs (Tableau 27). Nous discuterons de l'exploitation qui peut être faite de cette observation dans la partie discussion.

<i>PERSONNEL</i>	<i>NOMBRE D'HEURES TRAVAILLEES</i>	<i>NOMBRE D'HEURES D'EDUCATION</i>	<i>% D'EDUCATION</i>
Médecins : 3.5 ETP/35H/SEM	735 H	224 H	30.50 %
Diététiciennes : 1.8 ETP/35H/SEM	378 H	291H	76.96 %
IDE de jour : 5.1 ETP/35H/SEM	1071 H	359H15	33.55 %
IDE de nuit : 1ETP/32H30/SEM	195 H	11H45	0.06 %
Aide soignante: 1.5 ETP/35H/SEM	315 H	87H45	27.87 %
Podologue : vacances 12H/SEM	72 H	72 H	100 %
Psychologue : 0.70 ETP/35H/SEM	147 H	102H45	70 %
Cuisinier : 1 ETP/35H/SEM	210 H	60H30	28.89 %
1 FFCS : 1 ETP/35H/SEM	210H	9H	4.36 %
Temps global	3 333 H	1 218 H	36.55%

Tableau 27 - Temps consacré à l'éducation du patient en mai 2005, sur la base de 6 semaines d'HDJ

## D. Données financières

Le tableau ci-dessous résume le budget 2003 du CETRADIMN (Tableau 28). Afin de lever toute ambiguïté, nous rappelons que le CETRADIMN est le service de diabétologie du CHG de Roubaix.


	BUDGET EDUCATION THERAPEUTIQUE ET PREVENTION (MIGAC)	BUDGET T2A/MCO diabétologie CT COMPLET	TOTAL BUDGETS
service de diabétologie budget global - ref ex 2003			
<b>DEPENSES DIRECTES</b>			
DEPENSES DE PERSONNEL	416 353,00	639 351,00	1 055 704,00
DEPENSES MEDICALES	15 116,00	22 867,00	37 983,00
DEPENSES HOTELIERES	625,00	17 744,00	18 369,00
DEPENSES DE STRUCTURE	2 124,00	5 316,00	7 440,00
RECETTES ATTENUATIVES		-10 016,00	-10 016,00
<b>TOTAL DEPENSES DIRECTES NETTES</b>	<b>434 218,00</b>	<b>675 262,00</b>	<b>1 109 480,00</b>
<b>DEPENSES INDIRECTES</b>			
LOGISTIQUE MEDICALE	2 514,00	4 317,00	6 831,00
LOGISTIQUE GENERALE ET STRUCTURE	136 857,00	287 399,00	424 256,00
IMAGERIE		46 745,00	46 745,00
LABORATOIRE BIOLOGIE		208 209,00	208 209,00
LABORATOIRE ANAPATH		1 186,00	1 186,00
RESTAURATION		38 527,00	38 527,00
BLANCHISSERIE		15 717,00	15 717,00
<b>TOTAL DEPENSES INDIRECTES NETTES</b>	<b>139 371,00</b>	<b>602 100,00</b>	<b>741 471,00</b>
<b>TOTAL GENERAL</b>	<b>573 589,00</b>	<b>1 277 362,00</b>	<b>1 850 951,00</b>

Tableau 28 - données financières

## DISCUSSION

### *I. Analyse des résultats*

#### **A. Activité du CETRADIMN**

En comparant les résultats d'activité de notre base de donnée (I.A.2.a- p48) à ceux fournis par l'administration du CETRADIMN (II.A p120) on remarque une discordance. Lors d'une consultation effectuée durant un hôpital de jour ou de semaine, celle-ci sera comptabilisée comme une consultation ET un hôpital de jour ou de semaine dans le cadre administratif, alors que dans la base, seul un événement sera comptabilisé. Ceci est une première explication qui n'exclut pas un défaut de saisie. Malgré cette constatation, on peut mettre en rapport la Figure 10 et la Figure 97 qui semblent suivre une évolution similaire.

Sur la Figure 10 les hospitalisations de semaine, apparues en 1998, sont restées constantes. Cela peut s'expliquer par un nombre de lits constant entraînant un phénomène de saturation.

En mettant en rapport les Figure 11, Figure 12 et Figure 13, on peut s'apercevoir de l'effet de saturation de la prise en charge des nouveaux patients. En effet, même si l'on remarque un accroissement constant du nombre de patients distincts pris en charge, on note une part croissante des anciens patients, alors que le nombre de nouveaux patients semble lui stagner voire diminuer.

Si l'on met en rapport ces constatations avec la Figure 99, l'arrivée en 1988 d'un deuxième médecin a permis une nette augmentation de la patientèle, tant en terme de capacité de recrutement, qu'en terme de suivi mais rapidement, on note un effet de saturation, avec une poursuite du suivi mais une diminution nette du recrutement, qui sera stabilisé avec l'arrivée d'un mi-temps supplémentaire en 1993. Puis un temps complet en 1998 permettra à nouveau une augmentation légère de la capacité de prise en charge de nouveaux patients, qui saturera à nouveau 4 ans plus tard, pour même chuter très nettement.

Enfin sur la Figure 15, nous remarquons que le taux de fuite fluctue jusqu'en 1997 avec une remontée progressive à compter de cette année-là.

Il semble que l'arrivée d'un nouveau médecin permette d'accroître la prise en charge de nouveaux patients (phase de création de patientèle) dans un premier temps, puis qu'il existe un effet de saturation environ trois à quatre années plus

tard (phase de suivi de patientèle) avec un renouvellement correspondant aux patients perdus.

L'ensemble de ces constatations semble appuyer la thèse que le CETRADIMN est actuellement dans une phase de saturation, tant du point de vue du suivi que de l'augmentation de sa patientèle. Il semble nécessaire d'augmenter sa capacité de plage d'accueil, par la venue de nouveaux médecins, pour poursuivre sa croissance.

### ***B. Médecins traitants.***

Dans cette partie, nous avons voulu concevoir un indice de familiarité qui résumerait le degré de collaboration entre les médecins traitants correspondants et le CETRADIMN. Il est à noter qu'un patient peut très bien se présenter de lui-même au CETRADIMN. De même, un médecin traitant peut adresser un patient au CETRADIMN, sans que celui-ci ne décide de venir. Toutefois nous pensons que l'indice de familiarité a du sens et permet de qualifier un réseau sur une période définie (car bien évidemment il varie selon le temps de l'étude).

Nous n'avons, malheureusement, pas pu trouver dans la bibliographie de quoi nous comparer. Mais le résultat de la Figure 19 montre que le CETRADIMN est un réseau étendu mais peu familier, en effet la majorité des médecins traitants envoient moins de 4 patients. On note par le zoom sur la file active de la Figure 20, que le CETRADIMN a tendance à améliorer sa distribution de médecin plus familier.

### ***C. Présentation des patients : de la totalité de la base, de la file active et des diabétiques de la file active.***

#### ***1. Répartition hommes/femmes***

On peut noter que le rapport Hommes/Femmes pris en charge est significativement en faveur des femmes :

- Sur la totalité de la base (Figure 21) qui a tendance à globaliser les tendances variables dans le temps puisqu'elle décrit une pratique sur 17 années, on vérifie cette affirmation.
- Sur la file active (Figure 25) qui a tendance à montrer les pratiques récentes puisqu'elle décrit les pratiques de 2001 à 2005, on note un léger rééquilibrage de ce rapport.

Il est intéressant de voir sur les patients diabétiques de la file active, que ce rapport est fonction du type de diabète :

- Il est bien évident que le diabète gestationnel est forcément féminin.

- Par contre on remarque que ce rapport hommes/femmes s'équilibre dans le cas du diabète de type 1.

On peut supposer que les hommes seraient moins compliants au suivi, dans le cadre d'une maladie chronique et asymptomatique tel le diabète type 2, et que l'amélioration de leur compliance, dans le diabète de type 1, serait due à l'amélioration de la qualité de vie qu'induit le suivi, le diabète type 1 n'étant pas durablement asymptomatique.

Une étude [46] menée en 1999 sur 975 patients a montré, chez le diabétique de type 2, une meilleure compliance à la prise de sulfamide, chez les femmes.

## 2. Age et type de diabète.

Il est intéressant de comparer le résultat de la pyramide des âges de l'ensemble des diabétiques de la file active (Figure 33) avec celle de l'étude Entred [47]

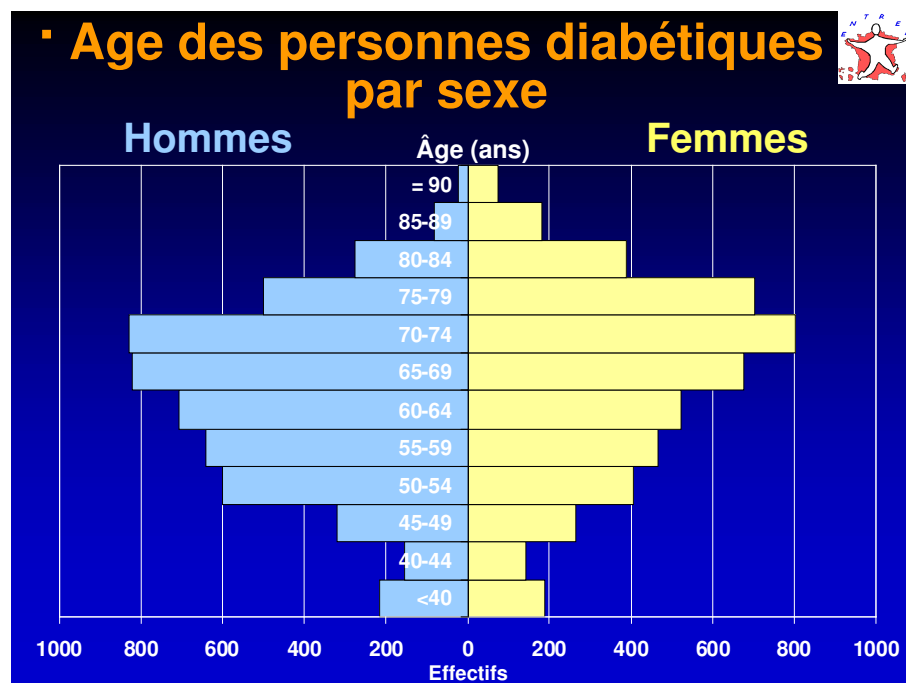


Figure 100 - Pyramide des âges des patients diabétiques de l'étude Entred

Si la répartition globale par tranche d'âge semble proche, nous remarquons que notre échantillon présente deux différences notables :

- la base de la pyramide semble décalée vers le bas
- la différence de répartition hommes/femmes par classe d'âge semble moins importante



Par comparaison nous pouvons affirmer que notre échantillon de patients diabétiques sur la file active, présente une répartition plus jeune et que la différence d'âges entre hommes et femmes, retrouvée dans l'étude Entred n'est pas retrouvée sur notre échantillon.

Si l'on compare les résultats trouvés (Figure 35 et Figure 37) par type de diabète, bien que ne disposant pas d'une réelle pyramide des âges dans l'étude Entred sur les différents types de diabète [47], on peut tout même voir une relative similarité des résultats concernant le diabète de type 1 :

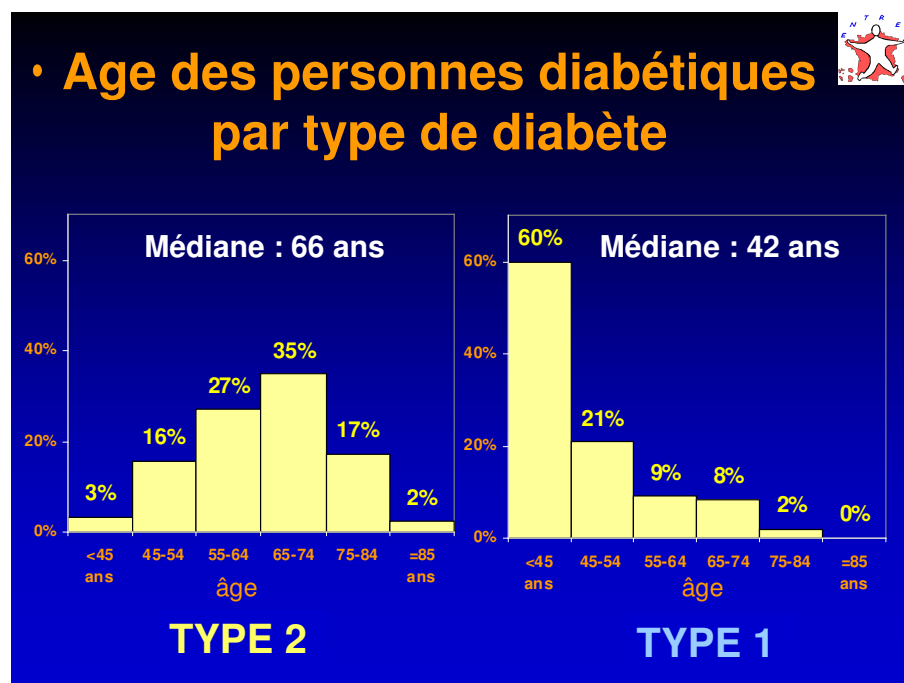


Figure 101 - Etude Entred : répartition des patients par type de diabète et par classes d'âge

Ce sont donc les caractéristiques du patient diabétique de type 2 qui semblent être différentes. Les patients diabétiques de type 2 de notre étude sont plus jeunes et notamment les femmes.

## ***D. Complications et morbidité***

### ***1. Complications Ophtalmologiques***

Selon notre base, 40,85% des diabétiques type 1 et 16,78% des diabétiques de type 2 présentent des complications ophtalmologiques (Figure 39).

Dans l'étude Entred [48], 3,5% de diabétiques tous types confondus ont répondu « oui » à la question « Avez-vous définitivement perdu la vue d'un œil ? » et

14,5% ont répondu « oui » à la question « Avez-vous déjà reçu un traitement par laser pour vos yeux ? ».

Les données ne sont pas superposables, notre définition de la complication est plus stricte que celle de l'étude Entred, qui ne prend en compte que le traitement de la complication ou son évolution péjorative terminale. De plus le mode de recueil est différent : dans notre cas, c'est une étude rétrospective portant sur un champ facultatif de formulaire, alors que dans l'étude des complications de l'étude Entred, il s'agit d'une étude prospective, sur échantillon, portant sur les déclarations du patient ou du médecin traitant dans un temps différé de l'événement.

