ÉCOLE DE TECHNOLOGIE SUPÉRIEURE
UNIVERSITÉ DU QUÉBEC

RAPPORT DE PROJET PRÉSENTÉ À
L'ÉCOLE DE TECHNOLOGIE SUPÉRIEURE

COMME EXIGENCE PARTIELLE
À L'OBTENTION DE LA
MAÎTRISE EN GÉNIE, CONCENTRATION
TECHNOLOGIES DE L'INFORMATION

PAR
MOHAMED ELYES BEN ABDELKRIM

# APPLICATION POUR LA GESTION D'UNE BASE DE DONNÉES CENTRALISÉE POUR UN LABORATOIRE DE RECHERCHE EN SANTÉ

MONTRÉAL, LE 20 NOVEMBRE 2016

.

**PRÉSENTATION DU JURY**

CE RAPPORT DE PROJET A ÉTÉ ÉVALUÉ

PAR UN JURY COMPOSÉ DE :

Prof. Alain April, directeur de projet
Département de Génie Logiciel et des technologies de l'information
École de technologie supérieure

Prof. Abdelaoued Gherbi, jury
Département de Génie Logiciel et des technologies de l'information
École de technologie supérieure

**REMERCIEMENTS**

Au terme de ce travail, je tiens à exprimer ma profonde gratitude et mes sincères remerciements à mon directeur de projet, professeur Alain April, professeur au département Génie Logiciel de l'ÉTS pour toute son aide.

Je remercie également Mme Anita Franco, assistante de recherche sénior au laboratoire Laboratoire Viscogliosi en Génétique Moléculaire des Maladies Musculo-Squelettiques, de l'hôpital Sainte-Justine pour son soutien et sa disponibilité.

Je remercie Christian St-Laurent, étudiant à la maîtrise en génie des technologies de l'information, qui a réalisé la première itération de ce projet.

Mes profondes gratitudes s'orientent vers mes parents Faiza EL SOUSSI et Rachid BEN ABDELKRIM pour leur support inestimable. Mes profonds remerciements vont aussi pour le membre de jury qui a accepté d'évaluer ce travail.

APPLICATION POUR LA GESTION D'UNE BASE DE DONNÉES CENTRALISÉE POUR UN LABORATOIRE DE RECHERCHE EN SANTÉ

Mohamed Elyes BEN ABDELKRIM

## RÉSUMÉ

L'exploitation et l'analyse des données cliniques, dans les centres de recherche hospitaliers, sont des activités de plus en plus importantes dans le processus de recherche. Pour cela les laboratoires de recherche médicaux utilisent de plus en plus de technologies afin d'obtenir des gains de performance, et ce selon leurs budgets. C'est dans ce contexte que le docteur Alain Moreau, chercheur principal du Laboratoire Viscogliosi en Génétique moléculaire des Maladies musculosquelettiques de l'hôpital Sainte-Justine, a demandé au professeur Alain April de l'École de technologie supérieure (ÉTS) d'établir une collaboration afin de concevoir une base de données centralisée contenant les différentes données utilisées par le laboratoire.

Ce rapport décrit la réalisation de la deuxième itération de ce projet qui consiste à la sélection, l'extraction, la transformation et le chargement des données contenues dans nombreux fichiers Excel vers une nouvelle base de données centralisée conçue à cet effet. Ce rapport présente la méthodologie de conversion ainsi que le choix des outils utilisés pour cette migration.

# APPLYING AN ETL APPROACH FOR THE INTEGRATION OF ALL CLINICAL DATA USED IN THE MOLECULAR GENETIC AND VISCOGLIOSI MUSCULOSKELETAL DISORDERS RESEARCH LABORATORY OF THE SAINTE-JUSTINE UNIVERSITY HOSPITAL

Mohamed Elyes BEN ABDELKRIM

## ABSTRACT

The exploitation and analysis of clinical data, in research hospital centers, is becoming a central point of contention in the overall research process. To improve this process, medical research laboratories are using more and more technology to achieve performance gains, and this according to their limited budgets. It is in this context that Dr. Alain Moreau, principal investigator at the Laboratory of Molecular Genetics and Viscogliosi Musculoskeletal Diseases of the Sainte-Justine hospital, asked Professor Alain April of the École de Technologie Supérieure (ÉTS) to design a centralized database containing various data used by the laboratory.

This report describes a second iteration of this project where the selection, extraction, transformation and loading of the data originating from many Excel spreadsheets is converted into a new centralized database. This report presents both the conversion methodology as well as the choice of the open source tools used for this migration.

## TABLE DES MATIÈRES

## LISTE DES TABLEAUX

Page

## LISTE DES FIGURES

## LISTE DES ABRÉVIATIONS, SIGLES ET ACRONYMES

| | |
|---|---|
| **ETL** | **Extract – Transform – Load** |
| **SQL** | **Structured Query Language** |
| **SSIS** | **SQL Server Integration Services** |
| **SGBD** | **Système de Gestion de Bases de données** |
| **MVC** | **Model View Controler** |
| **ORM** | **Object Relational Mapping** |
| **EF** | **Entity Framework** |
| **CHU** | **Centre Hospitalier Universitaire** |
| **HTML** | **Hyper Text Markup Language** |
| **DOM** | **Document Object Model** |

## INTRODUCTION

L'utilisation des technologies de l'information est un moyen incontournable afin d'optimiser le travail, et ce dans plusieurs domaines. Les technologies de l'information offrent la possibilité de créer des logiciels qui permettent de réduire les délais de travail et rendre plus fiables les opérations quotidiennes. C'est dans ce contexte que le laboratoire de recherche en musculosquelettique, de l'hôpital Sainte-Justine a demandé aux étudiants, de génie logiciel de l'École de technologie supérieure (ÉTS), d'étudier le traitement des données du laboratoire et de proposer une solution pour optimiser la gestion des données actuellement gérée à l'aide de nombreux fichiers Excel. L'utilisation de plusieurs fichiers Excel, au laboratoire, présente plusieurs défis (voir figure 1.1): la duplication des données, le manque d'information synthétique, le manque de validation des données (c.-à-d. la présence de données erronées et ayant différents formats), la difficulté à rechercher et trouver l'information, la possibilité de perte d'information (c.-à-d. ces fichiers sont sur les ordinateurs personnels de chaque employé), l'absence de traçabilité (c.-à-d. les journaux de modifications des données) et l'absence de droits d'accès précis.

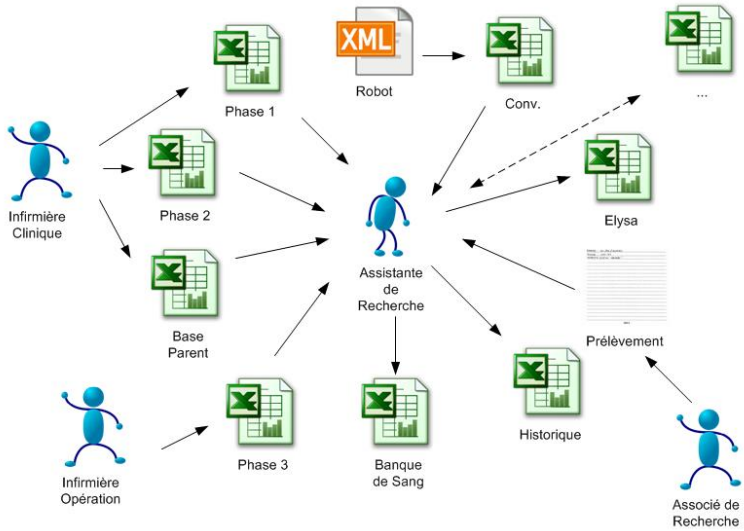



Figure 1.1 Situation actuelle de gestion des données au laboratoire [1]

La solution proposée par les étudiants de l'ÉTS, qui ont conçu un prototype de logiciel Web, a été réalisé à l'aide les technologies Microsoft ASP.net, Entity Framework, C# ainsi qu'une base de données centrale exploitant SQL serveur. Ce choix technologique est imposé par les responsables des technologies d'information de l'hôpital Sainte-Justine permettant ainsi une homogénéité des logiciels internes, l'accès à des licences, infrastructures et des services normalisés.

Ce projet de recherche appliquée, de 15 crédits, consiste à effectuer l'extraction, la translation et le chargement des données dans la base de données centralisée à partir des nombreux fichiers Excel. Il est donc nécessaire de normaliser, corriger, vérifier et intégrer les données vers cette base de données centralisée. Elle pourra ensuite être exploitée par le logiciel web. Pendant ce projet il est possible qu'il soit nécessaire d'effectuer des ajustements à la structure de données et à l'application Web suite aux essais de réception effectuée par Mme Anita Franco au laboratoire.

Afin de bien mener ce projet à terme, il est nécessaire de maîtriser les méthodologies de normalisation, d'extraction, de transformation et de chargement des données (c.-à-d. le domaine de l'ETL). Pour atteindre ce premier objectif, le premier chapitre de ce rapport est consacré à la revue littéraire des différentes approches d'extraction, de transformation et de chargement de données.

# CHAPITRE 1

# La revue de la littérature

## 1.1 Problématique

Il a été soulevé, dans l'introduction de ce rapport, que le laboratoire de recherche en maladies musculosquelettiques, situé à l'hôpital Sainte-Justine, utilise plusieurs fichiers Excel pour la gestion et la sauvegarde des données patients lors de ses activités de recherche clinique. La saisie de ces données est faite manuellement par différents intervenants du laboratoire ce qui cause plusieurs problématiques. De plus, les formats de ces différents fichiers ne sont pas standardisés ce qui rend la manipulation et la gestion de ces fichiers encore plus ardue. La première étape, afin d'améliorer la situation de ce laboratoire, est de centraliser les données dans une base de données centrale. Pour ce faire, il est intéressant d'explorer les différents processus d'extraction, de transformation et de chargement de donnée (connus sous le terme : ETL (Extract, Transform, Load)) proposés dans la littérature. Les prochaines sections de ce chapitre introduisent les concepts du processus d'ETL, son cycle de vie et les critères de sélection d'un outil d'ETL.

## 1.2 Processus ETL

Au tout début d'un processus ETL, il y a des données issues de plusieurs sources de données hétérogènes. Le processus ETL (présenté à la figure 1-2) se situe entre les différentes sources de données (à gauche de la figure) et l'entrepôt de donnée de l'organisation (au centre de la figure). Il vise à traiter les données de ces différents systèmes sources et les rendre homogènes. Homogène, signifie ici 'avoir le même modèle de données et même univers sémantique' [2]. Cela est possible en concevant des règles sous la forme de métadonnées, c'est-à-dire des informations sur les données (tel que présenté à la figure 1-3).

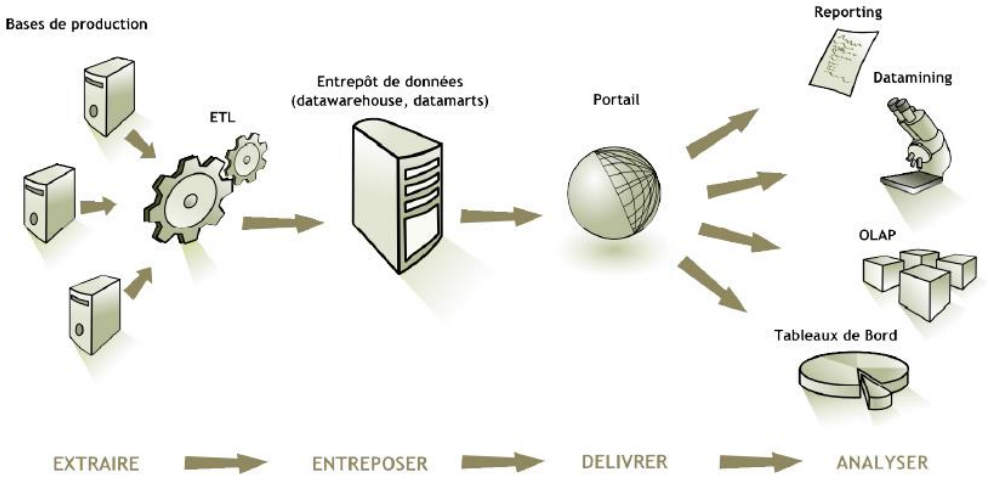



Figure 1.2 Identification de la localisation du processus ETL pour l'entrepôt de données [3]

Voici un exemple de règles appliquées aux données par un processus d'ETL:

- L'âge des participants, les participants pour les tests du laboratoire, doit être positive et correcte par rapport à la date de naissance : dans un des fichiers Excel sources (à traiter) nommées « Copie de Copie de # 3  PHASE 3  janvier 2014 » l'âge peut être négative à cause de la saisie manuelle, de l'utilisation erronée des formules Excel. Dans la figure 1.3, un exemple de ce fichier Excel où l'âge est négatif et incorrect.

| Famille | RANDOM | # DOSSIERS HSJ | SALIVE | DDN | SEXE | ÂGE |
|---|---|---|---|---|---|---|
| 725 | 3209 | Mère | | 17/03/1965 | f | -65,21 |

Figure 1.3 Anomalie de données dans un fichier (Âge négatif)

- Une autre règle, par exemple, est que les données qui ont la même signification (exemple : mâle et femelle) doivent avoir la même codification. Or dans plusieurs fichiers sources du laboratoire, cette règle n'est pas respectée (c.-à-d. mâle: m, 1, et femelle : f, 2) comme le montre la figure 1.4 ci-après.

| DDN | SEXE |
|---|---|
| 23/08/1999 | m |
| 26/03/2003 | f |
| 02/06/1965 | f |
| 01/11/1967 | f |
| 18/04/1998 | f |
| 21/05/1995 | 1 |
| 29/07/1996 | 2 |
| 12/07/1998 | 2 |
| 21/09/1967 | 1 |

Figure 1.4 Anomalie de données dans un fichier (valeur du sexe)

Un autre exemple de nécessité de mettre en place des règles sous la forme de métadonnées est présenté à la figure 1.5 : la même information sur le test du « Risser » ou bilan radiologique du bassin est représentée de différentes façons (c.-à-d. 0-1a la même signification que 0+).

| M |
|---|
| Risser |
| 0-1 |
| 0 + |
| 1 + |
| 1 + |
| 1-2 |
| 0 + |
| 0 + |
| 1 + |
| 0 + |
| 0 + |

Figure 1.5 Anomalie de données dans un fichier (valeur de risser)

Il existe plusieurs autres règles qui ont dû être établies pour le laboratoire. La définition de ces règles a nécessité plusieurs rencontres avec la spécialiste du domaine d'affaires afin de bien comprendre la signification de chaque donnée, son importance et sa relation avec les autres données.

Le Processus ETL est donc une étape incontournable dans les projets de mise en place d'un entrepôt de données qui pourra, par la suite être exploité. Effectuer la conversion des formats de données (c.-à-d. dans le format requis par l'entrepôt de données) constitue l'essentiel de la tâche de l'ETL. L'alimentation de la base de donnée centralisée, à l'aide d'un processus ETL, comporte quatre grandes phases (voir figure 1.6) : 1) la sélection des données; 2) l'extraction des données; 3) la transformation des données; et, 4) le chargement des données. Voici une explication détaillée de chacune de ces quatre phases :

1. La sélection des données : cette phase consiste à séparer, des informations qui sont nécessaires et importantes pour le client, les informations qui n'ont pas d'importance pour client et qui ne seront pas utilisées. Lors de cette phase, il est aussi nécessaire d'identifier les fichiers sources, les collecter et préparer les données pour leur extraction. Cette première phase est très importante et représente environ 75% de la durée du projet de mise en place d'un ETL [2].