Les résultats de l'étude J.Pirart [11, 12, 13] montrent que 50% des patients suivis sont atteints de rétinopathie après 25 ans de diabète.

Toutefois on notera qu'il s'agit d'une étude longitudinale, que l'on parle de sujet ayant 25 ans de diabète sans distinction de type, et que ce dénombrement s'est effectué entre 1947 et 1973. Dans notre cas, il s'agit d'une étude rétrospective, les patients inclus dans la file active ont des âges de diabète variables, le suivi s'est effectué à une période plus récente, où la prise en charge complète du diabète est plus efficace.

Il aurait été particulièrement intéressant de disposer, dans notre base, d'une échelle de gradation de la complication ophtalmologique. Malheureusement, cette information n'est actuellement pas recueillie.

## ***2. Complications rénales***

Nos résultats (Figure 43) montrent une complication rénale (selon les recommandations de la HAS [43]) chez 45,33% des diabétiques de type 1 et chez 41.51% des diabétiques de type 2.

Sur le site de l'INstitut de Veille Sanitaire, nous pouvons retrouver un diaporama présentant les complications retrouvées dans l'étude Entred [49]. L'étude Entred semble avoir retrouvé, chez 18% des personnes diabétiques de type 2 et 8% des diabétiques de type 1, un débit de filtration glomérulaire inférieur à 60 ml/min/1.73m<sup>2</sup>. Toutefois, nous n'avons pas pu retrouver la méthode d'obtention de ce résultat dans l'article [50] sur lequel ce diaporama s'appuie. Ce résultat serait à comparer avec les résultats que nous avons obtenus (Figure 42) après calcul de la clearance de la créatinine, qui retrouve, chez 23.3% des personnes diabétiques de type 2 et 20,4% des diabétiques de type 1, un débit de filtration glomérulaire inférieur à 60 ml/min/1.73m<sup>2</sup>.

Nos résultats sont nettement plus élevés, toutefois, il faut noter un premier biais de calcul puisque nous utilisons la formule de Cockcroft alors que les résultats de l'étude Entred ont été calculés par la formule du MDRD [51], ce que nous n'avons pu faire pour des raisons techniques. De plus, dans le cas de l'étude Entred, nous n'avons aucune information quant à l'obtention des données permettant ce calcul. Pour rappel, nous avons dû faire des choix afin de pouvoir effectuer notre calcul. (I.A.7.b- p72).

Les résultats de l'étude Pirart [11, 12, 13] déterminent la néphropathie diabétique sur la présence d'une protéinurie trouvée à plusieurs reprises par le réactif picrocitrique d'Ey Bach et/ou par la preuve histologique d'une glomérulosclérose nodulaire ou diffuse typique. Il retrouve ainsi seulement 14% de patients, tous types de diabète confondus, porteurs d'une complication rénale. Ces résultats seraient à comparer avec nos résultats de macro albuminurie qui sont de 9.36% chez le diabète de type 1 et de 10,64% chez le diabète de type 2.

Seul le croisement de l'ensemble des données nous a permis d'établir une approximation basée selon la définition établie par la HAS [43]. Ces résultats sont particulièrement élevés.

### **3. Complications neuropathie et pieds diabétiques**

Notre base de données retrouve :

- une neuropathie (Figure 44) chez 20.33% des diabétiques de type 1 et chez 23.50% des diabétiques de type 2. (Ces résultats proviennent du champ *complications* uniquement.)
- une complication trophique du pied (Figure 45) chez 23.88% des diabétiques de type 1 et chez 22.91% des diabétiques de type 2. (Ces résultats sont issus du croisement du champ *complications* et des consultations pour pieds diabétiques)

Les résultats de l'étude Entred retrouvent 6% de complications par mal perforant plantaire et 1,4% d'amputations [48]. L'étude Entred n'a pas publié de résultats sur les neuropathies en dehors du taux de recherche de la neuropathie.

Les résultats de l'étude Pirart [11, 12, 13] retrouvent une incidence de la neuropathie en fonction des années de diabète de 3 à 19% et une prévalence qui augmente avec les années de diabète et l'érosion de l'effectif de 9 à 50%.

Nos résultats présentent deux biais :

- les médecins examinateurs peuvent avoir une sensibilité et une spécificité différentes quant au diagnostic subjectif de neuropathie et à sa saisie
- les complications trophiques, peuvent être majorées du fait de l'ouverture spécifique d'une consultation « pied diabétique », dont le recrutement est individualisable.

Nous ne pouvons comparer nos résultats à la bibliographie, du fait que les méthodes de recueil dans le cadre de l'étude Entred différent et que l'étude Pirart -dont les méthodes de recueil sont proches- est une étude prospective longitudinale, contrairement à notre étude qui est rétrospective transversale. Toutefois, nous estimons que nos résultats sont cohérents.

#### **4. Complications cardiovasculaires et facteurs de risques**

Notre analyse de la base de données (Figure 50, Figure 51, Figure 52, Figure 54) retrouve :

- Une dyslipidémie, chez 17.53% des diabétiques de type 1 et chez 34.28% des diabétiques de type 2 (pour rappel il s'agit d'une évaluation basse)
- Une complication cardiaque, chez 2.23% des diabétiques de type 1 et chez 8.86% des diabétiques de type 2 (il s'agit ici du simple examen du champ *complications*, aucun croisement n'a pu être réalisé)
- Une complication AVC, chez 0.93% de diabétiques de type 1 et chez 1.81% des diabétiques de type 2. (il s'agit ici du simple examen du champ *complications*, aucun croisement n'a pu être réalisé)
- Une complication HTA, chez 11,75% des diabétiques de type 1 et chez 29.24% des diabétiques de type 2 (il s'agit ici du simple examen du champ *HTA*, aucun croisement n'a pu être réalisé)

Il est important de noter qu'en dehors des complications de dyslipidémies, nous subissons un biais de renseignement par le médecin examinateur.

Les résultats de l'étude Entred retrouvent [48] que 15,3% des personnes déclarent avoir eu un infarctus du myocarde ou un angor en répondant « oui » à la question: « Un médecin vous a-t-il dit que vous aviez fait un infarctus du myocarde, une crise cardiaque, de l'angor ou de l'angine de poitrine ? » (soit 306 000 diabétiques en France) et 8,5% déclarent avoir bénéficié d'une « intervention sur les artères du cœur ? (pontage coronarien ou angioplastie coronaire, pose d'un stent ou encore dilatation coronaire) », soit 170 000 diabétiques en France.

Ces chiffres, bien que non immédiatement comparables, peuvent être rapprochés. Il en ressort que l'impossibilité de différencier, sur le formulaire, une réponse « pas de complications » de l'absence de réponse, entraîne une sous-évaluation importante de certains champs ; toutefois, les champs que l'on peut confronter à d'autres champs (en particulier valeurs biologiques) échappent à ce constat.

## ***E. Indicateurs d'utilité***

### ***1. Préalables sur l'usage de l'HbA1c***

#### ***a- Du bon usage de l'HbA1c***

L'intérêt de l'hémoglobine A1c(HbA1c) comme marqueur rétrospectif de l'équilibre glycémique chez le patient diabétique est communément admis. Reflet cumulatif de la glycémie moyenne des quatre à six semaines qui précèdent le dosage, elle est utilisée en pratique courante pour évaluer de façon rétrospective l'efficacité du traitement, en complément des dosages glycémiques instantanés réalisés quotidiennement. Des objectifs thérapeutiques très précis, déterminés à partir de résultats d'études prospectives ou rétrospectives à grande échelle, ont été fixés à la fois dans le diabète de type 1 et dans le diabète de type 2 [6, 52, 53, 54]. Ces études ont permis de définir des objectifs thérapeutiques théoriques pour diminuer raisonnablement les risques de complications spécifiques et d'hypoglycémie. Ainsi, L'ALFEDIAM recommande, dans le diabète de type 1, des valeurs d'HbA1C comprises entre 7 et 7,5 %, et dans le diabète de type 2, un taux d'HbA1C inférieur à 6,5 %, pour un individu d'âge moyen et sans contre-indication au contrôle glycémique strict. [55]

Les opérations de contrôle de qualité menées en France depuis 1991 par la Société Française de Biologie Clinique (SFBC) et l'Agence du Médicament ont montré une grande inégalité de performance des méthodes de dosage, certaines étant compatibles avec un usage en biologie clinique, d'autres absolument pas. Ces enquêtes ont également révélé une grande dispersion des valeurs de référence indiquées par les laboratoires, génératrice de possibles erreurs d'interprétation des résultats [56, 57]. On peut ajouter à cela que, même si l'HbA1C doit être dosée spécifiquement, d'autres formes d'hémoglobine glyquée sont encore mesurées par certaines méthodes, ce qui engendre une confusion supplémentaire [58]. Il en résulte que, la plupart des diabétologues et des médecins prescripteurs n'accordent crédit qu'aux résultats d'un seul laboratoire, dont ils connaissent la technique et les valeurs de référence. L'amélioration de l'utilisation de ce dosage nécessite un travail portant sur la terminologie, la

standardisation des dosages, les critères de contrôle des méthodes et d'interprétation des résultats.[59]

Afin d'obtenir une standardisation de ses HbA1C, le laboratoire d'analyse de Roubaix s'est adapté aux recommandations de la SFBC [60], à savoir :

- Tous les résultats d'hémoglobine glyquée doivent être rendus sous forme d'HbA1C, exprimée en pourcentage de l'hémoglobine totale, à l'exclusion de tout autre mode d'expression.
- Les méthodes utilisées doivent soit doser directement l'HbA1C, soit pouvoir être reliées à une méthode de référence dosant l'HbA1C afin de corriger les valeurs brutes. Ceci inclut par exemple les méthodes par chromatographie d'affinité, qui dosent l'ensemble des hémoglobines glyquées [61].
- Les techniques utilisées par les laboratoires doivent faire la preuve de leur comparabilité à une méthode de référence recommandée par les sociétés scientifiques déjà citées : NGSP/DCCT ou IFCC. Il appartient aux biologistes de s'informer auprès des fournisseurs des conditions d'étalonnage de leurs méthodes, et d'exiger de ceux-ci un document de certification qui leur aura été fourni soit par le NGSP/DCCT soit par l'IFCC. Ces certifications garantissent que la méthode répond à des critères minimaux en matière de justesse et de précision.

Il faut insister sur le fait qu'il n'existe pas de méthode de référence disponible sur le marché qui soit utilisable pour les dosages courants, mais uniquement des méthodes ayant fait la preuve de leur comparabilité aux méthodes de référence proposées.

#### ***b- Stabilité du dosage au CH de Roubaix***

De nombreuses études reposent sur l'idée confortable que les outils utilisés sont fiables. On peut ainsi constater parfois le contraste entre d'une part des calculs à haute résolution et une discussion sur du détail numérique, et d'autre part des méthodes de mesure parfois peu fiables ou peu reproductibles, et en tout cas rarement testées. Cette remarque s'adresse essentiellement aux études rétrospectives, tant il est vrai que les investigateurs d'études prospectives prennent le soin de fixer une fois pour toutes les méthodes de mesure employées.

Dans les chapitres suivants, nous cherchons à déterminer des critères d'utilité de la prise en charge des patients diabétiques. Ces critères se basent essentiellement sur la mesure de l'hémoglobine glyquée / glycosylée, que nous dénommons indifféremment HbA1c, bien que cela soit approximatif. Nous avons

demandé au laboratoire de biologie du CH de Roubaix si les méthodes de dosage de l'HbA1c avaient changé au long de la période que nous étudions, soit 1988-2005. Nous rappelons que la base ne contient que des dosages réalisés par le laboratoire du CH de Roubaix. Les dosages réalisés en ville sont interprétés par les médecins mais non consignés.

L'énumération des méthodes de dosage utilisées pourra être lue dans les résultats (I.B Validité du dosage de l'hémoglobine glycosylée). Dans le même chapitre, nous montrons que le changement de méthode intervenu le 01/01/1994 fait probablement artificiellement baisser la mesure de l'HbA1c (Figure 69 - Box plots - distribution des hémoglobines glyquées, 50 jours avant et après le 01/01/1994, Tableau 16 – test de Student - comparaison des moyennes des hémoglobines glyquées 50 jours avant et après le 01/01/1994).

Nous aurions pu être tentés de modifier les dosages réalisés de l'HbA1c par une *application affine* afin de *centrer* les deux échantillons (« avant » et « après ») sur une moyenne commune, et de les *réduire* sur une variance commune. Nous n'avons pas effectué cette correction pour les raisons suivantes :

- en admettant que la différence soit significative, il n'est pas nécessaire d'aller jusqu'à l'identité des moyennes et des variances pour que la différence ne soit plus significative. Dès lors, nos deux échantillons ne représentant ni les mêmes malades, ni les mêmes instants, quelles valeurs utiliser pour définir nos corrections ?
- nous n'avons pas rencontré de telles corrections dans la bibliographie que nous avons explorée

Nous avons donc préféré nous abstenir et conserver ce biais à l'esprit lors de l'analyse des résultats. Nous noterons toutefois que lorsque l'élément temporel est la première consultation et non une date fixe, la date de changement de méthode devient alors une date mobile.

D'autre part, nous savons qu'en 1993 le CETRADIMN a invité l'ensemble des médecins correspondants présents dans sa base de données à un séminaire de formation, ayant pour but de les sensibiliser au dépistage précoce et à l'éducation thérapeutique du patient. 100 médecins correspondants ont ainsi été formés au cours de 4 séances.

Enfin, nous ne savons pas comment interpréter la remontée progressive des HbA1c dans les semestres suivant le 01/01/1994, la courbe « avant » semblant rattraper la courbe « après » (Figure 102 - tendance moyenne annuelle des HbA1c mesurées (tout confondu), Figure 103 – box plots - distribution annuelle des HbA1c mesurées (tout confondu)). On pourrait penser que le creux de 1994 est lié à cette conférence, et que son

effet s'est atténué par la suite. En effet, une modification liée à la biologie ne se serait peut-être pas amendée par la suite. Aucun élément numérique ne nous permet de trancher.