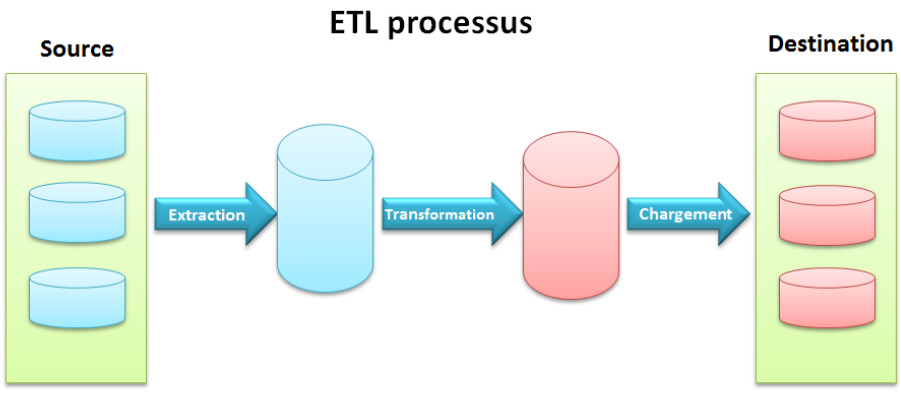



Figure 1.6 Processus ETL [4]

2. L’extraction des données : l’extraction des données est effectuée à l’aide d’un outil d’alimentation des données [5] qui interagit (c.-à-d. interface et est compatible) avec la technologie de la base de données centralisée voulue. Dans cette deuxième phase, il est nécessaire de traiter les anomalies dans les données qui bloquent l’extraction (c.-à-d. qui sont rejetées et doivent être traitées manuellement). Lors de cette phase, il est important de conserver une trace des modifications effectuées et de valider ces ajustements avec le propriétaire des données;
3. La transformation des données : le but de cette troisième phase et de rendre les données homogènes et cohérentes. Par exemple il est nécessaire de filtrer les données (c.-à-d. augmenter leur qualité), traiter les données sans valeurs ou avec une valeur manquante. Ces nombreuses transformations nécessitent une bonne compréhension des données (c.-à-d. si on doit séparer l’information, s’il y a des synonymes, si ces informations sont reliées …). La compréhension de la sémantique de ces données passe généralement par des rencontres avec le client et les spécialistes du domaine d’affaires;
4. Le chargement des données : cette dernière phase du processus ETL consiste à charger les données dans la base de données cible. Cette phase peut-être délicate surtout lorsque le volume de données est important.

Chaque phase du processus ETL contient des sous-phases. L’application des phases et des sous-phases forment le cycle de vie d’un ETL.

### 1.3 Cycle de vie d'un ETL

Afin de mieux comprendre les activités détaillées effectuées lors du processus ETL, le cycle de vie d'un processus ETL [6] est décrit plus en détail. Il comporte 8 activités :

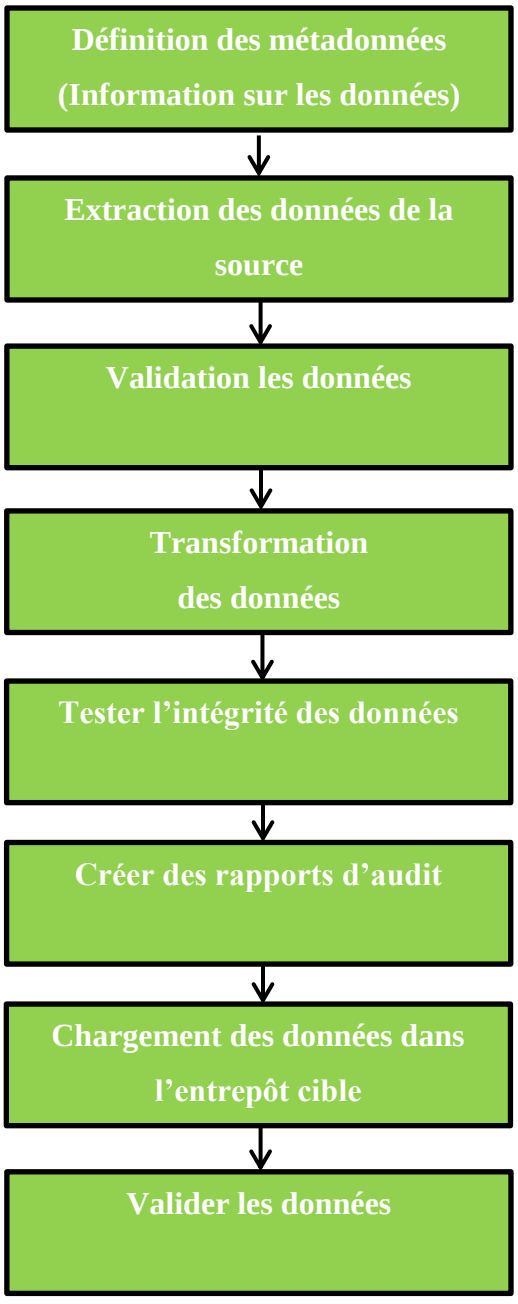



Figure 1.7 Le cycle de vie d'un processus ETL

Activité 1 : la définition des métadonnées : c'est la définition du format et de la nature des données. Il y a deux types de données :

- les données générales (exemple le nom, prénom, date de naissance, Numéro de téléphone…).
- Les données spécifiques du domaine d'affaire, par exemple dans notre cas, les types des COOB (les angles des os calculés par les médecins) sont : thoracique, lombaire, Thoracolombaire et cervical, leurs représentations successives est T, L, TL et C. la définition des métadonnées des données spécifiques nécessite des rencontres avec un spécialiste du domaine d'affaires.

Activité 2 : l'extraction des données de la source : c'est l'extraction des données sources avec un outil ETL. Il faut définir dans cette phase les différentes sources de données à traiter ainsi que les données à extraire dans chaque source.

Activité 3 : la validation des données : une fois les données sont extraient dans l'outil ETL et avant de commencer la transformation des données, il faut vérifier l'intégrité des données.

Activité 4 : Transformation des données: c'est la transformation, la normalisation et la correction des données.

Activité 5 : le test de l'intégrité des données: après la transformation des données, et avant de charger les données dans la base de données cible, il faut revérifier l'intégrité des données.

Activité 6 : la création des rapports d'audit: il faut créer des rapports d'audit afin d'assurer la traçabilité des modifications.

Activité 7 : le chargement des données dans l'entrepôt cible: c'est l'avant-dernière étape du processus ETL qui consiste à charger les données transformées.

Activité 8 : Validation les données :

Cette étape constitue à effectuer des tests sur les données chargées dans la base de données cible afin de valider l'intégrité des données.

Maintenant que nous connaissons les activités effectuées lors d'un projet d'ETL, la prochaine section traite du choix d'outils pour appuyer ces activités.

### 1.4 Choix d'un outil d'ETL

Il existe trois approches concernant l'utilisation d'outils d'aide à l'ETL qui peuvent être considérées lors de la planification d'un projet d'ETL. La première option, concernant les outils à utiliser, est de concevoir et programmer des logiciels soi-même (c.-à-d. des scripts de conversion) spécifiques pour le projet. La deuxième option est de choisir un/des outils, par exemple des ETL du domaine du logiciel libre, pour appuyer l'activité d'ETL. Finalement, la troisième option est d'acheter un outil d'ETL propriétaire. Bien sûr, chacune de ces approches présente des avantages et des inconvénients.

1. Les solutions spécifiques d'ETL (concevoir et programmer soi-même) :
   Cette approche présente certains avantages, car il est possible de programmer une solution qui répond à toutes les exigences de l'entreprise, une solution personnalisée et qui respecte la spécificité de l'entreprise. Mais, à long terme, cette solution est couteuse, car plus le projet devient grand et complexe plus c'est difficile de gérer l'application. En plus, l'instabilité des ressources humaines (c.-à-d. les concepteurs et les développeurs de l'application) diminue la robustesse, la disponibilité et surtout complique la maintenance de la solution.
2. Les solutions ETL propriétaires :
   Le grand avantage de cette approche est une mise en service souvent plus rapide de la solution. L'inconvénient est le cout plus élevé (c.-à-d. achat de licences, formation...).
   Voici quelques exemples de solutions propriétaire qui existent sur le marché :

- **SQL Server Integration Services (SSIS) :** est un produit de Microsoft SQL server, une plateforme d'intégration des données. C'est un outil performant et flexible pour l'extraction, la transformation et le chargement des données. Il permet aussi la mise à jour et la maintenance des bases de données SQL server [7].

- **IBM InfoSphere DataStage :** est un outil ETL qui fait partie de la plateforme IBM. Cet outil utilise des interfaces graphiques pour construire la solution d'intégration des données. Il y a plusieurs versions d'IBM InfoSphere : édition Serveur et édition entreprise [8].

- **Oracle Data Integrator (ODI) :** cet outil offre un environnement graphique pour intégrer les données afin de préparer un processus d'intelligence d'affaires [9].

- **SAS Data Integration Studio :** est un logiciel performant qui permet d'intégrer les données provenant de différentes bases de données, d'application ou de plateforme vers un entrepôt de données cibles. Le SAS Data Integration Studio est un environnement qui permet la collaboration de plusieurs utilisateurs et facilite le partage du travail à réaliser [10].

3. Les solutions ETL du domaine du logiciel libre :

Cette troisième approche utilise des solutions ETL disponibles en logiciel libre. Ces outils offrent une bibliothèque de fonctionnalités qui peuvent être utilisées afin de créer une solution personnalisée d'ETL gratuite. Voici quelques exemples de solutions libres disponibles :

- **Talend Open Source Data Integrator :** ce logiciel libre offre plusieurs solutions pour l'intégration des données. Il y a une version libre et une version commerciale. Talend possède plus de 400 connecteurs pour la connexion entre

plusieurs types d'entrepôts de donnée. La solution offre aussi des services Web et des applications pour le service Cloud [11].

- **Scriptella** : Ce logiciel libre est un script Java. Cette approche est idéale si une base de données SQL est utilisée comme source de données [12].

- **KETL** : un outil ETL libre en langage Java qui possède un module de configuration en XML, il offre plusieurs fonctionnalités qui peuvent rivaliser avec des solutions propriétaire. KETL supporte l'intégration, la sécurité et la gestion des données [13].

- **Pentaho Data Integrator :** possède un module ETL en langage Java, il permet la conversion et la transformation des données à travers de multiples connecteurs et des filtres. Il y a une possibilité d'ajouter à la version g, en logiciel libre, des modules payants comme le module « Agile BI » qui permet la visualisation des transformations étape par étape [14].

- **Jaspersoft** : Jaspersoft est un outil libre d'ETL qui permet la création de tableaux de bord et d'analyser les données transformées afin de présenter les résultats sous forme de rapport [15].

Le choix d'une des trois approches (c.-à-d. code spécifique, logiciel ETL propriétaire, ou logiciel ETL libre) a un impact sur le cout et la durée du projet. Un ETL propriétaire, bien qu'il a un cout élevé dès le début du projet (c.-à-d. achat de licences, formation) va être plus productif. L'approche d'ETL par le développement de code spécifique prendra aussi un investissement de départ. Enfin le choix d'un outil ETL en logiciel libre diminue initialement le cout d'acquisition, mais nécessitera l'adaptation spécifique. Les développeurs de l'ÉTL Pentaho présentent une courbe qui décrit cette relation à la figure 1.8 ci-dessous.

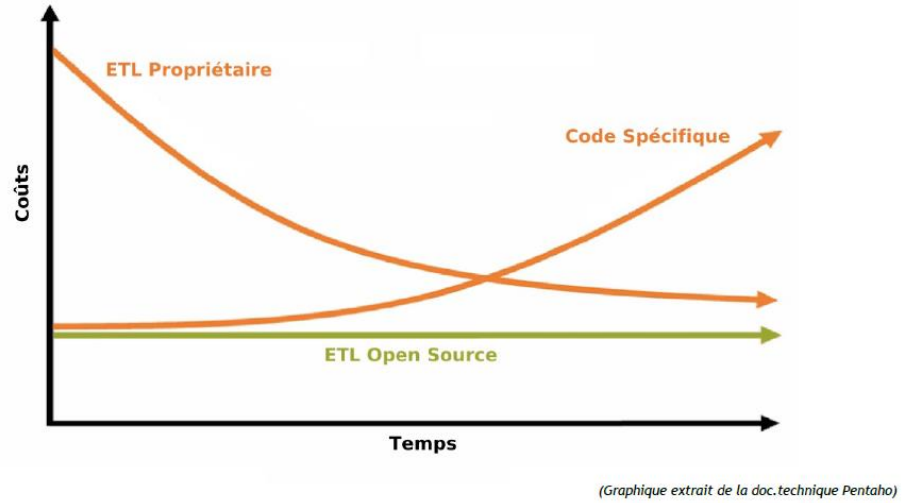



Figure 1.8 Évolution du cout dans le temps des différentes approches ETL [16]

## 1.5 Comparaison entre Talend Open Source Data Integrator et SQL Server Integration Services (SSIS)

Le développement d'une solution spécifique d'ETL (c.-à-d. concevoir et programmer soi-même) est un choix qui peut prendre un effort considérable, pour un projet de 15 crédits. Particulièrement dans ce cas-ci, la taille, la complexité et la spécificité du projet n'exigent pas nécessairement le développement d'une solution spécifique. Afin de finaliser le choix de l'outil ETL, une comparaison entre deux produits très présents dans le marché est effectuée. Le but de cette comparaison n'est pas de chercher la solution ETL la plus performante, car en général un produit propriétaire (c.-à-d. payant) offre plus de fonctionnalités et d'avantages qu'un produit du domaine du logiciel libre. Le but est donc d'identifier une solution la adaptée à ce projet qui n'a aucun financement, car le chercheur a perdu toutes ces subventions cette année.

Le SQL Server Integration Services (SSIS), qui est une solution propriétaire de Microsoft, et le Talend Open Source Data Integrator, qui est une solution du domaine du logiciel libre, sont parmi les meilleures solutions ETL disponibles, selon plusieurs rapports [17]. Les critères choisis pour la comparaison des deux produits sont : les critères de support et de

documentation (voir tableau 1.1) et les critères d'implémentation (voir tableau 1.2). Le tableau 1.1 précise que l'outil Talend n'offre pas de support pour les missions critiques contrairement à l'outil SSIS. Ce projet ETL est une itération de préparation de la base de données pour la prochaine itération qui est le développement d'outils analytiques sur ces données. L'évolution future du projet ainsi que la création de demandes de changements probables, de la part de la cliente, peuvent nécessiter un support pour des fonctionnalités critiques, d'où l'importance d'avoir ce service dès le début.

Tableau 1.1 Comparaison des critères de support et documentation [18]

| **Support et documentation** | **Talend Open Source** | **SSIS** |
|---|---|---|
| Communauté d'aide : forum, « bug-tracker » | oui | oui |
| Support pour entreprise avec SLA (Service Level Agreement) | non | oui |
| Support pour mission critique | non | oui |

L'utilisation du module SSIS est facile et rapide dans un environnement de Microsoft [18]. Par contre, tel qu'illustré au tableau 1.2 (parallélisation ), la version libre de Talend présente un inconvénient : c'est que l'outil est configuré pour un seul utilisateur à la fois, ce n'est pas possible d'avoir plus qu'un utilisateur à la fois, et un seul utilisateur par système ce qui pose des problèmes si l'implémentation nécessite plusieurs utilisateurs au même moment ou lorsqu'un utilisateur oublie sans mot de passe .

Tableau 1.2 Comparaison des critères d'implémentation [18]

| **Implémentation** | **Talend Open Source** | **SSIS** |
|---|---|---|
| Versionnage | oui | oui |
| Parallélisation | non | oui |
| Visualiseur de données | non | oui |
| Schéma dynamique | non | oui |

## 1.6 Conclusion

Ce premier chapitre a permis de mieux comprendre le processus ETL et d'étudier et de comparer les différentes approches possibles à adopter pour ce projet.

Le prochain chapitre présente une synthèse des activités techniques concernant la migration des données du laboratoire de recherche en musculosquelettique, de l'hôpital Sainte-Justine en conjonction avec l'utilisation de l'outil SQL Server Integration Services (SSIS) pour l'intégration des données.

# CHAPITRE 2

# Application du processus ETL

## 2.1 Introduction

Ce chapitre présente l'analyse des systèmes sources, la définition de la portée du projet et les différentes étapes du processus ETL appliquées aux bases de données source en utilisant l'outil SQL Server Integration Services (SSIS). Ce chapitre traite aussi les adaptations apportées à l'application web du projet.

## 2.2 Analyse des systèmes sources

Nous avons précisé que les données sources, sous forme de fichiers Excel, proviennent de plusieurs endroits :

- Base de la phase1, qui contient les données des cas les plus sévères de la maladie;
- Base de la phase 2, qui contient les données des cas modérés de la maladie;
- Base de la Phase 3, qui contient les données des sujets asymptomatiques avec antécédents familiaux;
- Base des contrôles « écoles », qui contient les données des contrôles effectués dans des écoles primaires;
- Base des contrôles « clinique pédiatrique », qui contient les données des contrôles effectués à l'hôpital;
- Base des Traumas, qui contient les données des participants ayant un trauma;
- Base des parents, qui contient les données des parents des participants.

L'étude du système existant a démontré que la saisie des données, dans chaque fichier Excel, est effectuée par une ou plusieurs infirmières et que ces fichiers ne sont pas normalisés .Il n'y a aucun contrôle de la validité de la codification des données et aucun journal des modifications effectuées sur la base de données.

L'exploitation des données est effectuée par l'assistante de recherche, qui vérifie manuellement l'intégrité et la validité des données récoltées lors de l'utilisation. Ce travail de vérification manuelle fait perdre, à l'assistante de recherche, un temps précieux qu'elle pourrait utiliser à des activités de recherche et à de l'analyse de données. De plus, le risque d'erreur en analysant et modifiant ces fichiers manuellement (c.-à-d. par l'utilisation de fonctions Excel) est présent. La figure suivante (figure 2.1) présente un diagramme de la saisie, traitement et analyse des données.

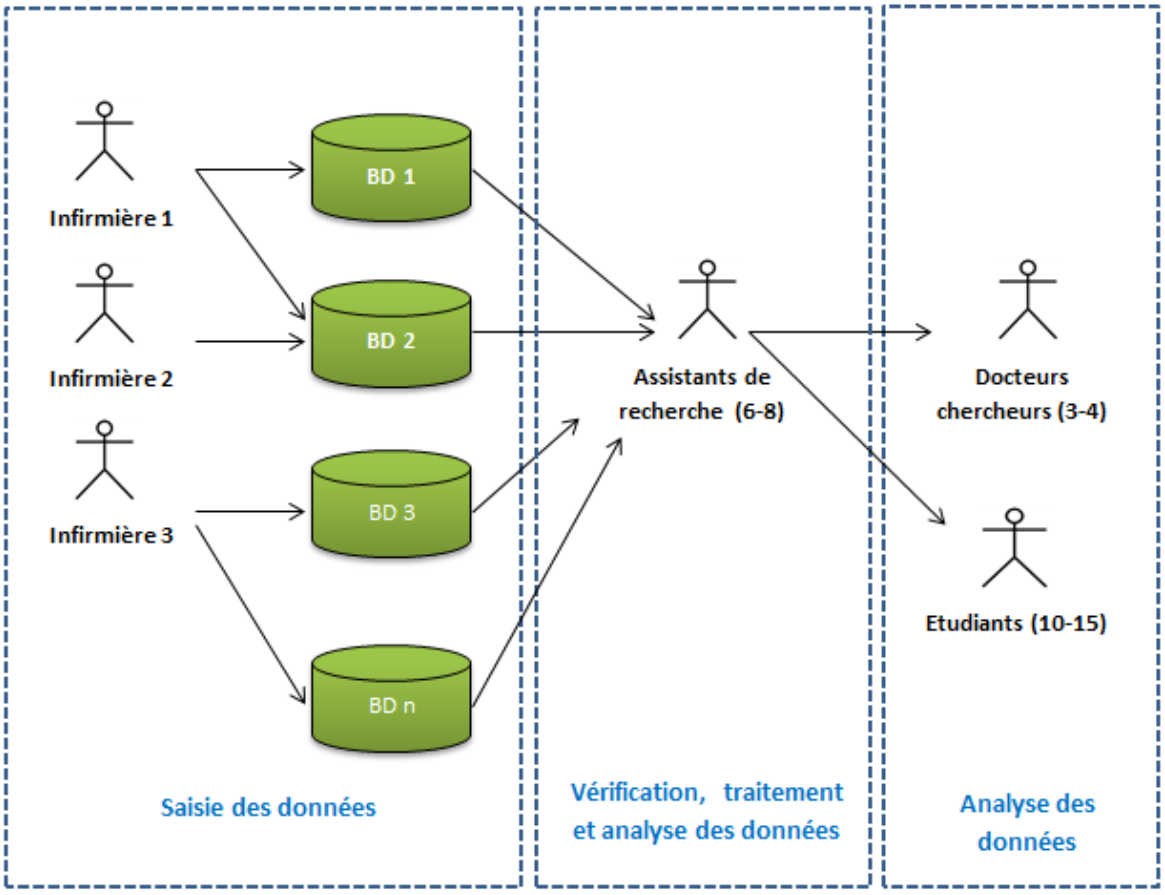



Figure 2. 1 Diagramme de la saisie, traitement et analyse des données

## 2.3 Définition de la portée des phases de projet

Nous avons déjà précisé que la portée du projet consiste à la sélection, l'extraction, la transformation et le chargement des données dans une base de données unique, homogène, centralisée et normalisée. Il faut aussi s'assurer qu'elle est compatible avec le prototype d'application Web qui va interagit (c.-à-d. ajouter, modifier et détruire des données) avec la base de données. La figure suivante présente le nouveau mode de saisie, traitement et analyse des données du nouveau système.

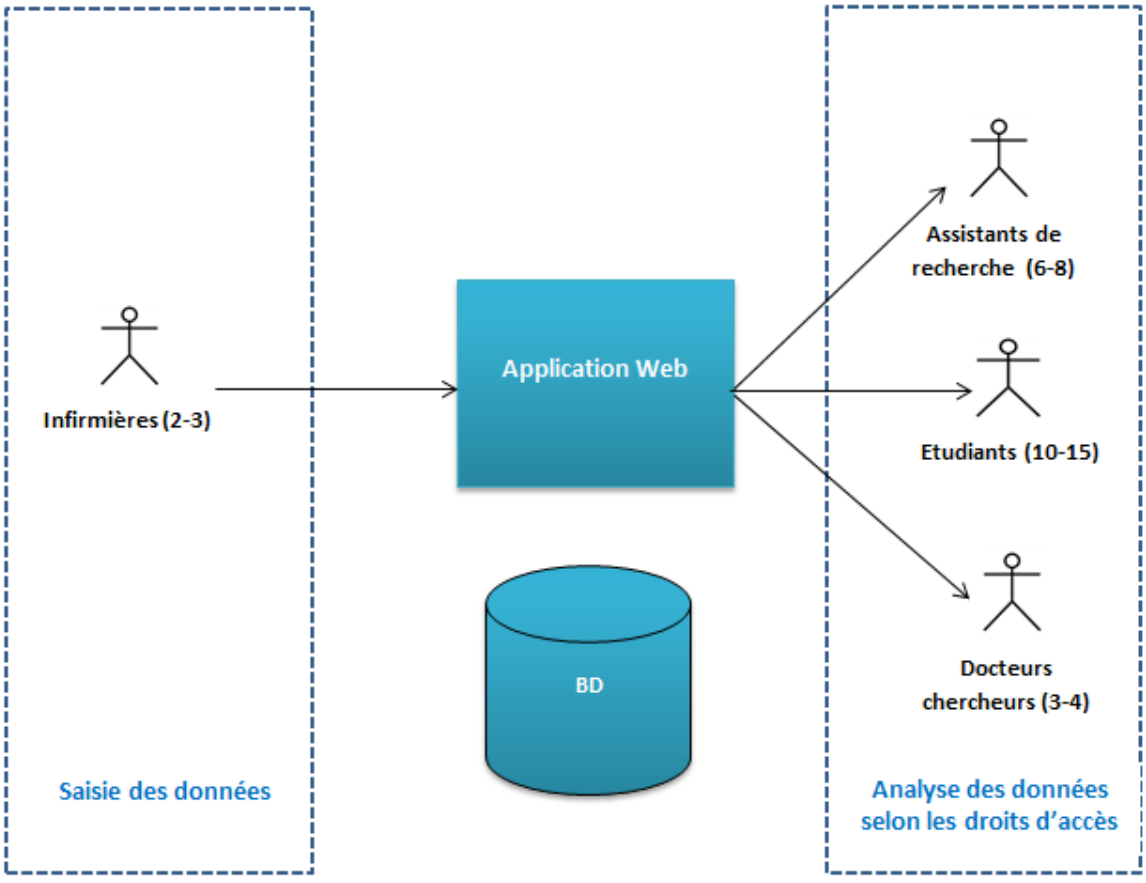



Figure 2. 2 Nouveau diagramme de la saisie, traitement et analyse des données

## 2.4 Sélection des données

La sélection des données est une étape délicate et importante dans le processus ETL. Il s'agit de choisir parmi toutes les sources de données celles qui seront utiles et utilisées par l'application. La sélection des données a nécessité plusieurs rencontres et échanges de courriels avec Mme Anita Franco, qui est assistante de recherche sénior dans le laboratoire et coordinatrice du projet. Aussi à cette étape il faut définir comment les données seront regroupées et dans quelle fenêtre de l'application les nouveaux champs seront intégrés. Encore, comme l'application peut être soit en français soit en anglais, et que dans les documents sources les champs des données sont soit en anglais ou en français il faut définir ,avec la coordinatrice du projet, pour chaque champ son équivalent dans l'autre langue.

## 2.5 Définition des métadonnées

Les métadonnées sont les informations sur les données (type, format …). La bonne définition des métadonnées est importante pour assure la réussite de l'application du processus ETL. Dans le contexte de notre projet, beaucoup de données sont des références médicales (type de maladies, type d'analyse, groupe ethnique, anatomie…) et une même information peut avoir plusieurs présentations. Dans le tableau en dessous (tableau 2.1) un exemple de métadonnées du champ groupe ethnique.

Tableau 2. 1 Exemple de métadonnée du champ groupe ethnique

| Groupe ethnique | Représentation correcte des données dans la base de données source |
|---|---|
| Asiatique | A |
| Haïtien | N |

| | |
|---|---|
| Indien | H |
| latino | L |
| Arabe | R |
| Italien | I |
| Turque | T |
| Arabe/ Italien | R/T |
| Italien/ Turque | I/T |
| Latino/ Italien | L/I |

D'après les spécifications de la coordinatrice du projet, le champ groupe ethnique contient une seule lettre, présenté au tableau 2.1 ci-dessus ou deux lettres différentes, qui existent dans le tableau, séparé par « / ».

Les spécifications des métadonnées, définies au début du projet, permettent de créer les critères de correction et validation des données lors de l'étape de transformation. Il faut noter que les spécifications de chaque champ peuvent être modifiées ou complétées au cours du processus ETL, puisqu'il n'y a pas une spécification pour l'insertion de données (c.-à-d. chaque infirmière, utilise ses propres spécifications actuellement).

## 2.6 Extraction des données de la source

Les fichiers sources du laboratoire sont des fichiers Excel et la base de données de l'application future et une base de données SQL server. Donc la première étape est de créer la liaison entre la source et la destination (figure 2.3).

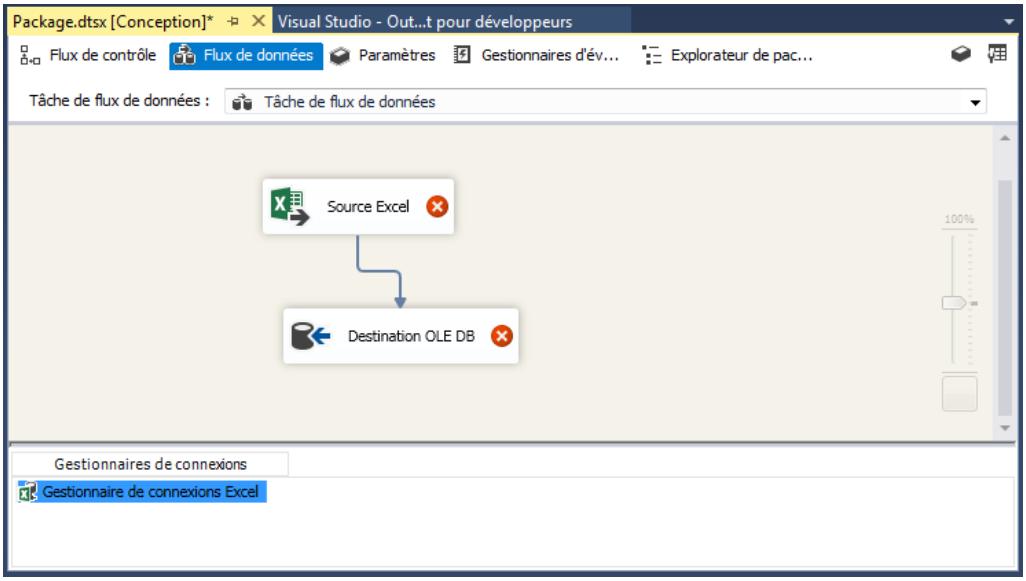



Figure 2. 3 Création de  la liaison entre la source et la cible

Une fois cette liaison créée, la première étape dans le processus d'extraction est de configurer la source, qui provient de fichiers Excel : c.-à-d.  choisir les fichiers Excel et les colonnes à utiliser dans le processus ETL (voir figure 2.4).

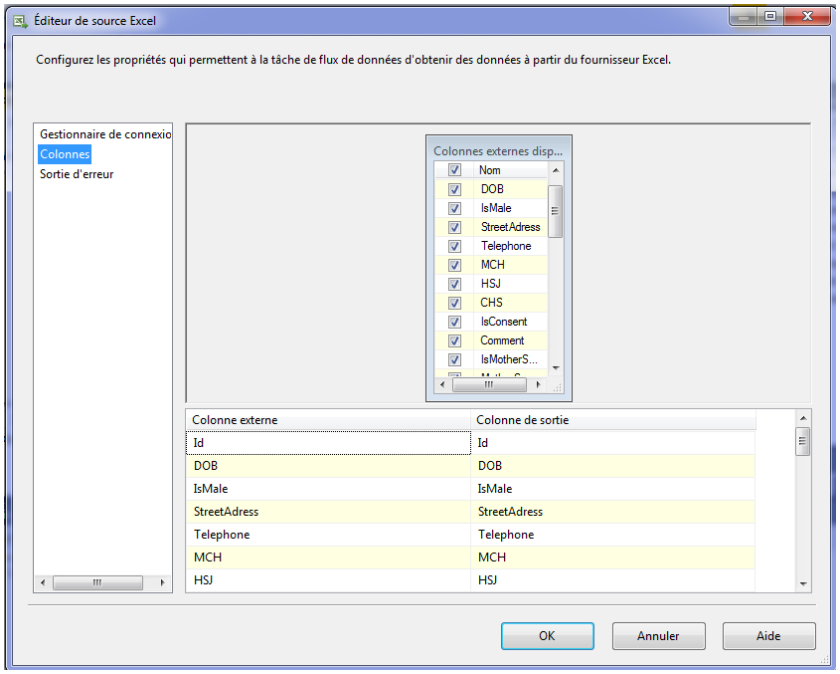



Figure 2. 4 Paramétrage des fichiers source (Excel)

La deuxième étape vise à configurer la destination et choisir la base de données SQL (voir figure 2.5).