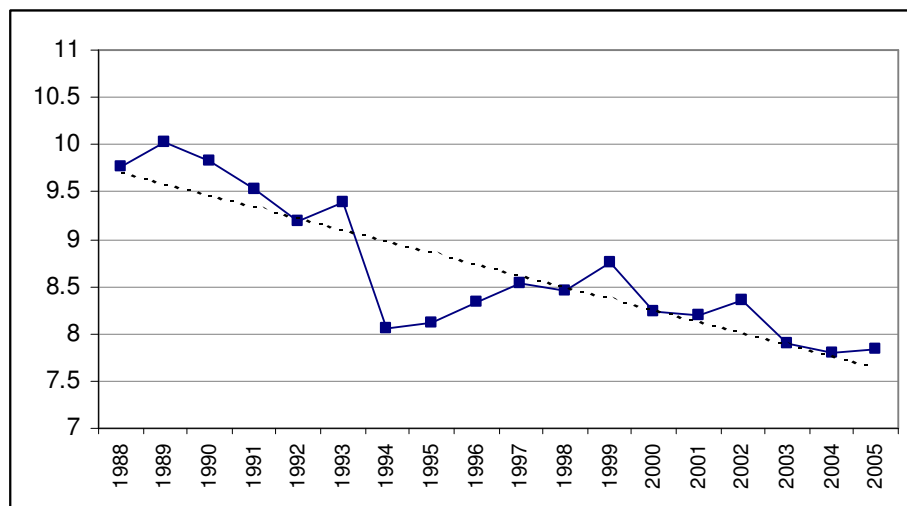


Figure 102 - tendance moyenne annuelle des HbA1c mesurées (tout confondu)

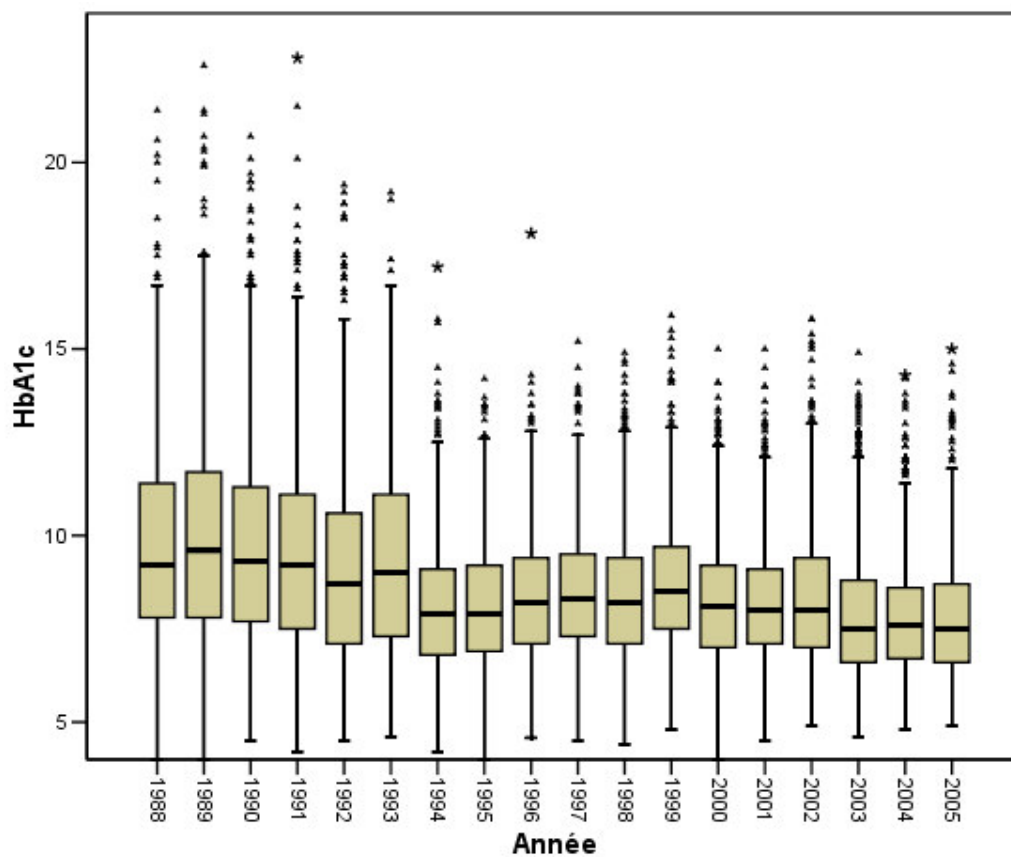


Figure 103 – box plots - distribution annuelle des HbA1c mesurées (tout confondu)



## **2. Utilité du suivi**

Nous avons tenu à comparer nos résultats de suivi des patients à l'indicateur d'évaluation défini par l'étude ENTRED [14] (Figure 71, p100). Toutefois, contrairement à l'étude ENTRED qui est partie du postulat qu'un patient hospitalisé bénéficiait systématiquement de son bilan annuel, nous n'avons pas eu besoin d'une telle approximation car notre base contenait tous les renseignements souhaités. De plus, nous rappelons que nous faisons une estimation basse qui ne tient pas compte du suivi déjà pratiqué par le médecin traitant du patient. Enfin en matière d'ECG et de fond d'œil, nous comptons bien les examens certains, et non les consultations de spécialité.

Le CETRADIMN présente un suivi nettement plus proche des recommandations de la HAS [17] que celui constaté sur l'étude ENTRED. Il est d'autant plus remarquable que le différentiel s'accroît avec le temps. Seul le suivi des trois HBA1C est proche des résultats de l'ENTRED. Toutefois, nous savons que d'autres dosages de l'HbA1c sont réalisés en ville à la demande du médecin traitant, ils ne peuvent être comptés ici.

On remarque que le CETRADIMN adopte une stratégie qui permet la réalisation complète du bilan annuel, laissant le reste du suivi aux soins du médecin traitant. Il s'agit là d'un partenariat clair qui permet une meilleur adéquation aux recommandations de la HAS.

Il est aussi intéressant de se comparer aux résultats 2003 exposés par les réseaux REVEDIAB, GENTIANE, DIABAIX et RESDIAB06 [62]. Nous retrouvons des résultats :

- de même ordre que les meilleurs résultats des réseaux dans le cadre du bilan lipidique et de la créatinine,
- nettement inférieurs dans le dosage des 3 HBA1C (probablement pour les causes sus cités)
- nettement supérieurs quant à la micro albuminurie, l'ECG et le Fond d'œil.

Cette comparaison reflète bien la différence de politique de prise en charge qui, pour le CETRADIMN, se situe bien dans la prise en charge du bilan annuel mais qui laisse le plein suivi du patient à son médecin traitant. Ceci permet une collaboration plus efficiente qui éviterait les oublis et redondance.

### **3. Utilité vis-à-vis des médecins traitants**

#### **a- Modèle**

L'hypothèse fondamentale est que, plus un médecin a déjà envoyé de patients au CETRADIMN, plus les relations qu'il entretient avec ses correspondants tendent à le sensibiliser au dépistage et au traitement du diabète de type 2. Cette sensibilisation pourrait être testée par une diminution de l'HbA1c des patients le jour de leur prise en charge au CETRADIMN. Dans la partie I.C.2 "Education" du médecin traitant nous testons l'effet suivant par une régression linéaire multiple :

- Variables explicatives :
  - X1=Année du recrutement du patient
  - X2=Numéro d'ordre du patient parmi les patients envoyés par le même médecin traitant
  - $\varepsilon$  = erreur résiduelle d'ajustement au modèle
- Variable expliquée :
  - Y=HbA1c du patient lors de son recrutement

$$Y = a.X1 + b.X2 + c + \varepsilon$$

#### **b- Interprétation**

Sous l'hypothèse de linéarité, les résultats de la régression linéaire s'interprètent au vu de l'intervalle de confiance à 95% (IC95) de la sorte :

- Le paramètre *a* décrit l'évolution des pratiques en ville, indépendamment de l'effet direct du réseau sur chaque médecin traitant :
  - IC95 strictement au dessous de zéro : les HbA1c lors du recrutement diminuent au fil du temps
  - IC95 strictement au dessus de zéro : les HbA1c lors du recrutement augmentent au fil du temps
  - IC95 encadrant zéro : les HbA1c lors du recrutement sont stables au fil du temps
- Le paramètre *b* décrit l'effet direct du réseau sur les pratiques de chaque médecin, indépendamment de l'évolution globale des pratiques :
  - IC95 strictement au dessous de zéro : les HbA1c lors du recrutement diminuent lorsque l'expérience entre le réseau et le médecin traitant augmente
  - IC95 strictement au dessus de zéro : les HbA1c lors du recrutement augmentent lorsque l'expérience entre le réseau et le médecin traitant augmente

- IC95 encadrant zéro : les HbA1c lors du recrutement ne sont pas modifiées par l'expérience entre le réseau et le médecin traitant

En résumé, si le paramètre  $b$  est significativement négatif (si son IC95 est strictement négatif), alors on déduit un effet d'éducation du médecin traitant par le réseau.

### ***c- Cas du CETRADIMN***

Dans le cas du CETRADIMN (voir I.C.2 "Education" du médecin traitant) nous ne mettons pas en évidence d'effet direct d'éducation du médecin traitant par le CETRADIMN au sujet du diabète de type 2. Toutefois, nous montrons une évolution globale des pratiques : les médecins traitants adressent leurs patients à des stades de la maladie de plus en plus précoces. Nous ne pouvons pas savoir si cela est l'effet des media et de la formation initiale et continue, ou bien l'effet d'une des campagnes d'information spécifiques initiées par le CETRADIMN à l'attention des médecins du versant Nord-Est.

### ***d- Perspectives***

L'indicateur que nous proposons ci-dessus pourrait simplement être calculé par les différents acteurs de la prise en charge ambulatoire (maisons du diabète, réseaux...) qui souhaiteraient évaluer leur effet d'éducation sur les médecins de ville.

A l'heure où les indicateurs quantitatifs d'utilité sont de plus en plus attendus, il nous paraissait opportun de formuler cette proposition.

## ***4. Utilité vis-à-vis des patients suivis***

### ***a- Remarques préalables***

Nous prenons le parti pris d'examiner l'évolution des hémoglobines glyquées des patients pris en charge par le CETRADIMN.

Dans le cas de la base que nous exploitons, il est très difficile d'examiner quantitativement l'évolution des hémoglobines glyquées des patients. Contrairement aux études prospectives, notre recueil présente les particularités suivantes :

- les patients sont suivis pendant des durées variables
- les patients disposent de mesures en nombre variable
- les mesures surviennent à des moments aléatoires :
  - à intervalles quasiment aléatoires
  - par conséquent sans synchronisation

Malgré les difficultés citées plus haut, il nous paraît opportun de poursuivre nos investigations car il importe d'évaluer l'utilité de la prise en charge en conditions réelles, or les obstacles que nous avons cités sont la réalité quotidienne.

## ***b- Approches descriptives***

### ***1) Nuage de points***

Dans un premier temps nous avons examiné des nuages de points, dans lesquels toutes les mesures participaient à poids égal, nonobstant leur rattachement à un patient. Un ajustement linéaire était à chaque fois réalisé (Figure 77 - nuage de points - HbA1c des diabétiques de type 1 au fil de la prise en charge ; Figure 78 - nuage de points - HbA1c des diabétiques de type 2 au fil de la prise en charge). Dans un tel ajustement, par construction, la participation de chaque patient à l'ajustement est d'autant plus forte qu'il vient souvent, et, à nombre de venues identique, qu'il vient longtemps après la première consultation. D'autre part, la diminution du nombre de points vers les abscisses élevées diminue la variance ; or la distribution n'est pas symétrique, les écarts diminuent donc plus pour les valeurs en excès que pour les valeurs en défaut et, l'ajustement s'appuyant sur les écarts carrés à la valeur prédite, cela conforte la décroissance d'un modèle d'ajustement linéaire. Nous ne pouvons donc pas interpréter l'apparente décroissance des HbA1c.

### ***2) Semestrialisation***

Dans un deuxième temps, nous arrondissons chaque date de mesure au numéro de semestre par excès. Nous examinons alors les HbA1c mesurées en fonction du semestre (Figure 79 – box plots - distribution de l'HbA1c des diabétiques type 1 au fil des semestres ; Figure 80 - courbe des moyennes - HbA1c des diabétiques type 1 au fil des semestres ; Figure 81 – box plots - distribution de l'HbA1c des diabétiques type 2 au fil des semestres ; Figure 82 - courbe des moyennes - HbA1c des diabétiques type 2 au fil des semestres). Cette autre approche présente les avantages suivants :

- les patients suivis longtemps sont certes les seuls à déterminer les abscisses hautes, en revanche leur poids dans les basses abscisses reste unitaire
- de manière générale le poids des patients est plus homogène : il dépend du nombre de mesures sur chaque paire de semestres consécutifs envisagée, donc de la seule densité, et non plus du nombre total de mesures
- l'évolution est suggérée plus finement qu'avec un ajustement linéaire
- la semestrialisation nous rapproche des problèmes classiques de régression : nous réalisons en quelque sorte une régression par paire de semestres, mais le calcul d'une ANOVA de régression n'a aucun sens puisque, sur seulement 2 abscisses, elle serait par essence « parfaite ».

En revanche les limites suivantes persistent :

- aucune preuve statistique n'est encore apportée à ce stade
- ces tendances moyennes ne rendent pas compte de la diversité des patients
- il existe un biais de sélection : les patients suivis de longues années sont probablement sérieux, le départ des autres patients optimise artificiellement l'évolution de l'HbA1c moyenne

### 3) Lissage

Dans un troisième temps, nous réalisons un lissage. Ce procédé permet une bonne visualisation de la tendance moyenne en évitant l'erreur liée à la discrétisation par semestre. En revanche, il ne met pas à disposition d'indicateur chiffré.

Stricto sensu, les lissages sont des méthodes de régression non paramétrique [44]. Nous présentons ci-dessous trois méthodes :

- La méthode Kernel est simplement une moyenne flottante calculée dans une fenêtre mobile dont la largeur est fixée par l'opérateur. Plus la fenêtre est large, plus la courbe est lisse. La méthode Kernel peut être améliorée en augmentant le poids des valeurs situées horizontalement en centre de fenêtre. [44]
- La méthode Spline suit un principe différent qui revient à pénaliser les valeurs qui rendraient la courbe moins lisse. [63]
- La méthode Lowess semble la préférée : il s'agit d'une méthode Kernel rendue moins sensible aux outliers. Cette méthode calcule dans chaque fenêtre un ajustement linéaire en augmentant le poids des données centrales. Puis le poids des valeurs dont les écarts carrés au modèle linéaire sont les plus élevés sont diminués. Ainsi, à chaque position de la fenêtre, un calcul itératif détermine un point finalement insensible aux outliers [64]. Le lissage Epanechnikov appartient à cette catégorie.