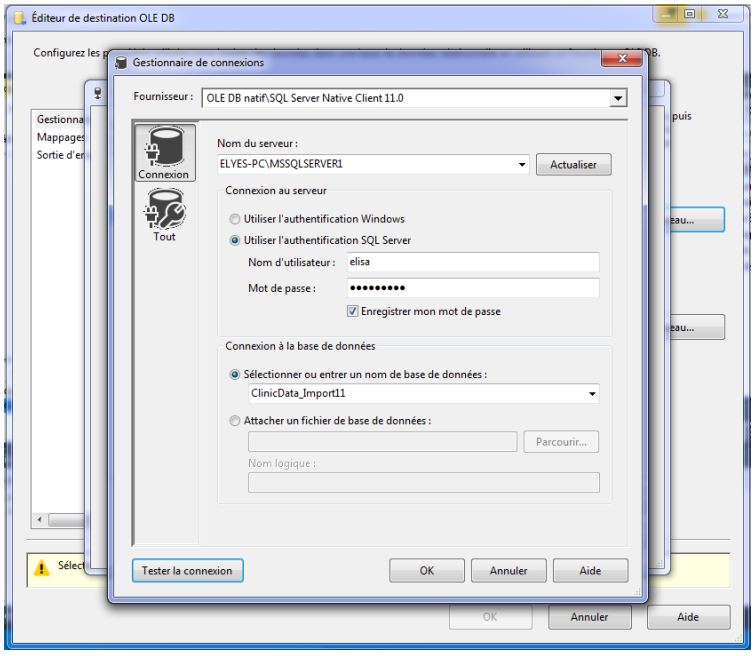



Figure 2. 5 Configuration de la base de données cible

Une fois la cible et la source de l'extraction configurées, il faut tester la connexion entre elles (voir la figure 2.6).

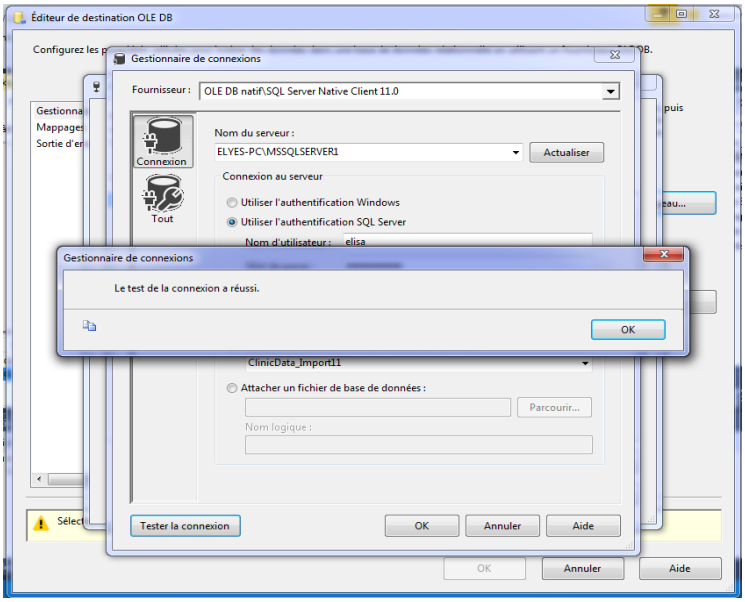



Figure 2. 6 Test de la connexion entre la base de données source et la base de données cible

La troisième étape consiste en la création de la table de données, dans la base de données cible, et à la définition du type de données de chaque colonne (voir figure 2.7).

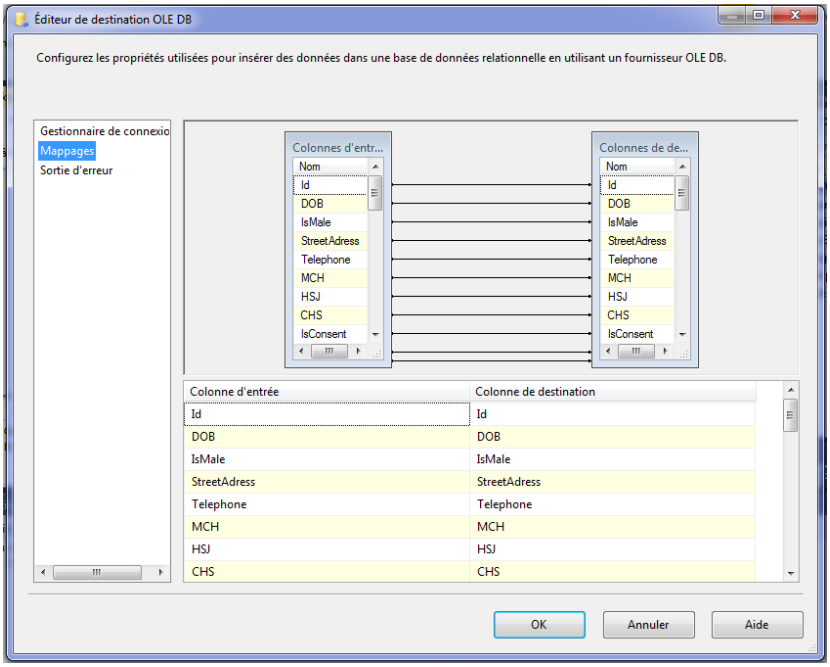



Figure 2. 7 Création des tables et définition du type de données

Il faut ensuite tester que la totalité des données sont transférées de la cible vers la destination.

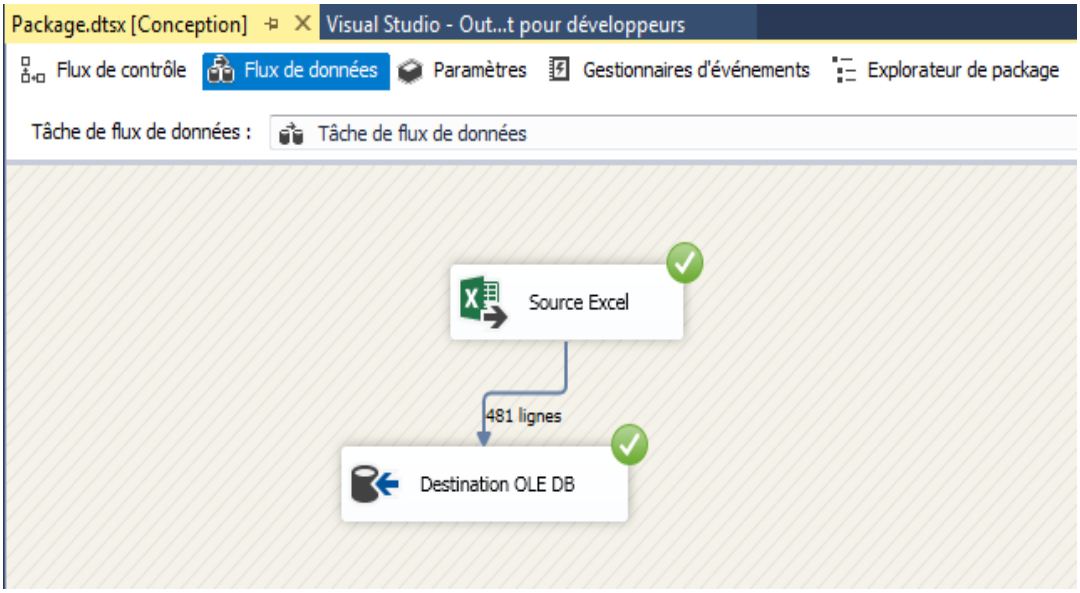



Figure 2. 8 Test de l'extraction des données

La quatrième étape vise à vérifier la création de la table dans la base de données SQL (voir la figure 2.9).

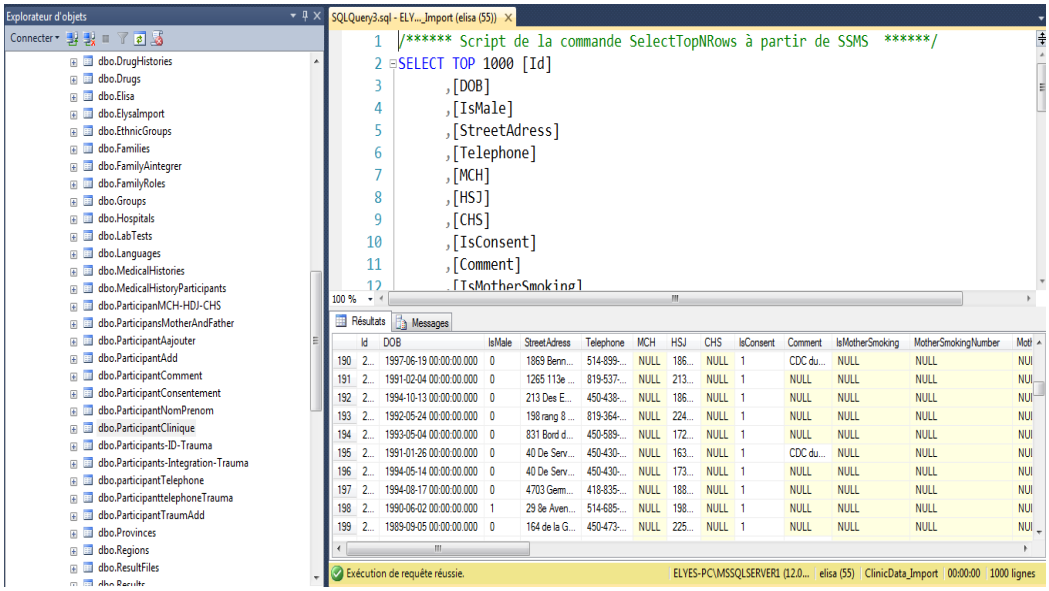



Figure 2. 9 Vérification de la création des tables dans la base de données cibles

## 2.7 Valider les données

Une fois les données transférées, une vérification du format de la table de données, des types de données et de la complétude des données transférées est nécessaire. La conception de différentes requêtes SQL permet d'effectuer la validation de l'extraction des données.

## 2.8 Transformer les données

La transformation et le nettoyage des données est l'étape qui a nécessité le plus d'effort dans ce projet. Le processus de transformation a nécessité plusieurs réunions avec la cliente afin de définir progressivement les critères de transformation (figure 2.10).

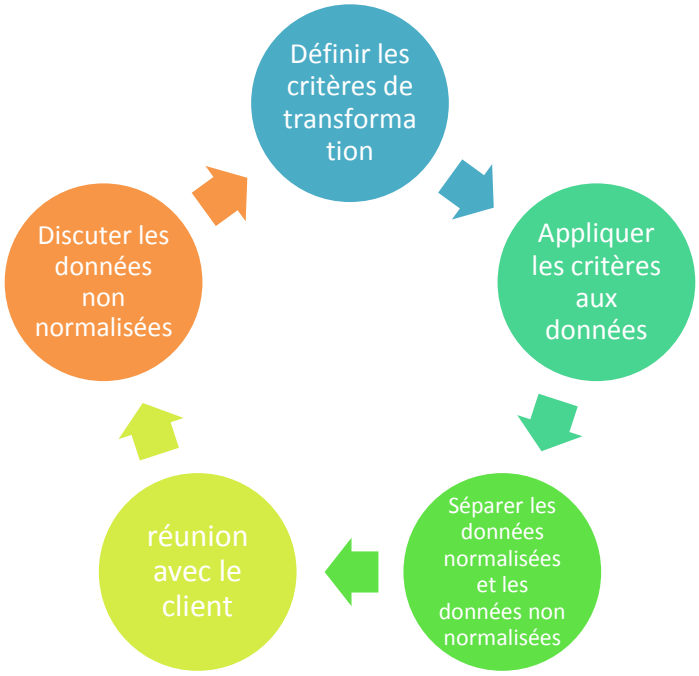



Figure 2.10 Processus des transformations

Parmi les transformations les plus importantes pour cette étape, on y retrouve:

Transformation de la date : Les représentations de la date, dans les bases de données sources, sont variées, voir le tableau 2.2.

Tableau 2. 2 Exemples de représentation des bases de données sources

| Exemples | Représentation |
|---|---|
| 1 | 20/10/1987 |
| 2 | 98/05/06 |
| 3 | 97-11-14 |
| 4 | 11 février 2015 |

De plus, il existe des représentations qui peuvent tromper les chercheurs, par exemple « 01/06/11 » qui peut être soit le premier juin 2011 soit le 11 juin 2001 (c.-à-d. ces deux représentations existent dans les fichiers Excels source).

Pour régler ces cas ambigus, l'utilisation des outils de transformation, dans notre cas le SSIS, permettra la normalisation de la représentation des dates. La transformation des dates a nécessité plusieurs itérations. Pour chaque itération (c.-à-d. cas de figure) il a été nécessaire d'appliquer des règles de transformations vérifiées avec la cliente. À la fin de chaque itération, il y a des cas résolus et des cas restants. Les cas restants ont nécessité des règles de transformation spécifiques. Ces données problématiques ont nécessité une nouvelle itération, avec de nouvelles règles de transformation. Le processus de transformation consiste à répéter ces itérations et à modifier les règles de transformation jusqu'à ce que toutes les données soient normalisées.

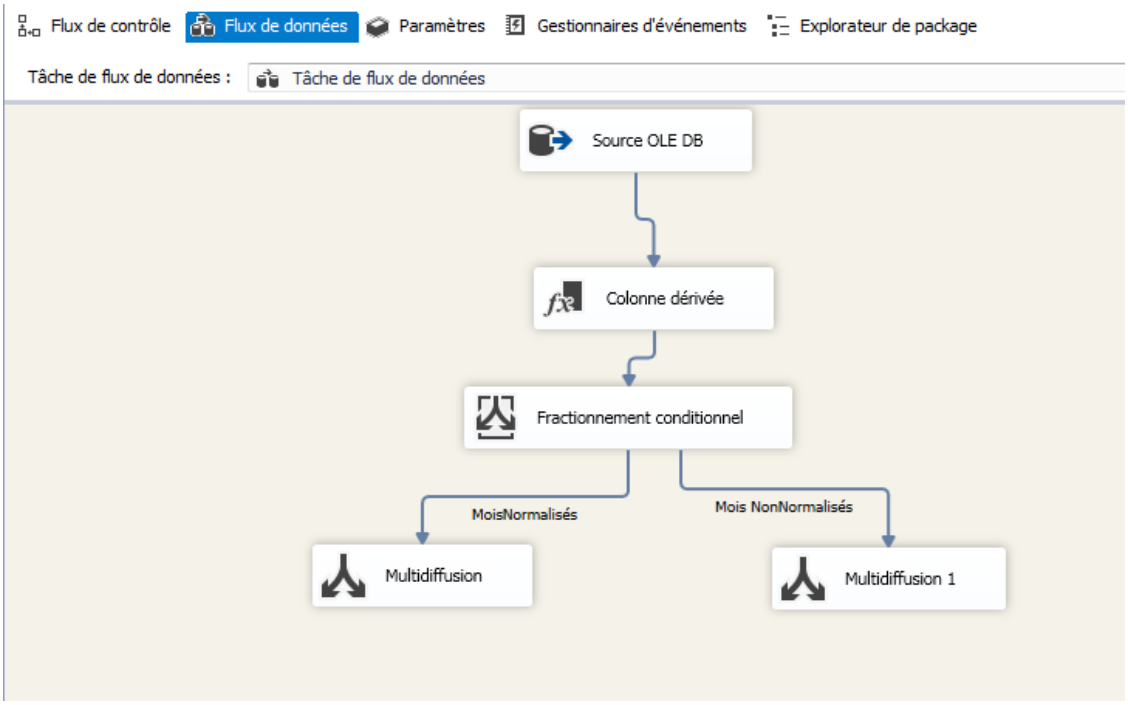



Figure 2. 11 Transformation et correction des dates par itérations

Pivoter les colonnes :

Les données dans les fichiers sources Excels sont stockées horizontalement sur des lignes. Afin de les transférer dans des tables (d'une base de données relationnelle), ces données doivent êtres transformées dans des colonnes, d'où l'utilité du la fonctionnalité du SSIS qui permet de pivoter les lignes en colonnes (voir figure 2.12).