Les lissages, contrairement à l'ajustement linéaire, ne sont pas influencés outre mesure par les points d'abscisse élevée. Toutefois, les patients qui viennent le plus fréquemment contribuent plus à la formation de la courbe que les autres patients. Les résultats suggèrent une évolution linéaire biphasique :

- Chez les diabétiques de type 1, l'HbA1c décroît nettement durant les 2000 premiers jours, puis suit une lente décroissance (Figure 83 - lissage Epanechnikov - HbA1c des diabétiques type 1 , Figure 84 - lissage Epanechnikov - HbA1c des diabétiques type 1 au fil du suivi - ZOOM).
- Chez les diabétiques de type 2, l'HbA1c décroît nettement durant les 500 premiers jours, puis remonte lentement (Figure 85 - lissage Epanechnikov - HbA1c des

diabétiques type 2 , Figure 86 - lissage Epanechnikov - HbA1c des diabétiques type 2 au fil du suivi - ZOOM).

Cette intéressante description nous fait envisager une modélisation qui pourrait permettre le calcul d'indicateurs chiffrés. Contrairement à ce que nous pouvions craindre, l'allure du lissage dans les abscisses élevées suggère que les points d'abscisse élevée n'auraient pas de propension à fausser un ajustement linéaire.

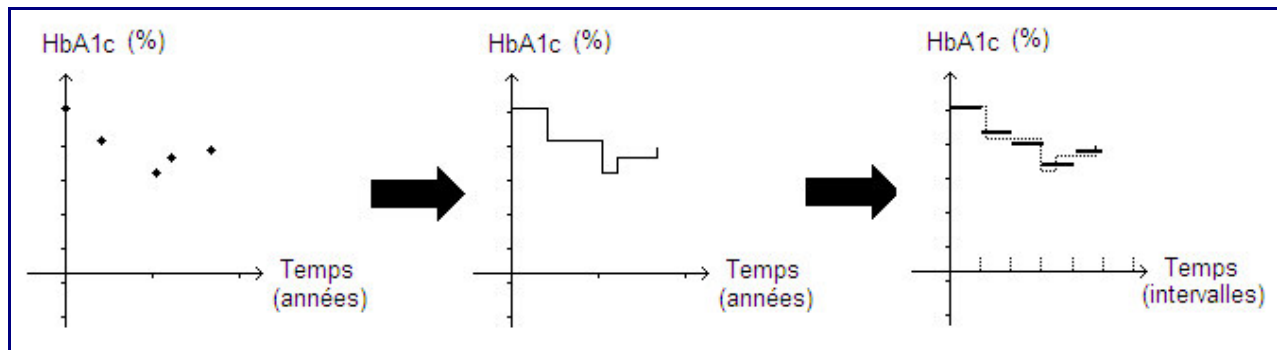
### **c- Analyses harmoniques**

#### **1) Principe**

Le but des analyses harmoniques est de décrire un profil d'évolution d'une variable dans le temps. Les analyses harmoniques s'accommodent de plusieurs caractéristiques qui sont habituellement des obstacles :

- les dates de mesure ne sont pas fixées
- le nombre de mesures par patient n'est pas fixé
- pour autant, le poids de chaque patient reste unitaire, nonobstant le nombre ou la densité des mesures
- en revanche, les patients doivent être suivis sur la même durée

Les analyses harmoniques supposent que la mesure soit fiable, et que la variation du résultat ne provienne pas d'une erreur de la mesure mais bien d'une variation de la variable évaluée. Elles reposent toutefois sur une interpolation « en marche d'escalier » de la variable entre deux dates de mesure. Dans un deuxième temps, l'analyse harmonique découpe chaque tracé de patient en intervalles réguliers fixés par l'utilisateur, et calcule une moyenne pondérée (Figure 104 - principe d'interpolation dans les analyses harmoniques).



**Figure 104 - principe d'interpolation dans les analyses harmoniques**

Il n'existe pas d'extrapolation : la courbe est interrompue à la dernière mesure. Enfin, une analyse en composantes principales est réalisée sur ces moyennes pondérées, qui constituent des variables. Dans notre cas, les variables sont

trimestrielles : notre période de suivi est de 10 ans, nous avons découpé la période en 40 intervalles réguliers.

## **2) Interprétation**

L'analyse harmonique permet d'interpréter directement des profils d'évolution des HbA1c des patients, à la lecture du premier plan factoriel. Pour les diabétiques de type 1, le premier plan factoriel (Figure 90 - positionnement des patients diabétiques type 1 sur le premier plan factoriel) explique 75.6% de la variance totale. Son interprétation est directe (Figure 91 - interprétation du premier plan factoriel - diabétiques type 1). Pour les diabétiques de type 2, le premier plan factoriel (Figure 95 - positionnement des patients diabétiques type 2 sur le premier plan factoriel) explique 76.66% de la variance totale. Son interprétation est encore plus directe dans ce cas précis (Figure 96 - interprétation du premier plan factoriel - diabétiques type 2).

## **3) Perspectives**

Les analyses harmoniques sont particulièrement adaptées à l'exploration de données biologiques évoluant avec le temps. La mise en évidence de différents profils d'évolution est un apport notable comparé aux analyses de données longitudinales classiques.

En revanche, bien qu'elles fournissent une réponse intuitive, elles ne permettent pas de fournir des indicateurs chiffrés d'utilité. Pour ce faire, il faudrait pouvoir ranger les patients dans des catégories d'évolution, et donc définir des seuils de significativité pour chaque composante principale (significativement négative, nulle, significativement positive), et ainsi quadriller le premier plan factoriel en neuf zones. Un décompte des effectifs de chaque zone pourrait constituer une sorte d'indicateur d'utilité, mais la définition des quatre seuils serait alors critiquable.

De plus, comme dans toute ACP, les composantes sont définies par l'échantillon, on ne pourrait donc pas définir de modèle fixe permettant de tester et comparer différentes populations, à moins de les analyser dans le même temps.

On peut en effet imaginer de produire une analyse harmonique dans le même temps sur deux populations à comparer ; ces deux populations seraient donc représentées sur le même plan unique ; les composantes seraient donc calculées sur la population totale. On pourrait ensuite tester la similitude des deux populations :

- soit par une approche qualitative : répartition des effectifs des deux populations dans les quadrants du plan factoriel
- soit par une approche quantitative : comparaison des moyennes des coordonnées sur une ou deux composantes

## **d- Analyses de données longitudinales**

### **1) Objectif**

En guise d'avertissement, nous précisons que le but de cette modélisation n'est pas de prédire l'HbA1c d'un patient donné à une date future, cet objectif serait déraisonnable. Notre but est de produire un indicateur numérique simple qui résume la partie la plus intéressante de la tendance moyenne. Ceci devrait permettre à des acteurs de la prise en charge ambulatoire du diabète de se comparer entre eux.

### **2) Modèle**

Les analyses de données longitudinales répondent au cas suivant : nous suivons plusieurs patients, nous dosons pour chaque patient un paramètre biologique (ici l'HbA1c) à des dates aléatoires, non synchrones, les mesures débutant à des moments aléatoires et ayant une durée aléatoire. Les analyses de données longitudinales proposent de modéliser ainsi un phénomène linéaire [44] :

$$y = a1.t0 + a2.(t-t0) + b + \varepsilon$$

y = variable à analyser, d'évolution supposée linéaire

t = date de la mesure

t0 = date de la première mesure chez le patient

$\varepsilon$  = erreur d'ajustement au modèle linéaire

La différence avec les approches suivies plus haut réside dans la prise en compte explicite de l'effet de la date d'inclusion t0.

Nous cherchons à modéliser cette linéarité en deux phases. Nous rappelons que les courbes d'évolution des HbA1c semblaient pouvoir se modéliser comme suit (VOIR Figure 83 - lissage Epanechnikov - HbA1c des diabétiques type 1 , Figure 85 - lissage Epanechnikov - HbA1c des diabétiques type 2) :



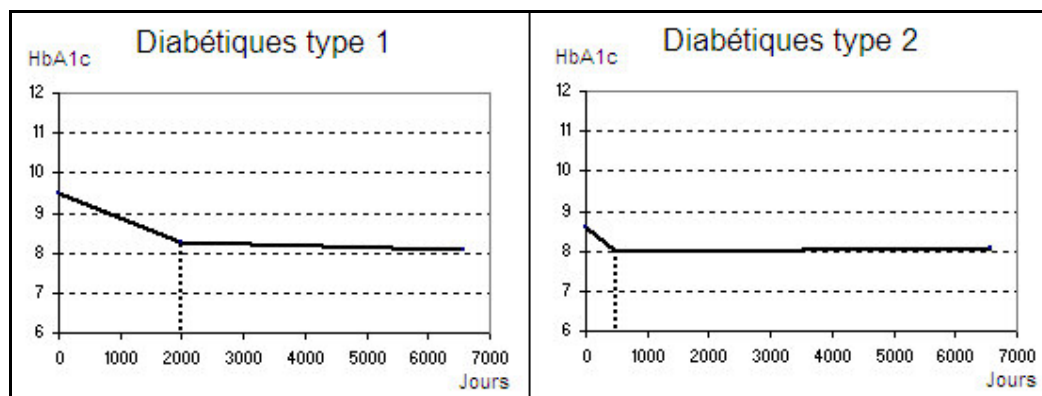


Figure 105 - modélisation longitudinale des HbA1c

Afin de tenir compte des deux phases dans le même modèle, que nous voulons synthétique et d'interprétation immédiate, nous créons des variables conditionnelles, dont nous écrivons ci-dessous les valeurs dans un formalisme ternaire, comme dans les langages SQL et Excel par exemple :

*SI( condition , valeur si vraie , valeur si fausse )*

Pour le calcul, les durées seront exprimées en années afin de faciliter l'interprétation.

$$Y = a1.t0 + a2.x2 + a3.x3 + a4.x4 + a5.x5 + \epsilon$$

#### Constante :

- @seuil = 2000 jours chez les diabétiques type 1
- @seuil = 500 jours chez les diabétiques type 2

#### Variables :

- t0 = date première mesure de l'HbA1c de ce patient
- t = date de la mesure de l'HbA1c
- x2 = SI( t-t0 <= @seuil , (t-t0) , 0)
- x3 = SI( t-t0 <= @seuil , 1 , 0)
- x4 = SI( t-t0 >= @seuil , (t-t0) , 0)
- x5 = SI( t-t0 >= @seuil , 1 , 0)
- y = HbA1c mesurée au temps t
- $\epsilon$  = erreur d'ajustement au modèle linéaire

#### Coefficients :

- a1 = pente générale liée à l'année d'inclusion
- a2 = pente du premier segment
- a3 = ordonnée à l'origine du premier segment
- a4 = pente du deuxième segment

- $a_5$  = ordonnée à l'origine du deuxième segment

### **3) Interprétation**

L'interprétation portera uniquement sur les pentes  $a_2$  et  $a_4$ , en particulier leurs intervalles de confiance à 95%. Nous précisons que les diabétiques peuvent se présenter à des stades divers de leur maladie. De ce fait, la comparaison des premiers segments (coefficient  $a_2$ ) peut poser des problèmes d'interprétation selon les habitudes de recrutement. En revanche, le paramètre  $a_4$  revêt une signification constante : il montre l'évolution à long terme, et mesure en particulier les phénomènes d'échappement ou d'évolution de la maladie.

Les deux variables binaires  $x_3$  et  $x_5$  apparaissent ici pour rappeler que les deux segments ne coupent pas l'axe des ordonnées au même point. Le logiciel ajoutera automatiquement une dernière constante, c'est pourquoi il est possible de supprimer l'une ou l'autre de ces variables. Il peut également arriver que le logiciel censure lui-même  $x_3$  ou  $x_5$ . C'est la raison pour laquelle les résultats présentés précédemment ne s'appuient pas exactement sur ce modèle (voir **Tableau 22 - Evolution des HbA1c des patients diabétiques au fil du suivi : proposition de modèle**). Quoi qu'il en soit l'interprétation directe n'est pas souhaitable.

La pente  $a_1$  sert à éliminer un facteur de confusion, son interprétation n'intervient pas directement dans l'évaluation de l'utilité.

### **4) Cas du CETRADIMN**

L'application du modèle aux données du CETRADIMN montre un excellent ajustement (**Tableau 23 - Modélisation des HbA1c des diabétiques type 1 : résultat**, **Tableau 24 - Modélisation des HbA1c des diabétiques type 2 : résultat**). On constate donc les tendances suivantes :

- les HbA1c des patients diabétiques type 1 diminuent pendant les 2000 premiers jours, puis continuent de diminuer lentement
- les HbA1c des patients diabétiques type 2 diminuent très fortement pendant les 500 premiers jours, puis continuent de diminuer très lentement.

Ces résultats nous semblent indiquer l'utilité de la prise en charge de ces patients par le CETRADIMN. En outre, nous ne constatons pas de remontée de l'HbA1c sur cet ajustement, alors qu'il est couramment admis que l'HbA1c des patients remonte inexorablement au fil du suivi. On peut penser qu'il s'agit là d'un biais de sélection, seuls les patients les plus sérieux étant suivis plus longtemps, pourtant le lissage des courbes ne semble pas montrer de modification vers les abscisses élevées.

## **5) Perspectives**

L'indicateur que nous proposons ci-dessus pourrait là aussi être simplement calculé rétrospectivement afin de quantifier l'évolution des glycémies des patients pris en charge. On peut supposer que l'évolution linéaire biphasique serait retrouvée dans de nombreux cas. Notons toutefois que le seuil (pour nous 500 jours en diabète type 2 et 2000 jours en diabète type 1) pourrait varier selon les habitudes de recrutement (en début ou non de prise en charge), il devrait alors être fixé empiriquement et perturberait la comparaison entre différents centres. Nous insistons toutefois sur l'idée que le coefficient  $a_4$  est de loin le plus intéressant, car il rend compte de l'évolution à long terme.

## **II. Un problème d'informatique médicale**

### **A. Les outils de la cohérence d'une base**

#### **1. Notions sur les contrôles de saisie**

Nous énonçons ici les principes qui devraient idéalement être appliqués chaque fois que possible dans la conception de formulaires destinés à alimenter une base de données :

➤ **Considérations générales :**

- le libellé de chaque champ devrait être clair et accompagné d'une info-bulle de définition
- les préfixes et suffixes constants devraient figurer à gauche et à droite du champ, afin d'éviter que l'utilisateur les saisisse parfois (exemple : pour le médecin traitant, écrire « docteur » à gauche du champ modifiable)
- certains champs non remplis devraient générer une alerte
- l'ordre d'enchaînement entre champs devrait être défini (touche tabulation)
- des sauts conditionnels devraient éventuellement exister entre pages de formulaire
- le cas échéant, le formulaire informatique devrait respecter l'ordre des documents papier qui constituent la source
- l'enregistrement et le non-enregistrement devraient être volontaires et explicites, si possible page par page
- afin de motiver les acteurs de la saisie et de contrôler les erreurs, un rendu immédiat devrait être proposé
- si le rendu révèle une erreur, elle devrait pouvoir être corrigée rapidement par lien direct sur la bonne page

- Champs numériques :
  - le libellé du champ devrait nécessairement comporter l'unité à utiliser
  - les caractères devraient être limités à {0123456789-} et {,} ou {.}.
  - une fourchette de validité devrait être imposée
  - certains contrôles de cohérence inter-champs pourraient être proposés
  - pour certains variables discrètes, une liste déroulante pourrait être proposée (exemple : nombre d'enfants, année de naissance...)
- Champs à motifs (date, numéro de téléphone, code postal, email...) :
  - un patron de saisie devrait être défini, si possible implémenté directement dans le champ, et dans tous les cas écrit dans le libellé (par exemple : *date de naissance (jj/mm/aaaa)* )
  - à défaut, une alerte bloquante devrait apparaître lorsque le champ perdant le focus ne respecte pas le patron
  - chaque composante du motif devrait être testée (exemple de la date de naissance : année comprise entre 1880 et 2005, mois compris entre 1 et 12, jour inférieur à 32 mais cohérent avec le mois et l'année)
  - dans certains cas, ce champ pourrait être décomposé en plusieurs sous-champs éventuellement munis de listes déroulantes
- Champs alphanumériques
  - chaque fois que possible, une liste déroulante devrait être proposée
  - lorsque le champ est multivalué mais que la liste peut être complétée par l'utilisateur, un schéma relationnel devrait être considéré afin de préserver l'intégrité des données (exemple du médecin traitant d'un patient)
  - lorsque le schéma relationnel n'a pu être utilisé, dans un formulaire idéal, il serait parfois utile de suggérer une correction automatique du champ lorsqu'une valeur très proche existe déjà (exemple classique de l'identité d'un patient)
  - il devrait rester possible d'affirmer qu'un champ est intentionnellement laissé vide :
    - pour les champs textuels libres, en leur associant une checkbox
    - pour les listes déroulantes, en incluant spécifiquement une valeur de signification nulle, et en forçant l'utilisateur à ne pas laisser le champ vide (exemple : traitement [pas de traitement/ ADO/ insuline/ADO+insuline] )

## 2. Schéma relationnel et intégrité des données

Comme expliqué plus haut (voir Tableau 6 - exemple de normalisation - contenu des tables , Tableau 7 - exemple de normalisation - structure des tables), seul un schéma relationnel garantit l'intégrité des données. Rappelons ce que signifie le mot « intégrité » :

- Intégrité lors de la création : dans notre exemple, si on ajoute un événement qui fait référence à un patient, l'opérateur est contraint de vérifier si le patient existe déjà et, si tel est le cas, crée un lien vers ce patient : il n'y a pas de recopie des propriétés de ce patient (nom, prénom, sexe, date de naissance) mais l'écriture d'un lien unique et univoque vers un enregistrement de la table des patients
- Intégrité lors de la modification : la modification des propriétés d'un patient affecte immédiatement et dynamiquement tous les événements se référant à ce patient, car il s'agit d'un lien et non d'une recopie
- Intégrité lors de la suppression : la définition de clefs étrangères en tant que contraintes d'intégrité (*constraint ... foreign key...*) interdit de supprimer un patient tant que des événements y font référence.

En revanche on conçoit combien l'administration d'un schéma relationnel peut paraître lourde, et interdit le design de bases locales indépendantes tel que le permettent les SGBD objet utilisés au CH de Roubaix.

Nous insisterons toutefois sur l'énorme apport que constitue dans ce contexte l'implémentation de l'IPP depuis 1995, qui, tel une clef étrangère (Patients.id=événements.id\_patient dans notre exemple), garantit l'intégrité au moins de la liaison patients-événements, à ceci près que les propriétés {nom, prénom, date de naissance, sexe} restent des copies locales certes exactes au moment de leur création, mais figées. De la sorte, seule l'intégrité lors de la création est assurée. Toutefois, les autres champs multi-valués, tels le médecin traitant, ne bénéficient pas d'un tel privilège.

## 3. Quelques préconisations supplémentaires pour NESSI

Actuellement, les formulaires mis en place dans CACHÉ reprennent strictement les anciens formulaires de SESAME. Nous pouvons suggérer d'y ajouter quelques préconisations formulées dans le chapitre II.A.1 Notions sur les contrôles de saisie. En particulier, nous insistons sur la nécessité de déclarer dans les listes multi-valuées des valeurs de signification vide, comme « pas de traitement », ainsi que sur la nécessité de préciser les unités de saisie des variables numériques, comme « taille en centimètres ».

Enfin, peut-être serait-il utile d'implémenter un système d'**identifiant personnel du médecin correspondant**, qui permettrait de centraliser sur l'hôpital la gestion

du carnet d'adresses des médecins en relation avec le CH, d'alléger ainsi les tâches redondantes de recherche d'adresses et de gestion des retours de courrier, et de recréer un quasi-schéma relationnel pour les liens médecin-événement ou, moins certainement, médecin-patient. Dans le cas des médecins libéraux, le numéro d'inscription à l'ordre des médecins pourrait fournir un identifiant provisoire ; ce numéro est rappelé sur chaque papier à entête (courrier, ordonnance).

### ***B. Système d'information et motivation des fournisseurs de données***

La mise en place d'un système de saisie d'informations est souvent mal vécue par les fournisseurs de données (que nous désignerons par habitude comme les utilisateurs, bien que ce terme soit ici ambigu) :

- Il représente une contrainte et des tâches supplémentaires
- L'utilisateur se sent méprisé dès que l'ergonomie n'est pas soignée ou en cas de dysfonctionnement
- L'utilisateur ne perçoit pas de bénéfice direct à la saisie
- L'utilisateur pense que son travail de saisie facilite le travail d'autres personnes qui, elles, n'ont pas de surplus de travail ; voire que sa saisie constitue un « flicage » de son travail
- Des saisies approximatives restent sans conséquence (exemple de l'utilisation au fil de l'eau, voir chapitre I.A.2.b-*Une utilisation au fil de l'eau / papier*)

Ces remarques valent tant pour l'utilisateur responsable (le médecin) que pour l'utilisateur effectif (la secrétaire lorsque la saisie est déléguée). L'exemple de la longue mise en place du PMSI à l'hôpital public est parlant : seule l'instauration de la tarification à l'activité (T2A) a permis une amélioration sensible de la qualité de codage en introduisant des conséquences budgétaires qui apparaissent à certains comme des sanctions.

Afin d'améliorer la qualité et l'exhaustivité de la saisie, il convient de veiller aux points suivants :

- Soigner au maximum le bon fonctionnement de l'application, communiquer sur les soins apportés au traitement des bugs (cellule d'aide, de signalement enregistrement et traitement des bugs, feed-back auprès de la personne qui a signalé l'anomalie, pendant et après résolution)
- Soigner au maximum l'ergonomie des applications, en utilisant les services de laboratoires d'évaluation, tels l'EVALAB de l'université Lille 2 : chaque

utilisation de l'application doit être aussi brève agréable et efficace que possible

- Former les utilisateurs à l'application
- Montrer aux utilisateurs l'utilisation qui est faite des données
- Ne pas demander de saisir une information dont tous savent qu'elle est inutile ; communiquer éventuellement sur l'utilité particulière de tel ou tel champ
- Proposer un retour le plus immédiat possible de la saisie, comme par exemple l'impression automatisée de courriers ou compte-rendus

L'opération de communication initiée au CH de Roubaix avant même le choix du système d'information est un bon exemple de politique pouvant impliquer l'utilisateur final. Il est également possible d'imaginer des systèmes d'intéressement pécuniaire de l'utilisateur, comme dans les entreprises privées :

- Après le calcul annuel d'un indice de précision ou d'exhaustivité, attribution d'une prime aux utilisateurs les plus zélés
- De la même manière que la boîte à idées de l'entreprise Peugeot Citroën SA, une prime pourrait être attribuée aux salariés formulant les propositions d'évolution ou les signalements de bugs les plus pertinents, une commission paritaire de suivi des effets des contributions pouvant être mise en place.

### ***C. Activité ou patients ? Place du patient dans le schéma relationnel***

La plupart des systèmes d'information sont centrés sur la notion de séjour, de passage, ou de consultation, et non de patient. Cette structure s'explique de la manière suivante :

- Les données recueillies ont essentiellement une vocation administrative, voire tarifaire [65]. L'information concerne donc bien le séjour ou la consultation, ou de manière plus générique l'événement. On ne s'étonnera donc pas de parler de « programme de médicalisation des systèmes d'information » et non de « programme d'informatisation de l'information médicale ».
- De nombreuses sources d'information plus ou moins formelles centrées autour du patient persistent à côté du système d'information (dossier papier, notes personnelles, base de données « illégale » dans le service, données non écrites...)
- L'anonymat des données est une préoccupation très ancrée dans la culture française. Les contraintes sont telles qu'elles représentent souvent un

obstacle majeur au partage d'information, à travers le concept de « données indirectement nominatives ». Nombreux sont ceux qui le jugent excessif en comparaison aux autres pays occidentaux, qu'il s'agisse du soin, du management de l'établissement, ou de la recherche médicale.

- La notion d'identité du patient pose un problème conceptuel lorsque le nombre d'identités est différent de 1 : dans des séjours légalement anonymes (IVG) ou dans des séjours ou consultations concernant plusieurs patients en même temps (séjour de grossesse et d'accouchement, consultation de couple...)
- La numéro d'affiliation à la sécurité sociale, objet de nombreux propos à l'emporte-pièce, ne peut être utilisé : sa promulgation est prohibée. De plus il ne représente pas une personne mais un bénéficiaire de prestations c'est-à-dire le plus souvent un salarié et ses ayant-droit (épouse et enfants). Une même personne peut changer de numéro d'immatriculation (passage à l'âge adulte, changement de caisse). Et enfin l'immatriculation française n'est pas obligatoire (étrangers, future mise en concurrence des caisses)

Cette orientation des systèmes d'information est telle que lorsqu'un décideur politique souhaite simplement connaître le nombre de patients souffrant de tel ou tel cancer, la réponse ne peut lui être donnée.

Dans le domaine du PMSI, nous signalerons toutefois la mise en place d'un numéro anonyme du patient permettant le chaînage des séjours hospitaliers nationalement par le logiciel MAGIC [66], ce numéro est abusivement appelé *numéro MAGIC*. Malheureusement, ce numéro ne concerne que le court-séjour MCO, et donc ni le moyen séjour SSR ni l'hospitalisation à domicile HAD ni le long séjour ni la psychiatrie, encore moins les consultations (en ville comme en établissement).

### **III. Education thérapeutique et financements**

L'éducation thérapeutique est au centre de la prise en charge réalisée au CETRADIMN. L'éducation thérapeutique en milieu hospitalier a un coût :

- elle consomme directement des ressources de personnel intervenant dans l'éducation
- elle consomme indirectement des ressources de personnel administratif (tâches de secrétariat supplémentaires, diminution des tâches liées à l'activité « rentable » des soignants)



- si elle est efficace, elle diminue le nombre de séjours, la lourdeur des séjours, et le nombre de consultations des patients qui deviennent plus autonomes et dont les complications apparaissent moins vite.
- elle fait appel à des modalités non couvertes par les tarifs ambulatoires : consultations d'infirmières, séquences d'éducation en groupe

Dans un contexte de convergence tarifaire future entre les secteurs public et privé, il est donc naturel que cette mission soit couverte par les MIGAC en sus des tarifs des séjours et des consultations.

Le financement MIGAC alloué au CETRADIMN correspond à 31% des ressources totales du CETRADIMN en 2003 (Tableau 28 - données financières), soit un supplément de 45% par rapport à la dotation globale redistribuée par le CHG.

Ce ratio qui paraît important est à rapprocher des importants moyens mis en œuvre (Tableau 26 - les séquences proposées en groupe de patients au CETRADIMN).

Par ailleurs, le CETRADIMN a procédé à un chronométrage des tâches d'éducation thérapeutique (Tableau 27 - Temps consacré à l'éducation du patient en mai 2005, sur la base de 6 semaines d'HDJ). Nous notons les éléments suivants :

- ce temps ne rend pas compte du temps administratif (secrétariat) qui découle de cette observation
- le temps médical utilisé est de 35h, bien que ce temps n'ait pas de valeur légale
- ce temps ne prend pas en compte les consultations d'infirmières

Ce temps est décompté sur la base de 6 semaines d'hospitalisation de jour. Nous ne pouvons pas mettre ce temps en regard des financements. Pour ce faire, il faudrait connaître sur une semaine type *toutes* les activités menées (consultations facturables, consultations officieuses, hospitalisation classique, séquences non facturables...) par toutes les catégories de personnel, une simple extrapolation pouvant être réalisée pour le personnel administratif. Le temps déjà facturé directement ne saurait être inclus. Ensuite, en disposant des salaires moyens par catégorie de personnel, ceci permettrait de calculer un coût réel des activités d'éducation. Nous ne sommes pas en mesure de le faire ici.

#### ***IV. Modèle conceptuel de la prise en charge du patient diabétique***

Le développement de l'éducation thérapeutique du patient s'est développé à partir de l'extension du concept de santé passant d'un modèle médical pur à un modèle global ou holistique.