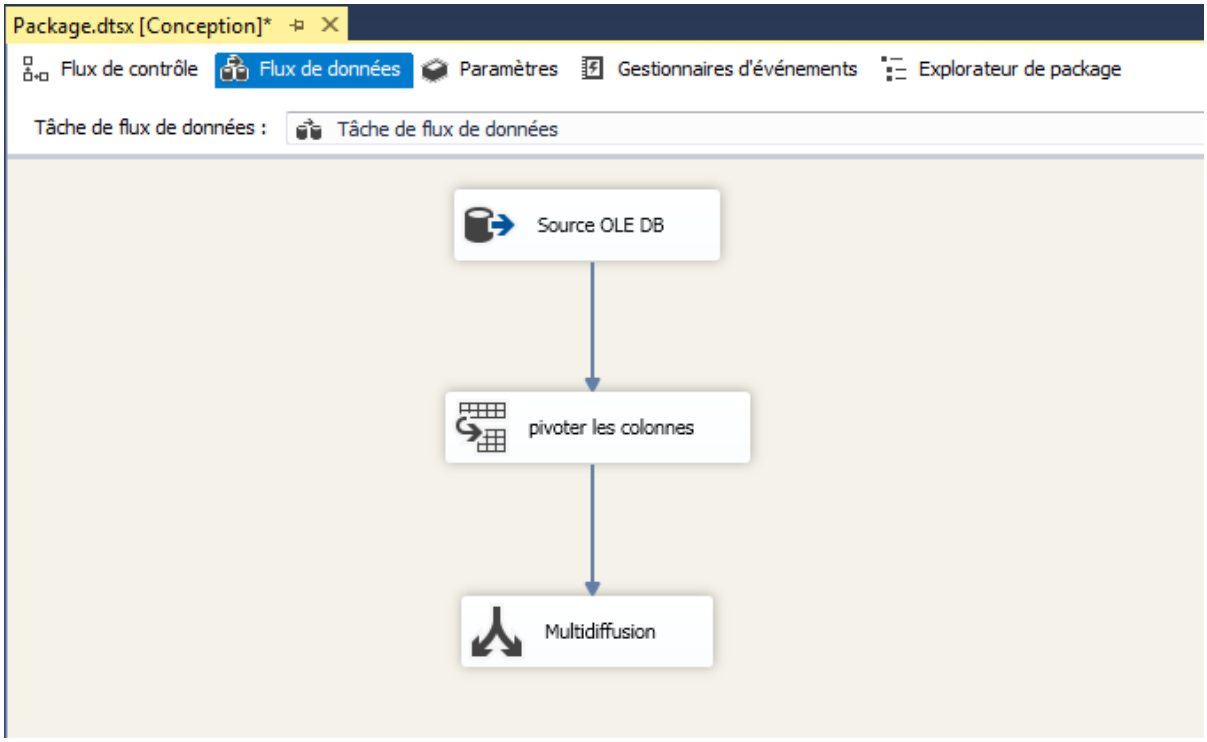



Figure 2. 12 Transformation des lignes en colonnes

## 2.9 Charger les données dans l'entrepôt cible

Une fois les données traitées, transformées et normalisées, elles sont chargées dans la base de données cible (c.-à-d. la figure 2.13 de Microsoft SQL server).

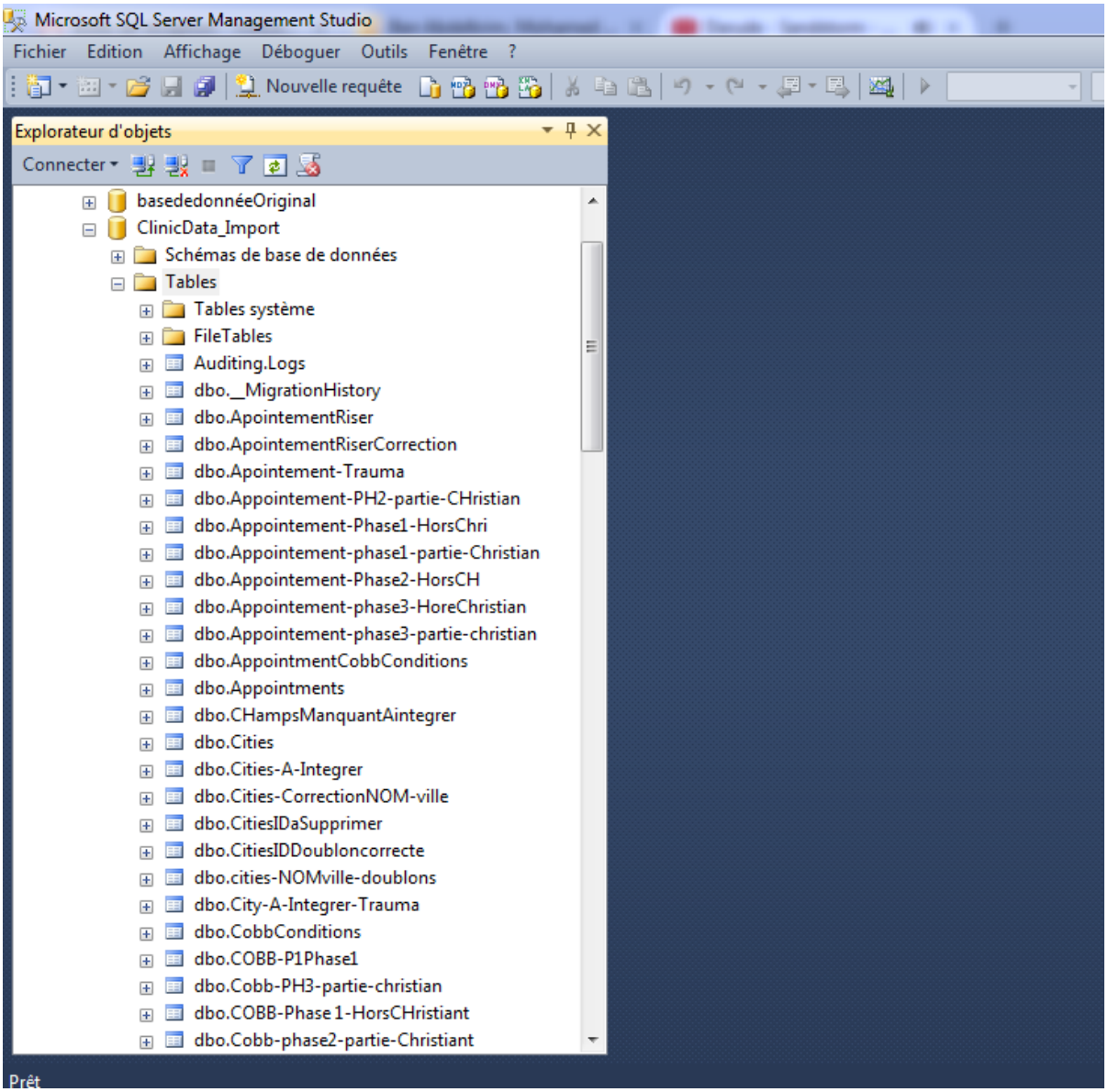



Figure 2. 13 Tables dans la Base de données

## 2.10 Adaptation de l'application Web

La liaison entre cette base de données SQL Server et l'application Web est assurée par l'utilisation du « Entity Framework », qui est un ORM (Object Relational Mapping), et qui permettra de travailler avec un niveau d'abstraction qui faciliter le développement d'une application orientée donnée.

La première étape de cette liaison consiste à créer la relation (c.-à-d. un mapping) entre les colonnes des tables de la base de données et les objets utilisés par l'application (figure 2.14).

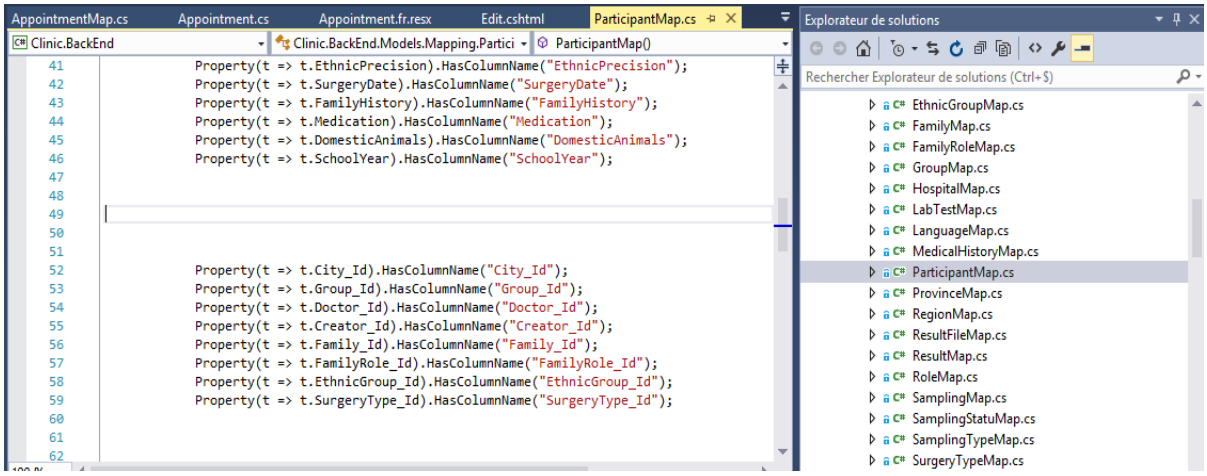



Figure 2. 14 Création du mapping entre les colonnes de la table Participant et les objets de l'application

Comme l'application est bilingue (c.-à-d. en français et en anglais), il est nécessaire de créer les fichiers ressources de contenant le nom de l'objet, à afficher dans l'application, dans ces deux langues.

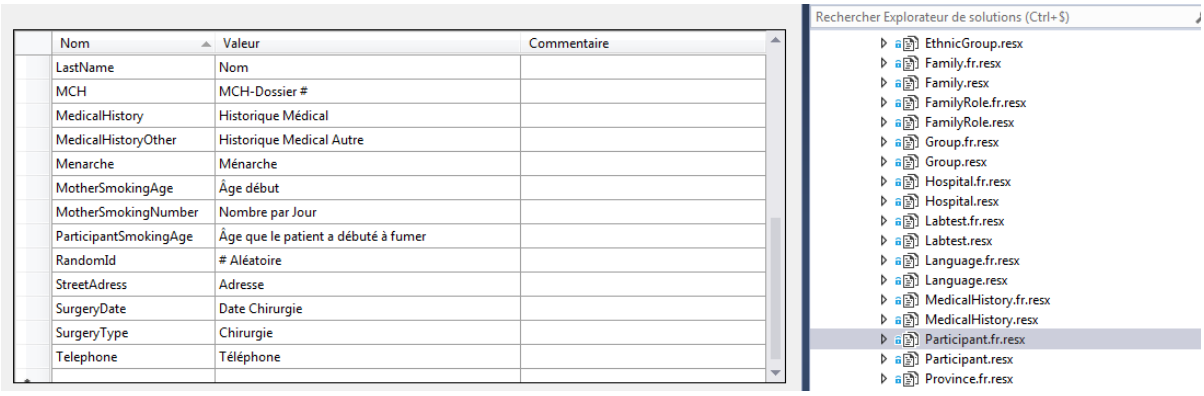



Figure 2. 15 Ressources pour la version française et anglaise de l'application

La dernière étape de la construction de l'application Web est l'utilisation du langage HTLM (figure 2.16) et CSS afin de fixer l'emplacement des champs dans les fenêtres de l'application Web, le style, la couleur…

```
        @Html.LabelFor(model => model.FamilyRole.Name, new { @class = "control-label col-md-4" })
        <div class="col-md-8">
            @Html.DropDownList("FamilyRole_Id", String.Empty)
            @Html.ValidationMessageFor(model => model.FamilyRole_Id)
        </div>
    </div>

    <div class="form-group">
        @Html.LabelFor(model => model.Doctor_Id, new { @class = "control-label col-md-4" })
        <div class="col-md-8">
            @Html.DropDownList("Doctor_Id", String.Empty)
            @Html.ValidationMessageFor(model => model.Doctor_Id)
        </div>
    </div>

    <div class="form-group">
        @Html.LabelFor(model => model.FamilyHistory, new { @class = "control-label col-md-4" })
        <div class="col-md-8">
            @Html.EditorFor(model => model.FamilyHistory)

            @Html.ValidationMessageFor(model => model.FamilyHistory)
        </div>
    </div>

    <div class="form-group">
        @Html.LabelFor(model => model.Medication, new { @class = "control-label col-md-4" })
        <div class="col-md-8">
            @Html.EditorFor(model => model.Medication)
            @Html.ValidationMessageFor(model => model.Medication)
        </div>
    </div>
    <div class="form-group">
        @Html.LabelFor(model => model.DomesticAnimals, new { @class = "control-label col-md-4" })
        <div class="col-md-8">
            @Html.EditorFor(model => model.DomesticAnimals)
            @Html.ValidationMessageFor(model => model.DomesticAnimals)
        </div>
    </div>
```

Figure 2. 16 Exemple de codage pour l'application en langage HTLM

Le résultat final de l'ajout des champs dans l'application est dans les figures (2.17 et 2.18).

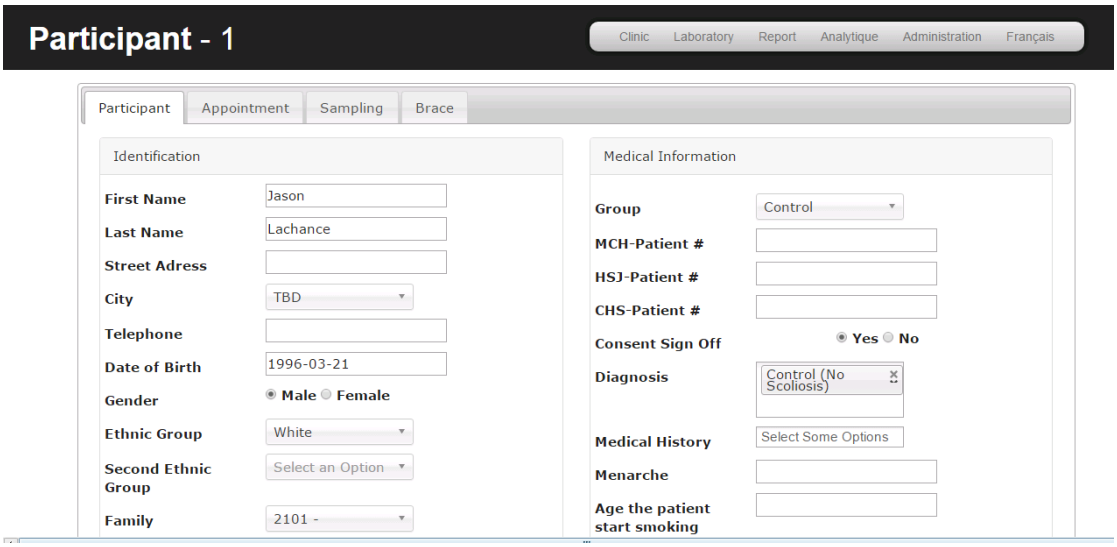



Figure 2. 17 Fenêtre de l'application pour les champs du « participant »

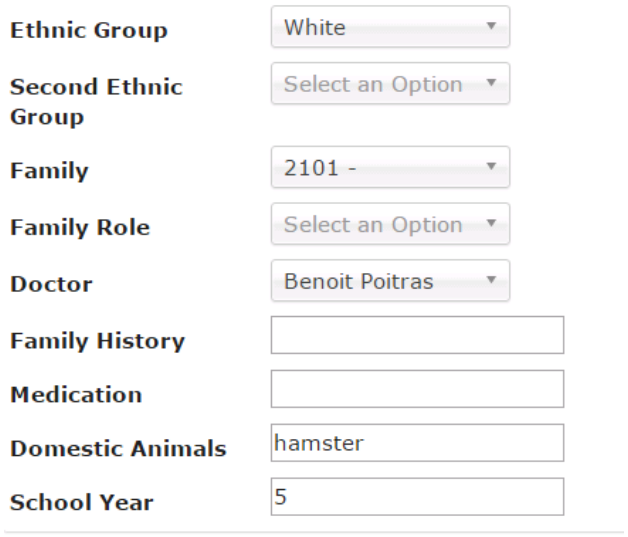



Figure 2. 18 Exemples des champs ajoutés pour la fenêtre participant

## 2.11 Conclusion

Ce chapitre a permis d'analyser les systèmes sources, de fixer la portée du projet et d'effectuer les différentes étapes du processus ETL pour ce cas pratique. Le chapitre suivant présentera les tests réalisés afin de valider les fonctionnalités d'accès aux données et le bon fonctionnement de la nouvelle application Web suite à la centralisation de toutes les données cliniques du laboratoire dans un seul entrepôt.

# CHAPITRE 3

# Tests et validation

## 3.1 Introduction

À chaque étape du projet des tests sont réalisés pour tester les nouvelles modifications, ainsi l'impact et le bon fonctionnement des fonctionnalités antérieures. Les résultats de ces tests sont discutés puis approuvés par le client à chaque étape. Plusieurs types de tests ont été nécessaires pour réaliser ce projet.

## 3.2 Tests fonctionnels et les tests IHM de l'application Web

Ces tests sont effectués manuellement, les tests fonctionnels permettent de vérifier que tous les liens fonctionnent : l'application web contient plusieurs fenêtres (les participants, les rendez-vous, les médicaments...), chaque fenêtre contient plusieurs fonctionnalités à chaque modification de l'application, une vérification du bon fonctionnement de ces fonctionnalités est réalisée.

Les tests IHM permettent de vérifier la présentation visuelle (la forme, les couleurs, la résolution...) des menus, des boutons et des fenêtres.

Ces tests sont réalisés en utilisant le « Nuget Package », qui est un gestionnaire de progiciel fourni par Visual Studio et qui permet d'utiliser des bibliothèques pour faciliter le développement. Les « Package » utilisés sont :

« Selenium.WebDriver » : son rôle est d'envoyer les commandes de test aux différents navigateurs web (Internet Explorer, Firefox, Chrome …).

« Selenium.Support » : ce « package » comporte des bibliothèques pour supporter les classes HTML, sélectionner les éléments et appliquer les conditions de test.

Le test de l'application Web en utilisant « Selenium » comporte plusieurs étapes, la première étape et de sélectionner l'élément à tester (champ, bouton...) et d'inspecter l'arbre DOM pour obtenir l'ID de cet élément (figure 3.1). L'ID de l'élément est utilisé par le code de test de l'élément, tester l'insertion du nom dans le champ «First Name», (figure 3.2). L'exécution du

programme permet d'accéder à l'application Web et de tester l'élément. Le résultat de l'insertion du nom est illustré par la figure 3.3.

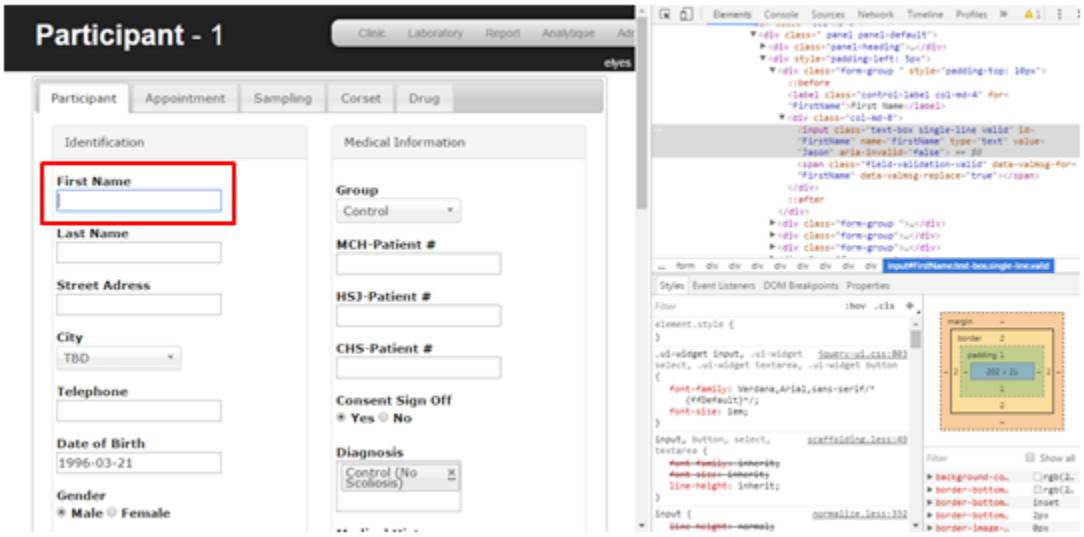



Figure 3. 1 Inspection de l'arbre DOM pour obtention de l'ID

```
using System;
using System.Collections.Generic;
using System.Linq;
using System.Text;
using System.Threading.Tasks;
using OpenQA.Selenium;
using OpenQA.Selenium.Firefox;

namespace TestAppWeb
{
    0 références
    class Program
    {
        0 références
        static void Main(string[] args)
        {
            IWebDriver driver = new FirefoxDriver();
            driver.Navigate().GoToUrl("http://localhost:10110/Participant/edit/1");
            driver.Manage().Window.Maximize();
            IWebElement searchInput = driver.FindElement(By.Id("FirstName"));
            searchInput.SendKeys("andré Jean");
            searchInput.SendKeys(Keys.Enter);
        }
    }
}
```

Figure 3.2 Code de test d'insertion de données dans le champ « First Name »

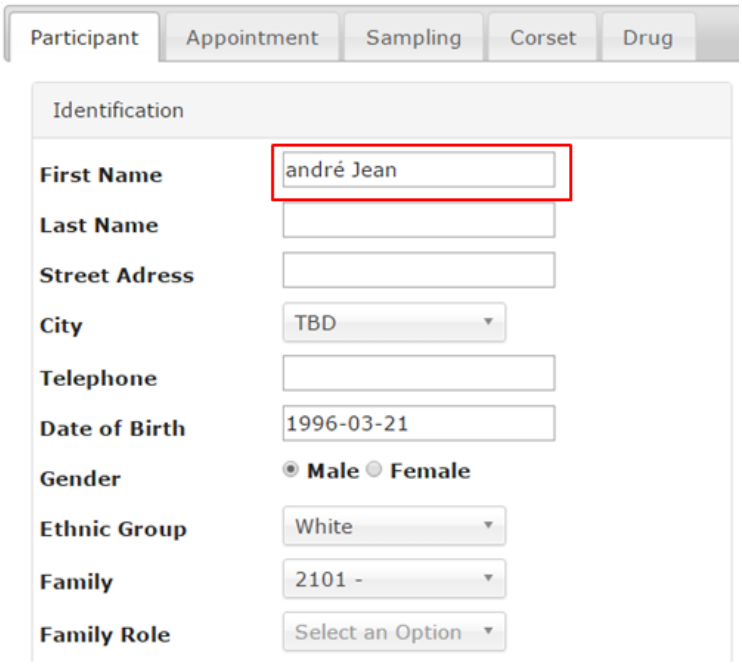



Figure 3. 3 Résultat de test d'insertion dans le champ « First Name »

La même méthode est appliquée pour tester les autres champs, boutons…et à chaque modification du code de l'application.

### 3.3 Tests manuels de la base de données

Pour chaque étape du processus ETL (extraction, transformation et chargement) une vérification l'intégrité ainsi que la bonne intégration des données est effectuée. L'utilisation des commandes SQL a permis de tester manuellement la base de données cible, vérifier les transformations et comparer avec la base de données source. Un exemple de code SQL est donné dans la figure 3.1.

```
SELECT id FROM [dbo].[Participants] WHERE
UPPER(FirstName) ='DA%'
and UPPER(LastName) ='BEK%'

SELECT Doctor_Id FROM [dbo].[Participants] WHERE [City_Id] =200
or Group_Id= 6

SELECT Count(*) FROM  [dbo].[Corsets]WHERE id is null

SELECT User_Id FROM [dbo].[Appointments]WHERE id = 4080
```

Figure 3.4 Exemple de tests manuels de base de données

Les tests SQL ont permis une vérification rapide et spécifique du résultat de transformation sur les cas particuliers et de les vérifier un par un quand c'est possible, dans le cas où le nombre de champs particulier à tester est grand, l'utilisation des tests automatiques devient plus adéquate.

### 3.4 Validation client

Une fois les tests sont effectués et validés, une lise de fonctionnalités et de données est établie pour être validé avec le client. La liste contient les nouvelles fonctionnalités ajoutées ainsi que les fonctionnalités réalisées dans la première itération. La liste retrace les exigences du client établi au début du projet. Le tableau 3.1 montre le gabarit à remplir par le client afin de valider les transformations. Après validation, le document de « validation des champs et des données » est approuvé et signé par le client (annexe V).

Tableau 3. 1 Gabarit à remplir par le client pour valider les transformations

| **Champ** | **Validation des Fonctionnalités** | **Validation des Données** |
|---|---|---|
| **Participants** | | |
| Liste des participants | | |

| | | |
|---|---|---|
| Ajouter un participant | | |
| Éditer un participant | | |
| Éditer les détails du participant | | |
| Ajouter un rendez-vous | | |
| Éditer un rendez-vous | | |
| Éditer un prélèvement du participant | | |
| **Corset** | | |
| Ajouter un corset | | |
| Éditer un corset | | |
| Supprimer un corset | | |
| **Laboratoire** | | |
| Liste des prélèvements | | |
| Éditer un prélèvement | | |
| Administration | | |
| Liste des utilisateurs | | |
| Ajouter un utilisateur | | |

| | | |
|---|---|---|
| Éditer un utilisateur | | |
| Liste de données de l'application | | |
| Conditions | | |
| Diagnostique | | |
| État d'échantillon | | |
| **Groupe** | | |
| Groupe ethnique | | |
| Rôle de famille | | |
| Type d'échantillon | | |
| Type de chirurgie | | |
| Type de Cobb | | |
| Type de Corset | | |
| Changer de langue | | |
| Français | | |
| Anglais | | |

## 3.5 Conclusion

Ce chapitre développe les tests effectués sur l'application Web en utilisant le progiciel « Selenium » ainsi que les tests de la base de données, en utilisant les commandes SQL. Ces tests ont confirmé le bon fonctionnement de l'application ainsi que la bonne intégration des données. De plus, un gabarit de validation a été rempli et approuver par le client confirme la réalisation des objectifs fixés pour ce projet.

# CONCLUSION

Ce projet de 15 crédits était une occasion précieuse d'appliquer un processus ETL dans un environnement professionnel avec les spécifications et les exigences d'un environnement médical. La transformation, le nettoyage et la restructuration, effectués dans ce projet, des bases de données sources dans une seule base de données faciliteront le travail de recherche de données.

En plus de créer une base de données conforme et de définir les relations entre les données, la nouvelle interface Web aide à une insertion correcte des données par l'utilisation des menus déroulants et message d'aide.

Ce projet m'a permis aussi de me familiariser avec les technologies .Net, totalement nouvelles pour moi et d'appliquer les bonnes méthodologies de suivi et de gestion de projet. Finalement, ce projet est considéré comme une base solide qui permettra le développement futur de fonctions analytiques.

## RECOMMANDATIONS

Ce projet qui consiste à une deuxième itération de l'application Web médicale pour l'insertion et la gestion de données a permis de normaliser et d'intégrer toutes les données dans un seul entrepôt.

Les recommandations pour les prochaines itérations qui visent à améliorer cette application sont :

- D'exploiter les données normalisées dans la base de données pour faire des analyses analytiques, selon les paramètres et les exigences des chercheurs. Il est recommandé aussi que le résultat de ces analyses peut être représenté de différentes façons (tableau, figure...).

- De sécuriser les données critique et sensible dans la base de données de l'application par le chiffrement de ces données stockées afin qu'en cas d'attaque, il n'ait accès qu'aux données chiffrées. Il faut aussi sécuriser les clés de chiffrement et les stocker dans un autre serveur que celui de la base de données.

- De protéger l'application contre les injections SQL, comme le but premier de l'application est l'intégration de données, il y a un risque d'injection SQL qu'il faille remédier.

- D'utiliser l'infonuagique pour enregistrer les données à distance, après la phase de cryptage. Ce qui permettra de sécuriser les données en cas ou les disques durs locaux sont endommagés.

# ANNEXE I

## Implémentation et restauration de la base de données dans la VM

La restauration de la base de données et une étape importante du projet

Première étape :

La première étape est de sauvegarder la base de données modifiée pour la restaurer dans la machine virtuelle.

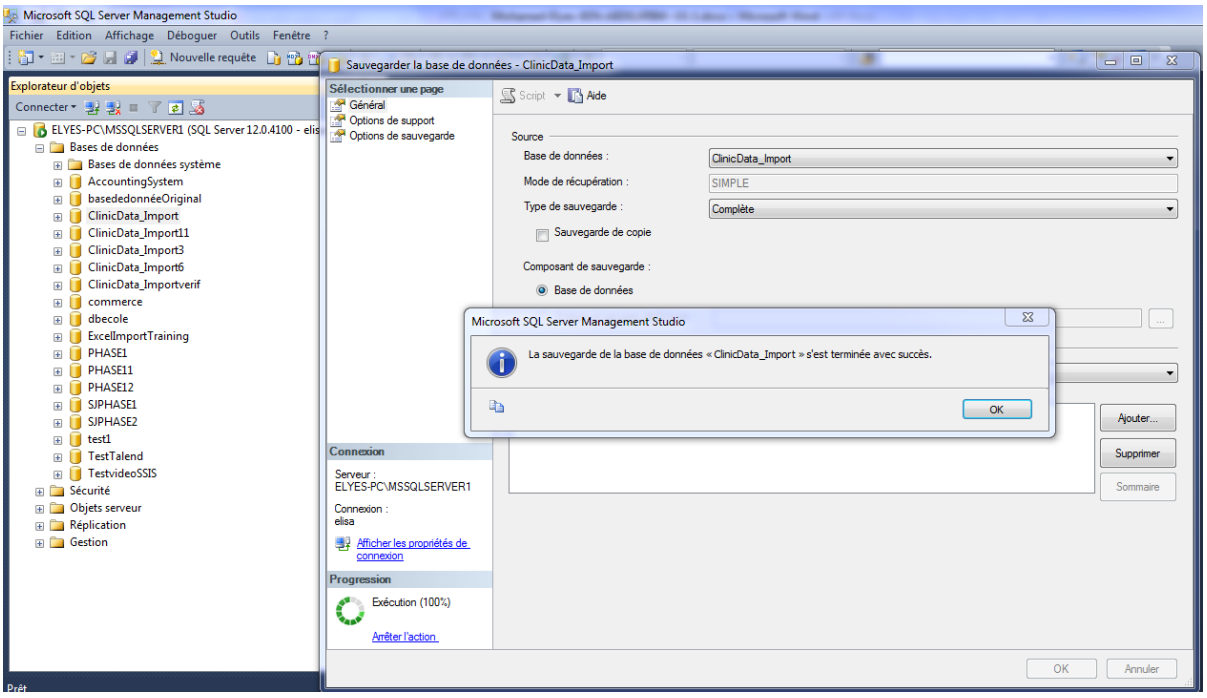



Figure I. 1 sauvegarde de la base de données modifiée

Deuxième étape :

La deuxième étape est la transmission et la restauration la base de données à la machine virtuelle. Il faut aussi redémarrer le service SQL server pour pouvoir restaurer la BD.