Le modèle médical est un modèle fermé où la maladie est principalement organique, elle affecte l'individu, elle doit être diagnostiquée et traitée (approche bio-médicale), la gestion est faite par les Médecins dans un système autonome centré autour de structures de soins dirigées par des Médecins et par une Institution organisée.

Ce modèle médical s'oppose au modèle global qui est un modèle holistique et ouvert. Dans ce modèle, la maladie résulte de facteurs complexes, organiques, humains et sociaux (souvent c'est le domaine de la maladie chronique). Elle affecte l'individu, la famille et l'environnement, et demande une approche continue de la prévention à la réadaptation. Cette approche doit tenir compte des facteurs organiques, psychologiques et sociaux,. C'est pourquoi elle nécessite des professionnels de santé travaillant en étroite collaboration, dans un système ouvert et interdépendant avec la communauté.

Ce concept de santé global couplé au concept de communication doit permettre à un soignant de communiquer avec un patient pour lui transmettre des messages de santé tout en évaluant, dans un deuxième temps, sa portée tant en ce qui concerne la connaissance pure, que le savoir-faire ou le changement de comportement (savoir, savoir-faire, savoir-être).

Le soignant doit analyser les déterminants liés aux patients : facteurs démographique, socio-économiques, psychologiques, représentations de la santé et de la maladie, acceptation de celle-ci, capacités intellectuelles et physiques... Suite à cette analyse, le soignant acquiert les pouvoirs, en fonction des besoins du patient et des objectifs médicaux, par une négociation avec le patient lui-même, de définir des objectifs de traitement à court terme, moyen terme et long terme.

Si bien sûr tous ces éléments sont quelques peu théoriques, différentes équipes nord-américaines et européennes, à partir des années 1970, les ont mis en pratique de manière efficace.

La modélisation de l'évolution du traitement du diabète est instructive quant aux progrès en éducation thérapeutique du patient et du suivi à long terme.

En effet, jusque dans les années 1970, le modèle de traitement du diabète était quasi-purement bio-médical et technique. C'est en 1921, avec la découverte de l'insuline que le pronostic du diabète a été révolutionné et que le taux de mortalité chez les diabétiques a diminué de manière très importante. Les antibiotiques et les anti-diabétiques oraux ont permis dans un deuxième temps de consolider ces progrès mais c'est à Léona Miller, à Boston aux Etats-Unis que l'on doit les premières études sur l'impact de l'éducation du patient sur le nombre de journées d'hospitalisation classique.

Grâce à un Centre fonctionnant totalement en ambulatoire, créé en 1973, Léona Miller a montré que la prise en charge des patients, dont les options principales étaient centrées sur l'éducation thérapeutique, permettait de diminuer le nombre de journées d'hospitalisation par patient et par an de 5,1% à 1,7%, très proche du nombre moyen d'hospitalisation par personne et par an chez des sujets diabétiques (1,2 %). Jean-Philippe Assal a pu montrer que l'éducation des diabétiques à risque (éviter la pression de lésions) a permis d'obtenir des résultats spectaculaires tant au niveau médical pur qu'au niveau de la qualité de vie des patients. Ainsi, il a pu, dans une étude prospective sur quatre ans, montrer qu'une telle éducation permettait de réduire de 80% les amputations au-dessus du genou chez les patients à risque, résultat qui ont été confirmés secondairement par d'autres études.

A l'époque, l'éducation thérapeutique du patient était principalement justifiée comme une réponse à la non-compliance. Le concept de compliance se conçoit en terme d'alliance thérapeutique et d'adhésion. L'adhésion est le résultat d'une négociation entre le soignant et le soigné ; la compliance-adhésion est donc le niveau de respect, de comportement thérapeutique et préventif librement consenti par le patient.

On sait par ailleurs que la non-compliance augmenterait avec l'âge et le nombre de médicaments prescrits (notamment supérieur à 3), ainsi 25 à 50 % des personnes âgées ne seraient pas compliantes. Les conséquences de la non-compliance au traitement ont des effets néfastes bien sûr, sur la santé du patient elle-même mais également sur son coût. Ainsi, la non-compliance augmente les périodes de maladies, favorise l'évolution de celle-ci en précipitant les complications, induit une augmentation des journées d'hospitalisation. Par ailleurs, elle favorise la perte de confiance et l'insatisfaction du patient, entraînant une surconsommation de soins et quelquefois l'utilisation technique parallèle. Enfin, elle augmente les coûts indirects de la santé en majorant les arrêts de travail, voire en induisant une invalidité.

La facilitation des soins en favorisant l'accès aux soins et en développant une recherche pharmacocinétique visant à faciliter la prise médicamenteuse est certainement un élément important. Pour autant, l'éducation thérapeutique du patient, qui dans son essence même peut être considérée comme une forme d'éducation pour la santé dirigée vers la gestion de la maladie, est une autre solution.

Mais comme nous l'avons expliqué antérieurement, cette éducation thérapeutique a un coût puisqu'elle mobilise de nombreux personnels et qu'elle réduit le nombre de journée d'hospitalisation de manière nette, ce qui pénalise l'institution

hospitalière. Le constat est simple à faire : sans un financement spécifique, l'éducation thérapeutique a et aura beaucoup de mal à se développer à un niveau adapté au développement des maladies chroniques et notamment des maladies de la nutrition, du diabète et des maladies générant un risque cardio-vasculaire. Le concept d'éducation thérapeutique du patient ira beaucoup plus loin sous l'impulsion de Jean-Philippe Assal qui, à la suite des études dans le domaine du diabète de l'UKPDS et du Sténo-Institut, développera le concept du suivi à long terme des patients chroniques.

L'efficacité médicale est incontestablement très grande dans le diagnostic et le traitement des situations aiguës ou sub-aiguës: il n'en est pas de même au niveau de la qualité des soins dans la chronicité. C'est la gestion de la monotonie issue de la chronicité qui nécessite une nouvelle approche et le développement de nouveaux savoir-faire que les médecins doivent acquérir pour améliorer la qualité des soins. C'est ainsi que Jean Philippe ASSAL introduit une réflexion sur la maladie chronique : "un long fleuve tranquille avec ses sables mouvants" [26, 69, 70, 71, 72, 73, 74, 75].

## CONCLUSION

En guise de bilan et de mise en perspectives, nous parcourons les trois axes de ce travail.

### ***I. Au sujet du recueil de données***

Nous avons montré les difficultés rencontrées, et comment des outils avancés permettaient toutefois de « faire parler » une base qui dort. Nous jugions utile d'exposer ces outils en détail dans la mesure où de nombreux services hospitaliers disposent de base de données à la fois riches en contenu et peu exploitables dans l'immédiat.

Les fruits de ce lourd travail de nettoyage dépassent toutefois cette seule thèse. Le dédoublement des patients que nous avons réalisé a été transmis sous forme d'une table d'agrégation aux informaticiens du CHG, cette table servira à rapatrier certains anciens dossiers dans les nouveaux dossiers dotés d'un IPP valide (postérieurs à 1995).

Nous noterons toutefois que l'exploitation faite ici se limite aux aspects biomédicaux. Cette première réflexion apporte des lumières au recueil prospectif qui sera prochainement mis en place au CETRADIMN. Ce recueil s'articulera autour de 5 axes plus larges de la prise en charge :

- La gravité biomédicale, seule prise en compte classiquement
- La chronologie d'apparition des complications : évolutivité du diabète
- Le handicap et ses répercussion psychologiques
- Un indice social
- Un indice prédictif de la qualité de vie, la qualité de vie étant un constat trop tardif par définition

### ***II. Au sujet de l'analyse de l'activité du CETRADIMN***

Nous avons montré la spécificité du recrutement de patientèle du CETRADIMN, aux lumières des particularités sanitaires et sociales du Versant Nord-Est de la Métropole Lilloise. Cette vision s'est voulue plus complexe que les classiques dénombrements. Nous regrettons de n'avoir pu exploiter de manière plus

dynamique l'évolution de la maladie : les variables non biologiques se seraient volontiers accommodées d'analyses de survie si les dates de survenue avaient été fiables.

En perspectives, il serait intéressant de quantifier réellement le coût de l'éducation thérapeutique du patient, et ce d'autant plus que le financement des MIGAC promet d'être de plus en plus fondé sur les surcoûts réellement démontrés.

### ***III. Au sujet des indicateurs d'utilité***

Nous avons proposé et testé des indicateurs numériques d'utilité, permettant d'évaluer le changement des pratiques des médecins de ville induit par le travail en réseau, le bénéfice de la prise en charge au regard des recommandations de suivi, et enfin le bénéfice de la prise en charge au regard de l'évolution à long terme de marqueurs biomédicaux.

Si ces approches sont certainement perfectibles, elles n'en demeurent pas moins intéressantes, à l'heure où les financements devraient de plus en plus s'appuyer sur les résultats chiffrés, et non plus sur une logique de moyen où les sommes investies paraissent être le seul indicateur d'efficacité.

## Lexique des sigles

ADO	<i>Anti-Diabétiques Oraux</i>
ANAES	<i>Agence Nationale d'Evaluation en Santé, actuelle HAS</i>
ATIH	<i>Agence Technique de l'Information Hospitalière</i>
BMI	<i>Body Mass Index, indice de masse corporelle mesuré par la formule (poids en Kg)/((taille en m)<sup>2</sup>)</i>
CETRADIMIN	<i>Centre d'Education pour le Traitement du Diabète et des Maladies de la Nutrition</i>
CH, CHG	<i>Centre Hospitalier, Centre Hospitalier Général</i>
CLI	<i>Command Line Interpreter, moteur de PHP capable d'interpréter un script en ligne de commande, en architecture client seul.</i>
CNAM TS	<i>Caisse Nationale d'Assurance Maladie des Travailleurs Salariés</i>
CSP	<i>Caché Server Pages, langage utilisé par le SGBD Caché pour générer des pages web (formulaires et états)</i>
DID	<i>Diabète Insulino-Dépendant, type 1</i>
DIM	<i>Département d'Information Médicale</i>
DIT	<i>Diabète Insulino-Traité</i>
DNID	<i>Diabète Non Insulino-Dépendant, type 2</i>
ENTRED	<i>Echantillon National Témoin REprésentatif des personnes Diabétiques, étude de la CNAMTS</i>
FTP	<i>File Transfert Protocole, protocole de transfert de fichiers</i>
GUI	<i>Graphic User Interface, programmes interagissant avec l'utilisateur grâce à des conventions graphiques (fenêtres, boutons, souris). Ces programmes s'opposent aux programmes en ligne de commande.</i>
HAD	<i>Hospitalisation à domicile</i>
HAS	<i>Haute Autorité en Santé</i>
HbA1c	<i>(ce n'est pas un sigle) Hémoglobine glyquée, ou hémoglobine glycosylée</i>
HDL	<i>High Density Lipoprotein, lipoprotéine de haute densité</i>
HMJ	<i>Hyperglycémie Modérée à Jeun</i>

HTML	<i>HyperText Markup Language</i> , langage permettant de décrire, à l'aide de texte brut agencé sous forme de balises, des mises en forme complexes. Le HTML est le langage utilisé par presque toutes les pages internet.
HTTP	<i>HyperText Transfert Protocol</i> , protocole de transfert de données utilisé par les navigateurs web
IC95	<i>Intervalle de Confiance à 95%</i>
IPP	<i>Identifiant Permanent du Patient</i> , numéro unique pour un patient dans un établissement donné, rappelé pour chacun de ses séjours. Ce numéro varie entre établissements différents.
LDL	<i>Low Density Lipoprotein</i> , lipoprotéine de basse densité
MERRI	<i>Missions d'Enseignement, Recherche, Référence et Innovation</i>
MIGAC	<i>Missions d'Intérêt Général et d'Aide à la Contractualisation</i>
MCO	<i>Médecine Chirurgie Obstétrique</i> , hospitalisation court séjour
PHP	<i>Php Hypertext Preprocessor</i> , langage de programmation orienté objet dans sa version 5.
SGBD	<i>Système de Gestion de Base de Données</i> , logiciel serveur de base de données, ce terme s'entendant tant côté interrogation et modification des données, qu'administration.
SPSS	<i>Statistic Program for Social Sciences</i> , logiciel de statistiques.
SQL	<i>Structured Query Language</i> , langage structuré d'interrogation des bases de données, présentant un fort niveau d'abstraction, implémenté par tous les produits phares du marché. Le SQL est défini par des standards que les éditeurs décident de respecter plus ou moins scrupuleusement.
SSR	<i>Soins de Suite et de Réadaptation</i> , moyen séjour
T2A	<i>Tarifification A l'Activité</i>
VNE	<i>Versant Nord-Est de la métropole lilloise</i>
W3C	<i>Word Wide Web Consortium</i> , organisme mondial en charge de la normalisation des langages internet, dont HTML et XML.



## Bibliographie

Avertissement : la bibliographie intègre de nombreuses références techniques, dont le formalisme déroge à la règle de présentation.

- [1] Brindel P, Fagot-Campagna A, Petit C, Coliche V, Isnard H, Simon D, Varroud-Vial M, Weill A. Entred : lancement d'une étude sur un échantillon national de 10 000 personnes diabétiques traitées. *BEH* 2002; 20-21:88-9.
- [2] Livinec F, Weill A, Varroud-Vial V, Simon D, Vallier N, Fagot-Campagna A. Caractéristiques régionales de la population diabétique traitée et de sa prise en charge médicale. Entred 2001. *BEH* 2003; 49-50:240-42.
- [3] Diaporama sur l'étude ENTRED, Institut National de Veille Sanitaire [http://www.invs.sante.fr/publications/entred/html/methodologie\\_fichiers/frame.htm](http://www.invs.sante.fr/publications/entred/html/methodologie_fichiers/frame.htm)
- [3bis] Questionnaire de l'étude ENTRED [www.invs.sante.fr/publications/entred/](http://www.invs.sante.fr/publications/entred/)
- [4] Site officiel de l'Ancred <http://www.ancred.org>
- [5] UK Prospective Diabetes Study (UKPDS) Group. Effect of intensive blood-glucose control with metformin on complications in overweight patients with type 2 diabetes (UKPDS 34). [published erratum appears in *Lancet* 1998 Nov 7;352(9139):1557]. *Lancet* 1998, 352, 854-65.
- [6] UK Prospective Diabetes Study (UKPDS) Group. Intensive blood-glucose control with sulphonylureas or insulin compared with conventional treatment and risk of complications in patients with type 2 diabetes (UKPDS 33). [published erratum appears in *Lancet* 1999 Aug 14;354(9178):602]. *Lancet* 1998, 352, 837-53.
- [7] UK Prospective Diabetes Study (UKPDS) Group. Cost effectiveness analysis of improved blood pressure control in hypertensive patients with type 2 diabetes: UKPDS 40. *Bmj* 1998, 317, 720-6.

- [8] UK Prospective Diabetes Study (UKPDS) Group. Efficacy of atenolol and captopril in reducing risk of macrovascular and microvascular complications in type 2 diabetes: UKPDS 39. *Bmj* 1998, 317, 713-20.
- [9] UK Prospective Diabetes Study (UKPDS) Group. Tight blood pressure control and risk of macrovascular and microvascular complications in type 2 diabetes: UKPDS 38. [published erratum appears in *BMJ* 1999 Jan 2;318(7175):29]. *Bmj* 1998, 317, 703-13.
- [10] UK Prospective Diabetes Study (UKPDS) Group. UKPDS 28: a randomized trial of efficacy of early addition of metformin in sulfonylurea-treated type 2 diabetes. *Diabetes Care* 1998, 21, 87-92.
- [11] J.Pirart. Diabète et complications dégénératives présentation d'une étude prospective portant sur 4400 cas observés entre 1947 et 1973. *Diabetes Metab* , 1977, 3, 97-107
- [12] J.Pirart. Diabète et complications dégénératives présentation d'une étude prospective portant sur 4400 cas observés entre 1947 et 1973. *Diabetes Metab* ,1977 , 3, 245-256
- [13] J.Pirart. Diabète et complications dégénératives présentation d'une étude prospective portant sur 4400 cas observé entre 1947 et 1973. *Diabetes Metab* ,1977 , 3, 173-182
- [14] Fagot-Campagna A, Simon S, Varroud-Vial M, Ihaddadène K, Vallier N, Scaturro S, Eschwège E, Weill A, pour le comité scientifique d'Entred. Caractéristiques des personnes diabétiques traitées et adéquation du suivi médical du diabète aux recommandations officielles. Entred 2001. *BEH* 2003; 49-50:238-39.
- [15] Ricordeau P, Weill A, Vallier N, Bourrel R, Fender P, Allemand H. L'épidémiologie du diabète en France métropolitaine. *Diabetes Metab* 2000, 26:11-24.
- [16] Données issu du SROS III, sur la base du rapport du COTER de diabétologie sur « la prise en charge du diabète dans la région Nord-Pas de Calais » de juin 2004 présidé par le Pr Fontaine CHRU Lille.
- [17] Comité d'experts de l'Agence nationale d'accréditation et d'évaluation en santé. Stratégie de prise en charge du patient diabétique de type 2 à l'exclusion de la prise en charge des complications. Recommandations de l'ANAES. *Diabetes Metab* 2000; 26

- [18] Conférence sur : versant Nord-Est de la Métropole : un territoire de santé, de la reconnaissance à la réalité, vendredi 16 septembre 2005, Sous le patronage de l'ARH, à l'initiative des CHG de Roubaix, Tourcoing et Wattrelos
- [19] Deccache A et al , la santé des milieux défavorisés. Synthèse des données épidémiologiques, sociologiques, culturelles et psychologiques collectées dans le cadre du projet « Education pour la santé en milieux défavorisés. ». Bruxelles, janvier 1990
- [20] Colin C, Lavoie Jp, Poulin C, Les personnes défavorisées. Et la santé, ça va ? *Les publications du Québec*, Québec, 1989, 119p
- [21] Schaetzel F, Querouil O, Etat de santé des bénéficiaires du RMI : synthèse des données existantes, novembre 1991
- [22] Dahlgren G et Whitehead M, politiques et stratégies en faveur de l'égalité devant la santé, *OMS*, Copenhague, 1992
- [23] Dussart AF , Prevost M, les comportements alimentaires des adultes en Belgique et dans la CEE, *Santé pluriel* n°6, 1991
- [24] Grenier JL, Bulle M, Delepaut A, La santé dans nos quartiers c'est la santé de nos quartiers.
- [25] Grenier J.L, Qafli M, Gross P. Bilan du CETRADIMN après 10 ans de fonctionnement en alternative à l'hospitalisation. *Diabetes Metab* , 1993, 19, 528-532
- [26] Jacquemet S. Le patient peut-il apprendre à devenir un malade chronique. *Bulletin d'éducation du patient* vol 27 n°3 sept 1998
- [27] Lawrence Philips. Metaphone algorithm  
<http://aspell.sourceforge.net/metaphone/>
- [27] Le CHG de Roubaix récupère 20 ans de données médicales grâce à Caché, témoignage sur le site Intersystems/Caché  
[http://209.61.190.221/france/proweb/system/downloadPage.csp?filename=Cas\\_CHGroubaix-.pdf&dataID=4686](http://209.61.190.221/france/proweb/system/downloadPage.csp?filename=Cas_CHGroubaix-.pdf&dataID=4686)
- [28] Russell R, Odell M. Soundex algorithm, *U.S. Patent* 1,261,167 & *U.S. Patent* 1,435,663
- [29] Knuth D. The Art of Computer Programming, Reading, Massachusetts : *Addison-Wesley*

- [30] Binstock, Rex. Practical Algorithms for Programmers, *Addison Wesley*, 1995
- [31] MySQL 5.0 Notice : installation, langage, utilisation, implémentation du SQL et tutoriaux d'introduction au SQL en général  
<http://downloads.mysql.com/docs/refman-5.0-fr.chm>
- [32] PHP5 Notice : installation, langage, utilisation  
[http://fr.php.net/get/php\\_manual\\_fr.chm/from/this/mirror](http://fr.php.net/get/php_manual_fr.chm/from/this/mirror)
- [33] CACHÉ, base de données post-relationnelle de la société Intersystems  
<http://209.61.190.221/france/proweb/frontend/start.csp?pageID=59>
- [34] Haute Autorité en Santé. Compte rendu d'accréditation du centre hospitalier de Roubaix, mai 2004
- [35] CLINICOM, Système d'Information Clinique communiquant centré sur le patient, plaquette de présentation  
[http://www.siemens.com/Daten/siecom/France/MED/Internet/MED\\_Unitwide/WORKAREA/edshs3/templatedata/Francais/file/binary/Fiche\\_Clinicom\\_1143551.pdf](http://www.siemens.com/Daten/siecom/France/MED/Internet/MED_Unitwide/WORKAREA/edshs3/templatedata/Francais/file/binary/Fiche_Clinicom_1143551.pdf)
- [36] ROUbaix Logiciel d'Aide au Codage (CCAM v0-v2, CIM10)  
<http://www.roublac.com/>
- [37] Sophie MANNESSIEZ. Passage à la CCAM v2 au CH de Roubaix grâce à Roublac
- [38] MySQL, système de gestion de bases de données, documentation et téléchargement  
<http://dev.mysql.com/>
- [39] EasyPHP, programme d'installation d'Apache PHP et MySQL pour Windows, documentation et téléchargement  
<http://www.easyphp.org/>
- [40] MySQL Query Browser, client graphique léger de requêtes SQL  
<http://dev.mysql.com/downloads/query-browser/>
- [41] PHP: Hypertext Preprocessor, documentation et téléchargement  
<http://www.php.net/>

- [42] HTML HyperText Markup Language : spécifications syntaxiques par le W3Consortium  
<http://www.w3.org/MarkUp/>
- [43] ANAES. Service des recommandations et références professionnelles  
Septembre 2002 : Diagnostic de l'insuffisance rénale chronique chez l'adulte
- [44] Diggle PJ, Heagerty P, Liang KY, Zeger SL. Analysis of Longitudinal Data, *Oxford university press*, ISBN 0 19 852484 6
- [45] Preda CA. Approximare în Analiza Armonică, teză de doctorat, Academia Română, centrul de Statistică Matematică "Gheorghe Mihoc"
- [46] Sclar DA, Robison LM, Skaer TL, Dickson WM, Kozma CM, Reeder CE. Sulfonylurea pharmacotherapy regimen adherence in a Medicaid population: influence of age, gender, and race. *Diabetes Educ.* 1999 Jul-Aug;25(4):531-2, 535, 537-8.
- [47] Institut National de Veille Sanitaire, diaporama sur l'étude ENTRED  
[http://www.invs.sante.fr/publications/entred/html/caracteristiques\\_fichiers/frame.htm](http://www.invs.sante.fr/publications/entred/html/caracteristiques_fichiers/frame.htm)
- [48] Prévalence des complications macro vasculaires et niveau de risque vasculaire des diabétiques en France, étude Entred 2001 », *BEH* n° 12-13/2005
- [49] Institut National de Veille Sanitaire, diaporama sur l'étude ENTRED  
[http://www.invs.sante.fr/publications/entred/html/complications\\_fichiers/frame.htm](http://www.invs.sante.fr/publications/entred/html/complications_fichiers/frame.htm)
- [50] Romon I, Fosse S, Weill A, Varroud-Vial M, Fagot-Campagna. Prévalence des complications macrovasculaires et niveau de risque vasculaire des diabétiques en France, étude Entred 2001 », *BEH* n° 12-13/2005.
- [51] Froissart M. Modification of Diet in Renal Disease, *J Am Soc Nephrol* 2005
- [52] Goldstein DE, Little RR. Monitoring glycemia in diabetes, Short term assessment. *Curr Therap Diab*, 1997, 26,475-86.
- [53] Bernard M, Bardas-Fonfrède M, Grimaldi A, et al. Intérêts respectifs des dosages d'hémoglobine glyquée et de fructosamines dans la surveillance du diabète sucré. *Ann Biol Clin*, 1995, 53,321-7.

- [54] The Diabetes Control and Complications Trial Research Group. The effect of intensive treatment of diabetes on the development and progression of long-term complications in insulin-dependent diabetes mellitus. *N Engl J Med*, 1993, 329,977-86.
- [55] Drouin P, Blicke JF, Charbonnel B, et al. Diagnostic et classification du diabète sucré. Les nouveaux critères. *Diabetes Metab*, 1999,25,72-83.
- [56] Gillery P, Labbé D, Dumont G, Vassault A. Glycohemoglobin assays evaluated in a large-scale quality-control survey. *Clin Chem*, 1995,41,1644-8.
- [57] Gillery P, Dumont G, Vassault A. Evaluation of GHb assays in France by national quality control surveys. *Diabetes Care*, 1998, 21,265-70.
- [58] Gillery P, Guillemain C, Delpech M. Hémoglobine glyquée méthodes de dosage et problèmes de standardisation. *Ann Biol Clin*, 1994, 52,157-63.
- [59] Gillery P, Bordas-Fonfrède M, Chapelle Jp, Drouin P, Hue G, Levy-Marchal C, Périer C, Sélam JL, Slama G, Thivolet C, Vialettes B. Rapport des experts de l'ALFEDIAM et de la SFBC. HbA1C : concertation clinico-biologique pour la standardisation des méthodes de dosage. *Diabetes métab*, 1999, 25, 283-287
- [60] Gillery P, Bordas-Fonfrède M, Chapelle J.P, Hue G, Périer C. Hémoglobine glyquée le temps de la standardisation est venu. *Ana Biol Clin*, 1998, 56,249-51.
- [61] Nuttall FQ. Comparison of percent total GHb with percent HbA1C in people with and without known diabetes. *Diabetes Care*, 1998,21,1475-80.
- [62] Varroud-Vial M. Données régionales d'Entred et évaluation des réseaux [Symposium]. Congrès de l'Association de langue française pour l'étude du diabète et des maladies métaboliques, Lyon, 22-26 mars 2005.
- [63] Silverman B.W. Spline smoothing : the equivalent variable Kernel method, *Annals of Statistics*, 12, 898-916
- [64] Cleveland W.S. Robust locally weighted regression and smoothing scatterplots, *Journal of the American Statistical Association*, 74 829-36
- [65] Agence Technique d'Information Hospitalière, ATIH, site officiel <http://www.atih.sante.fr>

- [66] CIRCULAIRE DHOS-PMSI n° 106 du 22 FEV. 2001 relative au chaînage des séjours en établissements hospitalier
- [67] Pons M. Validity of the diagnostic tests for diabetes Birules, *Aten Primaria* 2004 Sep 30;34(5):222-8.
- [68] UK Prospective Diabetes Study (UKPDS) Group. Etude UKPDS *Lancet* 1998; 352: 837-853 – *Lancet* 1998 ; 352 : 854-865
- [69] Assal JP, Golay A. Le suivi à long terme des patients chroniques, les nouvelles dimensions du temps thérapeutique. *Médecine et Hygiène* ; n° 2353 ; 27 juin 2001 :1446.
- [70] Lacroix A, Assal JP. L'éducation thérapeutique du patient, nouvelle approche de la maladie chronique. Editions Vigot, Paris, 1998
- [71] Deccache A, Lavendhome E. Information et éducation du patient. Editions Deboeck, Université de Bruxelles, 1989
- [72] D'Ivernois JF, Gagnaire R. Apprendre à éduquer le patient. Editions Vigot, Paris, 1995
- [73] Massé R. Analyse anthropologique et éthique des conflits de valeurs en promotion de la santé. Université de Laval, Québec
- [74] Education pour la santé et l'éthique. Séminaire international Doudan, 27-28 janvier 2000. Editions CFES.
- [75] Lecorps P. Education pour la santé des patients, un enjeu pour le système de santé. Le patient comme sujet. Ecole Nationale de Santé Publique de Rennes. Colloque européen, hôpital européen Georges Pompidou les 10 et 11 juin 1999, Paris. Editions CFES.