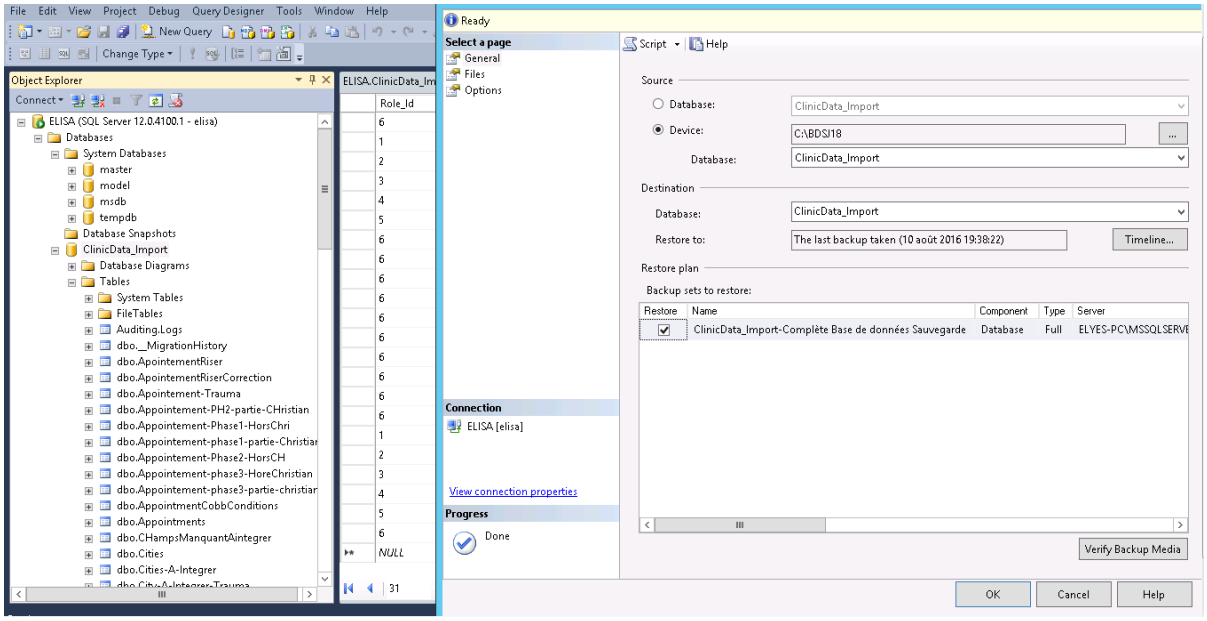


Figure I. 2 Restaurer la base de données dans la machine virtuelle

Troisième étape :

La troisième étape consiste à lancer l'application et à mettre à jour les droits d'accès.

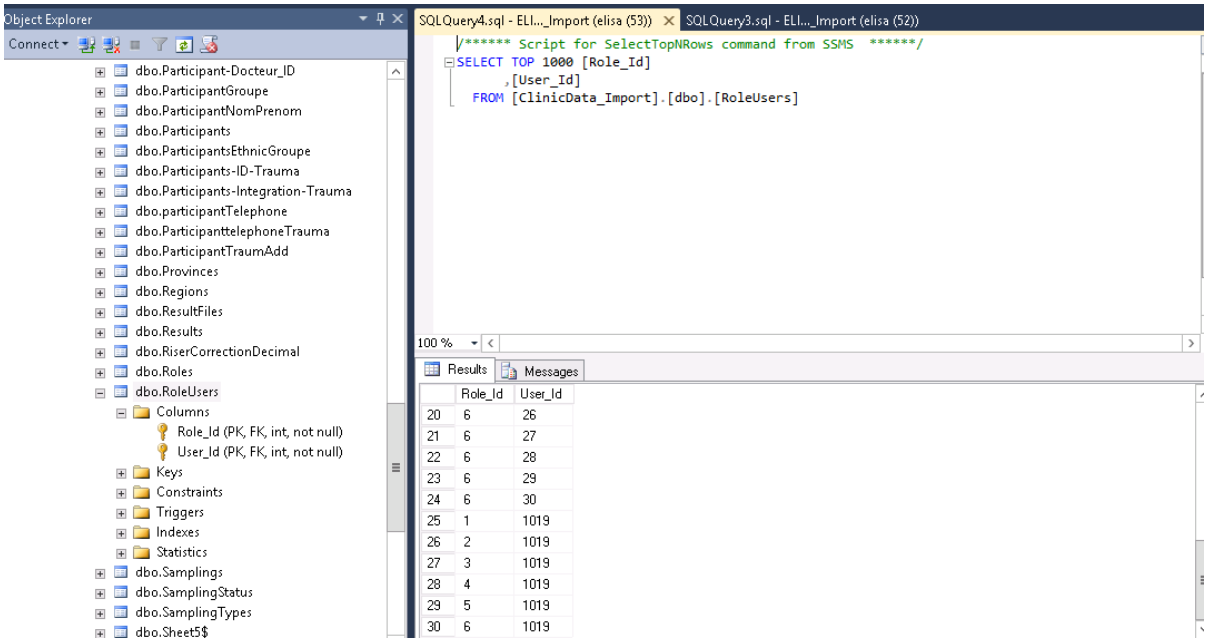


Figure I. 3 Mise à jour des droits d'accès

Quatrième étape :

La Quatrième est la dernière étape consiste à redémarrer le service IIS afin d'actualiser les modifications.

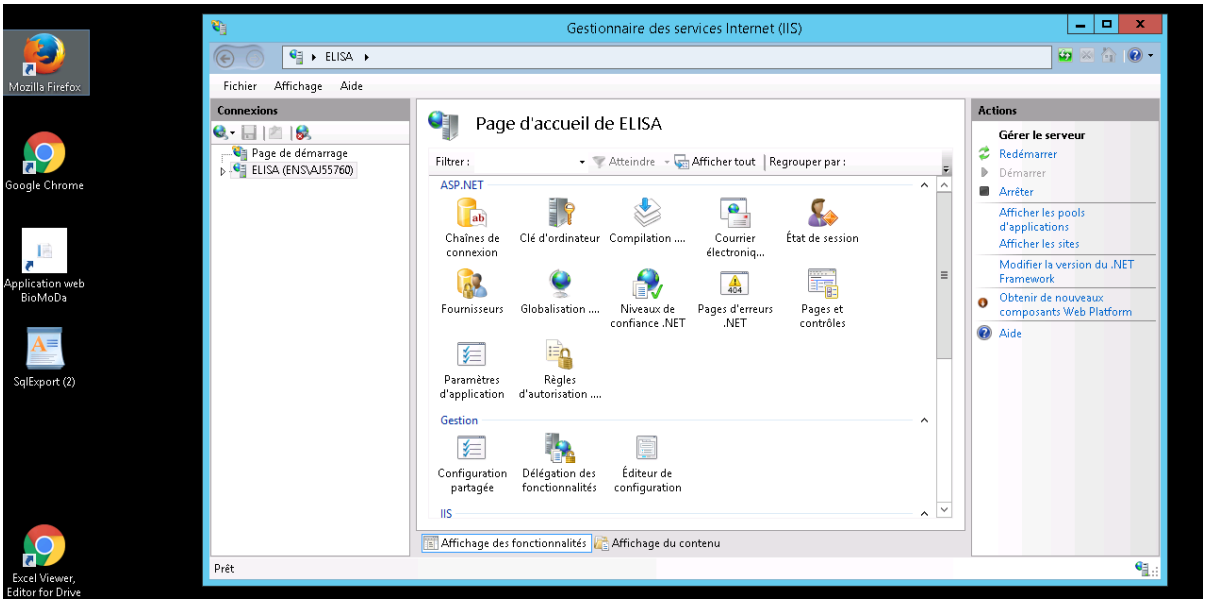
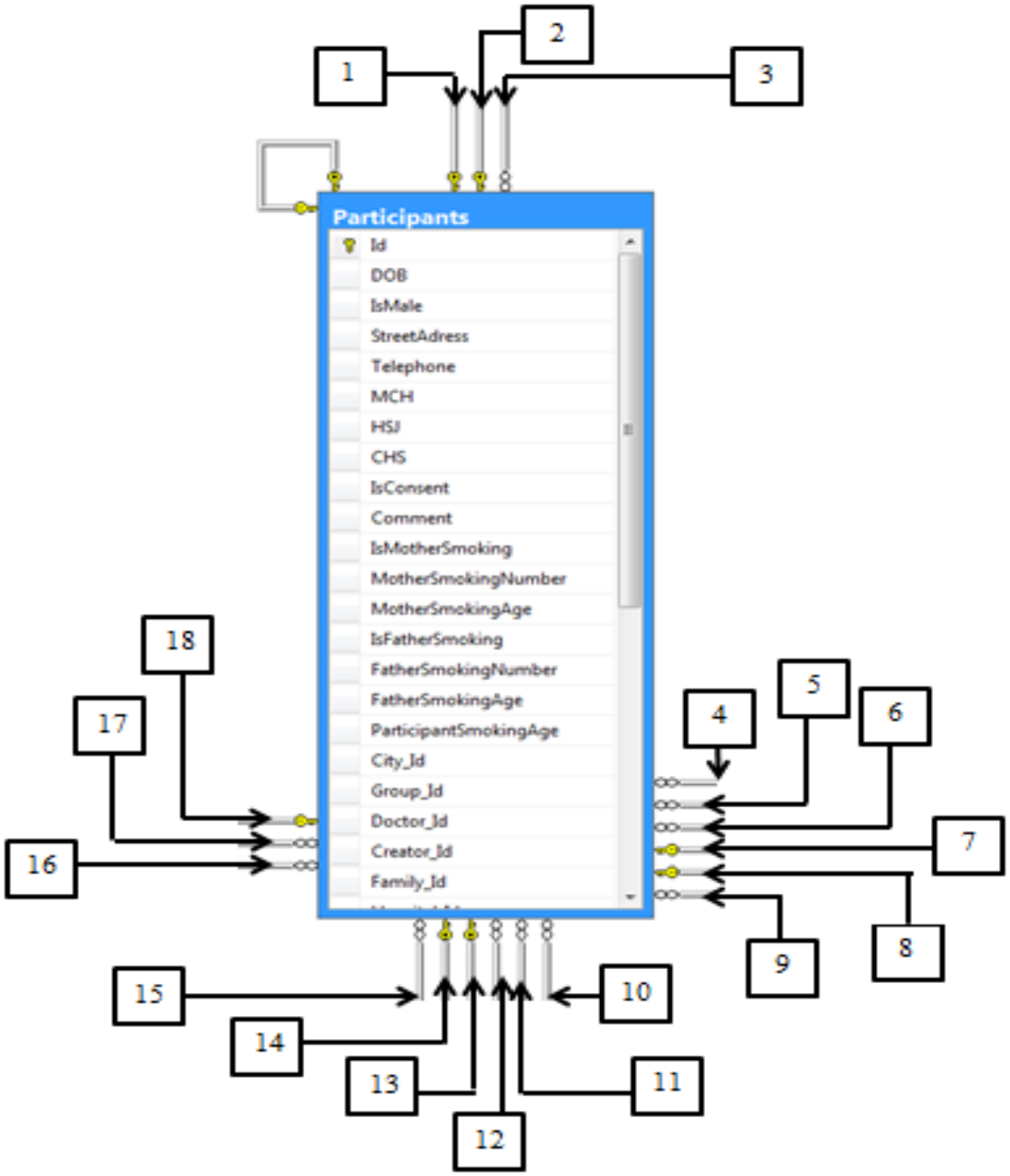
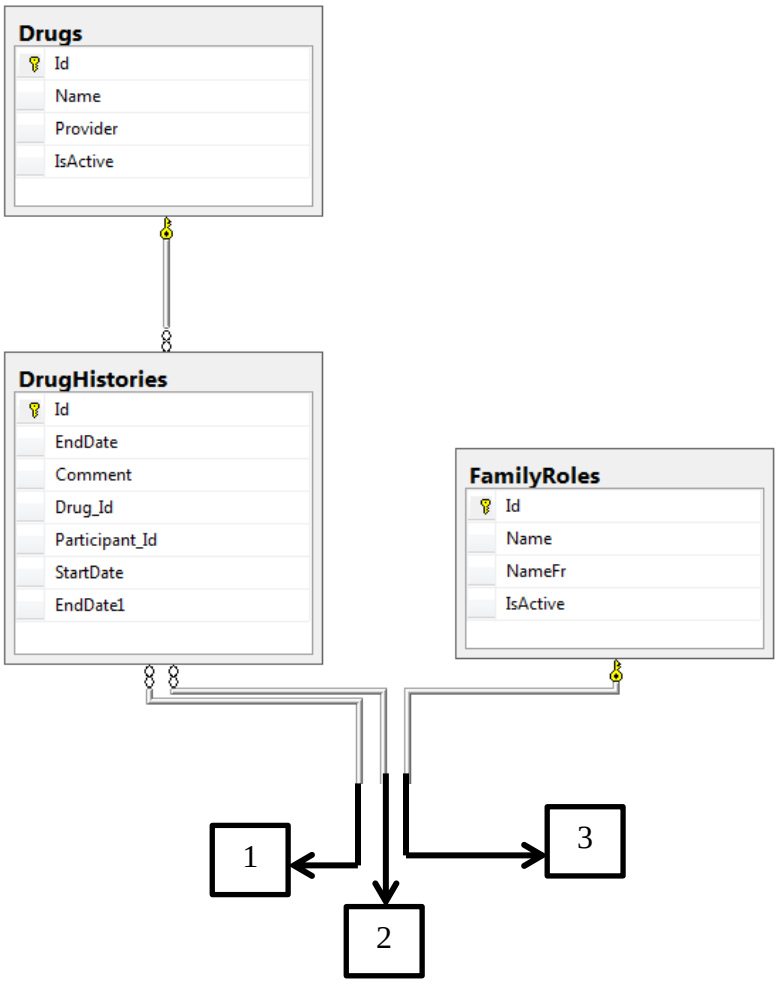