## Figures

Figure 1 - adéquation du suivi des diabétiques aux recommandations de la HAS	12
Figure 2 - évolution du chômage sur le territoire du Versant Nord-Est de la métropole lilloise (VNE)	15
Figure 3 - proportion de minima sociaux sur le VNE (source CAF 2004)	15
Figure 4 - Carte de l'indice de défaveur sociale sur 13 variables (ORS déc. 2004)	16
Figure 5 - illustration des pertes de données avant 1999 (ici avec l'HbA1c)	29
Figure 6 - architecture de CLINICOM, © Siemens Health Systems	30
Figure 7 - Fonctionnement de l'application	43
Figure 8 - Notre application : lancement	45
Figure 9 - Notre application : parsing et exécution automatisée de requêtes d'interrogation	46
Figure 10 - Evolution du nombre de consultations, d'hôpital de jour et d'hôpital de semaine.	49
Figure 11 - Evolution du nombre de patients distincts pris en charge chaque année.	49
Figure 12 - Nombre de nouveaux patients chaque année	50
Figure 13 - Part des nouveaux patients sur les patients distincts vus chaque année	51
Figure 14 - recrutement des nouveaux patients par année et par diagnostic	52
Figure 15 - Fuite et fidélisation des nouveaux patients	53
Figure 16 - taux de fuite par pathologie	54
Figure 17 – box plots - distribution par diagnostic des durées de suivi	55
Figure 18 – box plots - distribution par diagnostic des durées de suivi, patients venus plus d'une fois	55
Figure 19 - dénombrements des médecins traitants par indice de familiarité sur l'ensemble des patients	57
Figure 20 - dénombrements des médecins traitants par indice de familiarité sur la file active.	57
Figure 21 - répartition hommes/femmes sur la totalité de la base.	58



Figure 22 - Répartition par diagnostic sur l'ensemble de la base.	59
Figure 23 - évolution du nombre de nouveaux patients par type de pathologie.	60
Figure 24 - évolution du nombre de nouveaux patients par type de diabète	61
Figure 25 - répartition hommes/femmes de la file active	61
Figure 26 - Pyramide des âges de la file active	62
Figure 27 - répartition des patients de la file active par diagnostic.	63
Figure 28 - Nombre de venues moyen par type de pathologie	64
Figure 29 - box plots - distribution du nombre de venues par patient selon le diagnostic.	65
Figure 30 – box plots - distribution, par diagnostic, de l'intervalle moyen entre les venues du patient	66
Figure 31 - Répartitions par type de diabète des patients diabétiques de la file active.	67
Figure 32 - Patients diabétiques de la file active, représentation de l'âge moyen et de la répartition hommes/femmes	68
Figure 33 - Pyramide des âges des patients diabétiques de la file active	68
Figure 34 - répartition hommes/femmes des diabétiques de type 1 de la file active.	69
Figure 35 - Pyramide des âges des diabétiques type 1 de la file active	69
Figure 36 - répartition hommes/femmes des diabétiques type 2 de la file active.	70
Figure 37 - Pyramide des âges des diabétiques de type 2 de la file active	70
Figure 38 - Pyramide des âges des patientes présentant un diabète gestationnel	71
Figure 39 - dénombrement des complications ophtalmologiques	72
Figure 40 - Dénombrement des complications rénales sur le champs complications.	72
Figure 41 - Dénombrement des complications rénales sur le champs albuminurie	73
Figure 42 - Complications rénales sur le champ clearance de la créatinine.	74
Figure 43 - Complications rénales synthèse.	75
Figure 44 - dénombrement des complications neuropathies	76
Figure 45 - dénombrement des complications de pied	77
Figure 46 - dénombrement des complications d'artérite	77
Figure 47 - dénombrement des complications de dyslipidémie sur le champ complications	78
Figure 48 - dénombrement des complications hypercholestérolémie sur le champ LDL.	79

Figure 49 - dénombrements des complications hypertriglycéridémie sur le champ triglycérides.	79
Figure 50 - synthèse des dénombrements des complications de dyslipidémie	80
Figure 51 - dénombrements des complications coronarite ou infarctus du myocarde par type de diabète	81
Figure 52 - dénombrement des complications d'AVC par type de diabète	81
Figure 53 - dénombrement des complications surcharge pondérale par type de diabète	82
Figure 54 - dénombrement des complications d'HTA par type de diabète.	83
Figure 55 - synthèse des patients ayant une complication par type de diabète	83
Figure 56 - Evolution de la première HBA1C en fonction du temps	84
Figure 57 - box plots - distribution des HbA1C lors du recrutement en fonction de l'année de recrutement chez l'ensemble des diabétiques	85
Figure 58 - Nuage de points - HbA1C en fonction de l'année de recrutement sur l'ensemble des diabétiques.	85
Figure 59 - Evolution de l'HBA1C de recrutement des diabétiques type 1 et type 2 en fonction de l'année de recrutement	87
Figure 60 - Box plot - distribution des HbA1C lors du recrutement en fonction de l'année de recrutement chez les diabétiques de type 2	87
Figure 61 - Nuage de points - HbA1C en fonction de l'année de recrutement sur les diabétiques de type 2.	88
Figure 62 - Evolution de la première glycémie à jeun en fonction de l'année de recrutement chez l'ensemble des diabétiques	89
Figure 63 – Box plots - distribution des premières glycémies à jeun selon l'année de recrutement, tous diabétiques	90
Figure 64 : nuage de points - distribution des premières glycémie à jeun selon l'année de recrutement, tous diabétiques	90
Figure 65 - Evolution de la première glycémie à jeun en fonction de l'année de recrutement des diabétiques de type1 et de type2.	92
Figure 66 - Box Plots - distribution des première glycémies à jeun selon l'année de recrutement, diabétiques de type2	93
Figure 67 - Nuage de points - distribution des première glycémies à jeun selon l'année de recrutement, diabétiques de type2	93
Figure 68 - Evolution de la proportion des patients diabétiques de type 1 insulinés et non insulinés à l'entrée.	95
Figure 69 - Box plots - distribution des hémoglobines glyquées, 50 jours avant et après le 01/01/1994	96

Figure 70 - Box plots - distribution des hémoglobines glyquées, 50 jours avant et après le 01/04/2004	97
Figure 71 - Comparatif du suivi des recommandations HAS selon les années par rapport à l'étude ENTRED 2001	100
Figure 72 - Box plots - première HbA1c du patient en fonction de son numéro d'ordre	101
Figure 73 - Nuage de points - première HbA1c du patient en fonction de son numéro d'ordre	102
Figure 74 - Box Plots - année de recrutement du patient en fonction de son numéro d'ordre	103
Figure 75 - Box Plots - première glycémie à jeun du patient en fonction de son numéro d'ordre	104
Figure 76 - nuage de points - première glycémie à jeun du patient en fonction de son numéro d'ordre	105
Figure 77 - nuage de points - HbA1c des diabétiques de type 1 au fil de la prise en charge	106
Figure 78 - nuage de points - HbA1c des diabétiques de type 2 au fil de la prise en charge	107
Figure 79 – box plots - distribution de l'HbA1c des diabétiques type 1 au fil des semestres	108
Figure 80 - courbe des moyennes - HbA1c des diabétiques type 1 au fil des semestres	108
Figure 81 – box plots - distribution de l'HbA1c des diabétiques type 2 au fil des semestres	109
Figure 82 - courbe des moyennes - HbA1c des diabétiques type 2 au fil des semestres	109
Figure 83 - lissage Epanechnikov - HbA1c des diabétiques type 1 au fil du suivi	110
Figure 84 - lissage Epanechnikov - HbA1c des diabétiques type 1 au fil du suivi - ZOOM	111
Figure 85 - lissage Epanechnikov - HbA1c des diabétiques type 2 au fil du suivi	112
Figure 86 - lissage Epanechnikov - HbA1c des diabétiques type 2 au fil du suivi - ZOOM	113
Figure 87 - première composante principale - diabétiques type 1	114
Figure 88 - deuxième composante principale - diabétiques type 1	114
Figure 89 - contribution des variables trimestrielles au premier plan factoriel - diabétiques type 1	115

Figure 90 - positionnement des patients diabétiques type 1 sur le premier plan factoriel	115
Figure 91 - interprétation du premier plan factoriel - diabétiques type 1	116
Figure 92 - première composante principale - diabétiques type 2	116
Figure 93 - deuxième composante principale - diabétiques type 2	117
Figure 94 - contribution des variables trimestrielles au premier plan factoriel - diabétiques type 2	117
Figure 95 - positionnement des patients diabétiques type 2 sur le premier plan factoriel	118
Figure 96 - interprétation du premier plan factoriel - diabétiques type 2	118
Figure 97 - Evolution de l'activité médicale du CETRADIMN sur la base des données administratives	121
Figure 98 : Evolution de l'activité paramédicale du CETRADIMN sur la base des données administrative	122
Figure 99 - Evolution des effectifs médicaux et paramédicaux du CETRADIMN	123
Figure 100 - Pyramide des âges des patients diabétiques de l'étude Entred	128
Figure 101 - Etude Entred : répartition des patients par type de diabète et par classes d'âge	129
Figure 102 - tendance moyenne annuelle des HbA1c mesurées (tout confondu)	136
Figure 103 – box plots - distribution annuelle des HbA1c mesurées (tout confondu)	136
Figure 104 - principe d'interpolation dans les analyses harmoniques	142
Figure 105 - modélisation longitudinale des HbA1c	145

## Tableaux

Tableau 1 - traitement des patients diabétiques en France .....	11
Tableau 2 - densité au Km <sup>2</sup> des villes du versant Nord est (source INSEE) .....	17
Tableau 3- recopie automatique et blocs - exemple 1/3 .....	22
Tableau 4 - recopie automatique et blocs - exemple 2/3 .....	22
Tableau 5 - recopie automatique et blocs - exemple 3/3 .....	23
Tableau 6 - exemple de normalisation - contenu des tables.....	36
Tableau 7 - exemple de normalisation - structure des tables.....	37
Tableau 8 - effectifs utilisés pour les représentations de durées de suivi .....	56
Tableau 9 - Dénombrement des patients par diagnostic .....	59
Tableau 10 - dénombrement des patients de la file active par diagnostic.....	63
Tableau 11 - dénombrements par type de diabète des patients diabétiques de la file active .....	66
Tableau 12 - régression linéaire simple - évolution de l'HbA1c lors du recrutement en fonction du temps .....	86
Tableau 13 – Régression simple - Evolution de l'HbA1c des patients lors de leur recrutement en fonction de l'année de recrutement.....	88
Tableau 14 - régression linéaire simple - évolution de la première glycémie à jeun selon l'année de recrutement, tous diabétiques .....	91
Tableau 15 - régression linéaire simple - évolution de la première glycémie à jeun selon l'année de recrutement, diabétiques type 2 .....	94
Tableau 16 – test de Student - comparaison des moyennes des hémoglobines glyquées 50 jours avant et après le 01/01/1994 .....	96
Tableau 17 – test de Student - comparaison des moyennes des hémoglobines glyquées avant et après le 01/04/2004.....	98
Tableau 18 - Suivi des diabétiques et adéquation aux recommandations de la HAS .....	99
Tableau 19 - régression linéaire simple - première HbA1c du patient en fonction de son numéro d'ordre .....	102
Tableau 20 - régression linéaire multiple - première HbA1c du patient en fonction de l'année et du numéro d'ordre.....	104

Tableau 21 - régression linéaire simple - première glycémie à jeun du patient en fonction de son numéro d'ordre .....	105
Tableau 22 - Evolution des HbA1c des patients diabétiques au fil du suivi : proposition de modèle .....	119
Tableau 23 - Modélisation des HbA1c des diabétiques type 1 : résultat.....	119
Tableau 24 - Modélisation des HbA1c des diabétiques type 2 : résultat.....	120
Tableau 25 - Récapitulatif des données administratives d'activité du CETRADIMN .....	121
Tableau 26 - les séquences proposées en groupe de patients au CETRADIMN .....	124
Tableau 27 - Temps consacré à l'éducation du patient en mai 2005, sur la base de 6 semaines d'HDJ .....	124
Tableau 28 - données financières .....	125

**DESEINE Anne-Sophie**  
**CHAZARD Emmanuel**  
**LEGRAND Bertrand**

**Titre : Etude rétrospective de 17 ans d'activité du Centre d'Education pour le TRAitement du Diabète et des Maladies de la Nutrition du CHG de Roubaix : traitement de base de données, analyse, indicateurs d'utilité**

**Thèse, Médecine, Lille, 2006, n°**

**Cadre de classement :** Santé publique, Médecine générale

**Mots clés :** diabète, évaluation, informatique, rétrospectif

**Résumé :**

Le CETRADIMN (Centre d'Education pour le Traitement du Diabète et des Maladies de la Nutrition), rattaché au CHG de Roubaix, propose une prise en charge ambulatoire complète des patients diabétiques, en complément avec la médecine de ville. Le CETRADIMN dispose d'une base de données épidémiologique renseignée depuis 1988, conçue dans un système de gestion de base de données orienté objet. Ce travail s'articule autour de trois axes :

1- réaliser une extraction et un nettoyage de la base de données existante afin de permettre une exploitation rétrospective et participer à la migration des anciennes données vers le nouveau système d'information ; exposer les outils de récupération d'une base ; discuter des moyens permettant de garantir la cohérence et l'intégrité des données

2- porter un regard rétrospectif statique et dynamique sur l'activité, les patients, les pratiques et le réseau de médecins correspondants du CETRADIMN, plus particulièrement dans le domaine de la prise en charge des patients diabétiques ; exposer les spécificités du recrutement du CETRADIMN

3- proposer et tester in situ des indicateurs d'utilité, à l'aide de méthodes statistiques de régressions, d'analyses harmoniques et d'analyses de données longitudinales. Nous proposons et testons un indicateur de formation des médecins correspondants, un indicateur d'application des recommandations de suivi de la HAS et, après une modélisation des profils individuels et moyens d'évolution de l'HbA1c, un indicateur de contrôle glycémique à long terme.

**Composition du jury :**

Président : Monsieur le Professeur SALOMEZ

Assesseurs : Monsieur le Professeur BEUSCART  
Monsieur le Professeur FONTAINE  
Monsieur le Professeur GLANTENET  
Monsieur le Docteur GRENIER

Directeur de Thèse : Monsieur le Docteur GRENIER

**Adresse des auteurs :**

Mme DESEINE,	8 rue Albert Samain, 59700 Marcq en baroeul	asdeseine@yahoo.fr
M CHAZARD,	appt. 541, 6 rue Carnot, 59110 La Madeleine	emmanuelchazard@yahoo.fr
M LEGRAND,	8 rue Albert Samain, 59700 Marcq-en-Baroeul	bertrandlegrand59@yahoo.fr