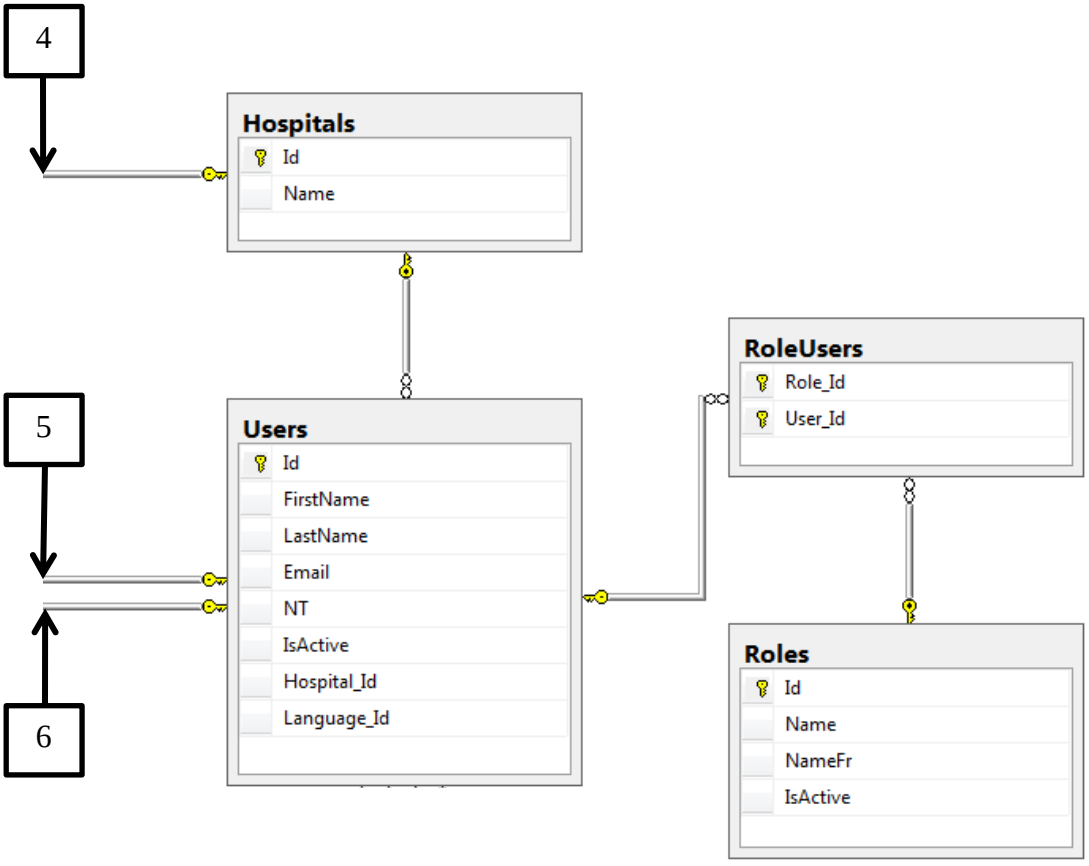

Figure I. 4 Redémarrage de service IIS

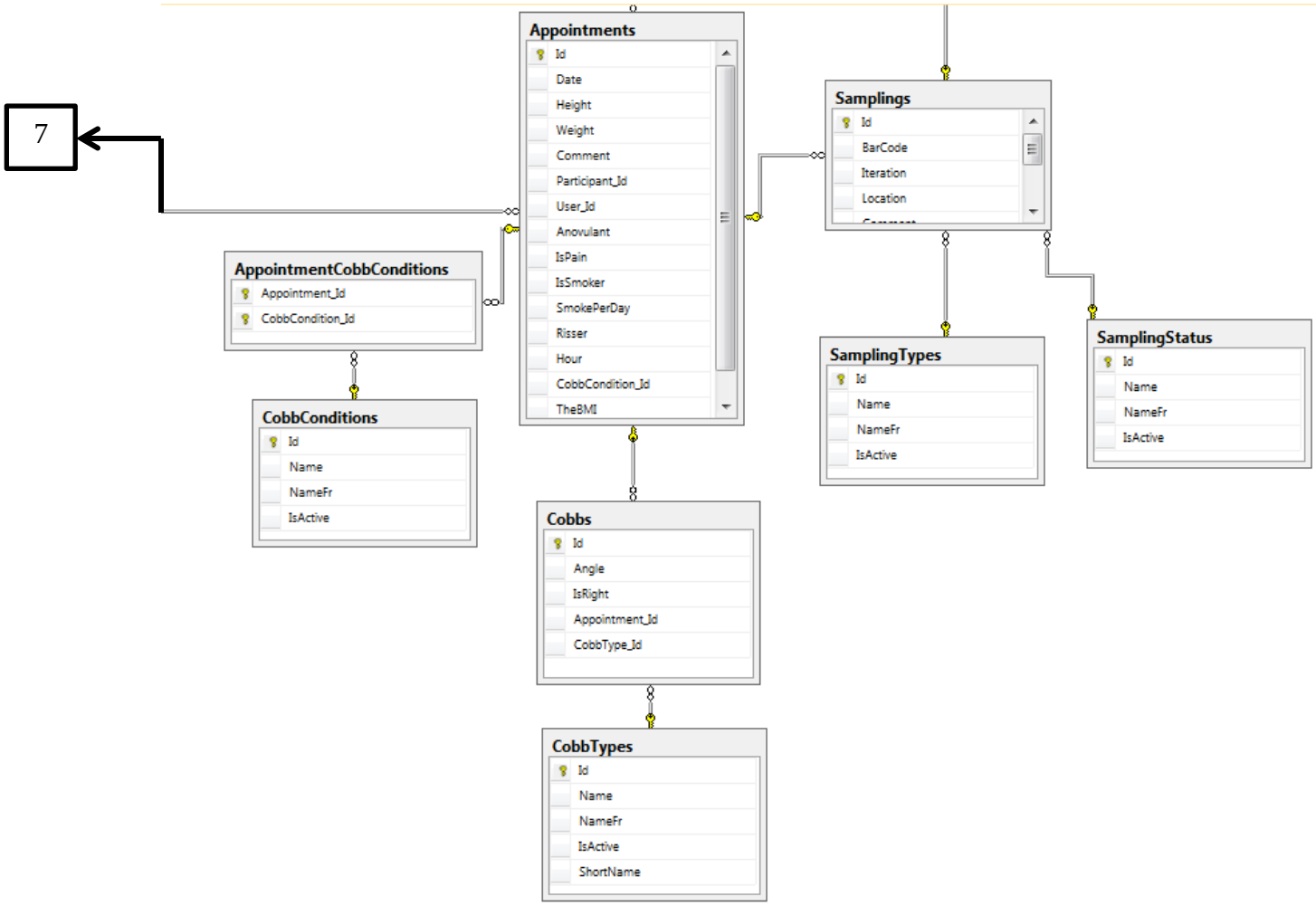
# ANNEXE II

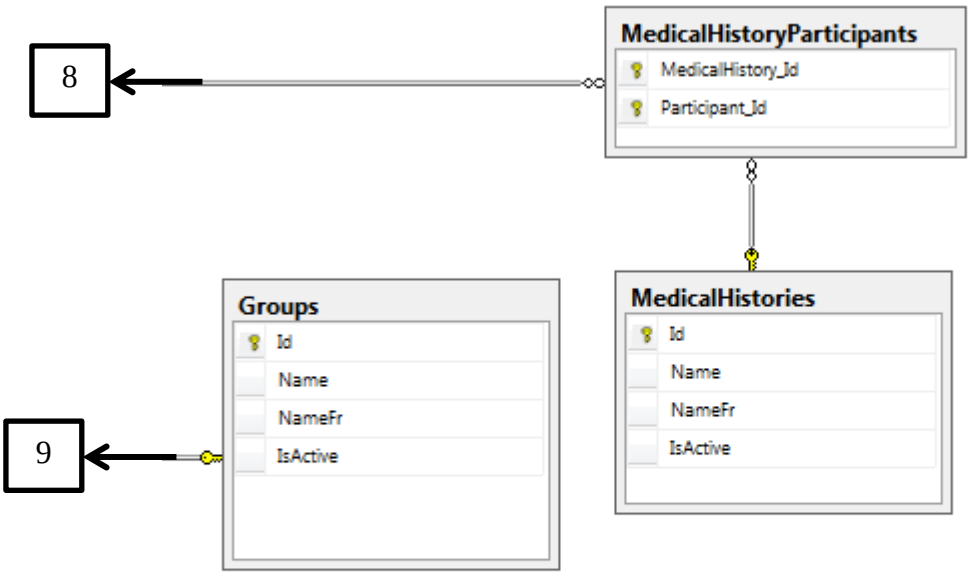
## Schéma de la base de données

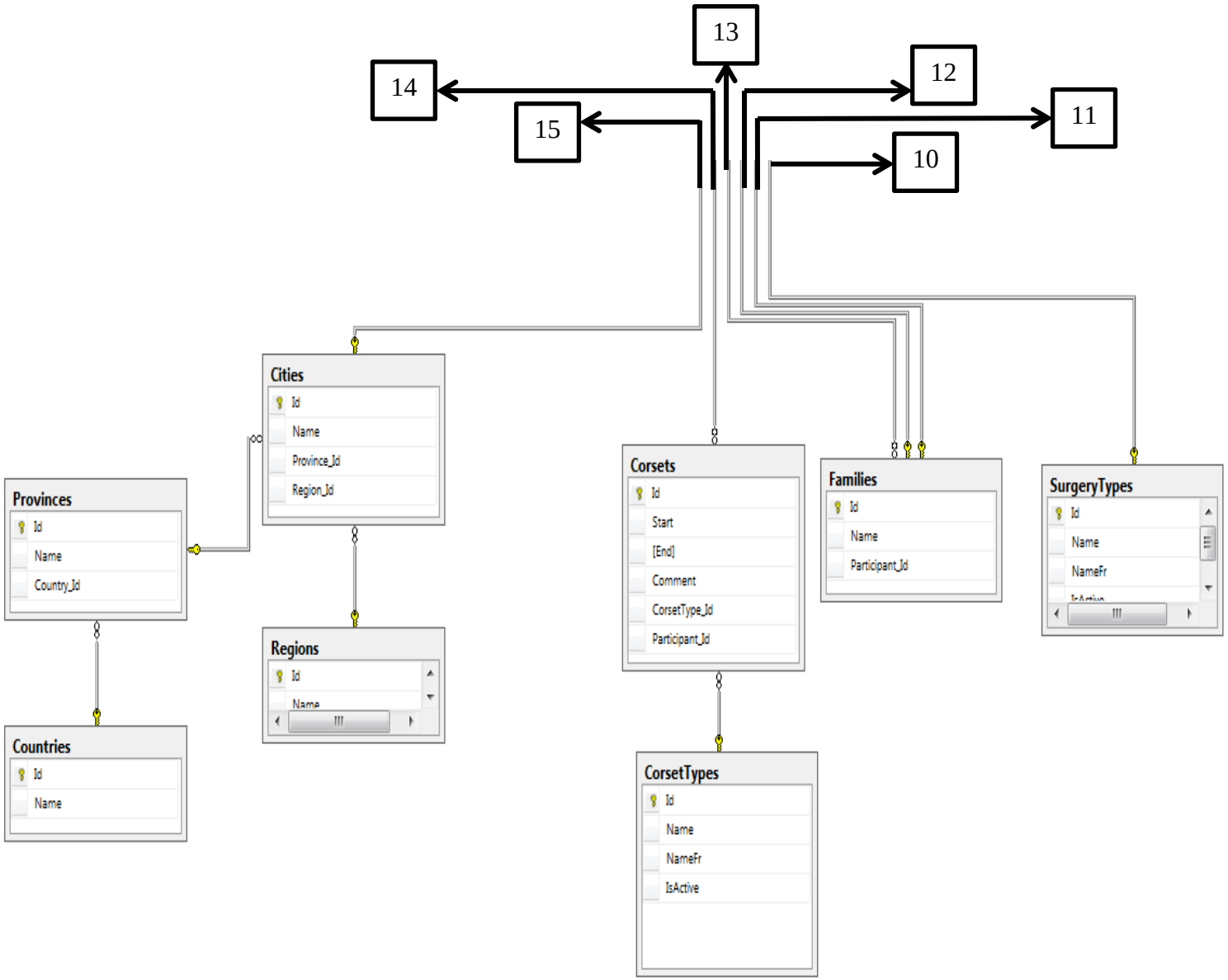


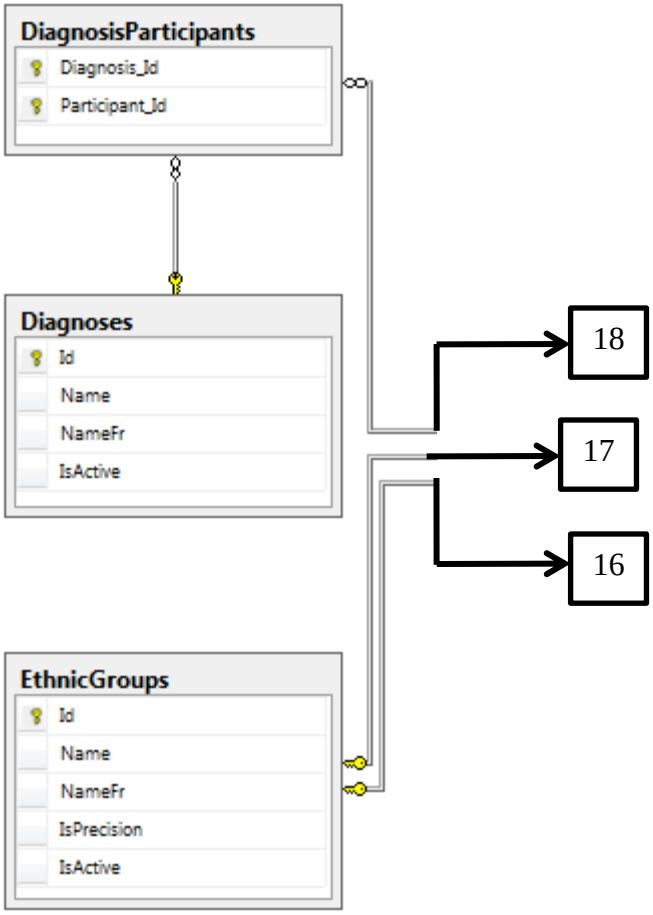

Drugs
Id
Name
Provider
IsActive
DrugHistories
Id
EndDate
Comment
Drug_Id
Participant_Id
StartDate
EndDate1
FamilyRoles
Id
Name
NameFr
IsActive
1
2
3


4
Hospitals
Id
Name
5
Users
Id
FirstName
LastName
Email
NT
IsActive
Hospital_Id
Language_Id
6
RoleUsers
Role_Id
User_Id
Roles
Id
Name
NameFr
IsActive


7
Appointments
Id
Date
Height
Weight
Comment
Participant_Id
User_Id
Anovulant
IsPain
IsSmoker
SmokePerDay
Risser
Hour
CobbCondition_Id
TheBMI
Samplings
Id
BarCode
Iteration
Location
AppointmentCobbConditions
Appointment_Id
CobbCondition_Id
CobbConditions
Id
Name
NameFr
IsActive
Cobbs
Id
Angle
IsRight
Appointment_Id
CobbType_Id
CobbTypes
Id
Name
NameFr
IsActive
ShortName
SamplingTypes
Id
Name
NameFr
IsActive
SamplingStatus
Id
Name
NameFr
IsActive


8
MedicalHistoryParticipants
MedicalHistory_Id
Participant_Id
MedicalHistories
Id
Name
NameFr
IsActive
Groups
Id
Name
NameFr
IsActive
9


13
14
12
15
11
10
Cities
Id
Name
Province_Id
Region_Id
Provinces
Id
Name
Country_Id
Regions
Id
Name
Countries
Id
Name
Corsets
Id
Start
[End]
Comment
CorsetType_Id
Participant_Id
CorsetTypes
Id
Name
NameFr
IsActive
Families
Id
Name
Participant_Id
SurgeryTypes
Id
Name
NameFr


DiagnosisParticipants
Diagnosis_Id
Participant_Id
Diagnoses
Id
Name
NameFr
IsActive
EthnicGroups
Id
Name
NameFr
IsPrecision
IsActive
18
17
16

**ANNEXE III**

**Exemple de données sources à traiter**

| **COMMENTAIRES** | **COBB RÉSIDUEL** | **RISSER** | **MÉNARCHE** | | **VISITE PRÉ-OP** | **COBB Pour décision S.O.** | **COBB juste PRÉ-OP** |
|---|---|---|---|---|---|---|---|
| | | | | | | | |
| | 19-15-15 TgTdLg | 4 | Juillet 2004 | lg | 14/01/2008 | 65 Td Risser 4 (OK pas copie) | |
| | | | | | | | |
| | 8-16 TgLd | 5 | Février 2006 | | 25/03/2008 | 40 Ld (OK Copie) | |
| | | | Juin 2008 | | 16/06/2008 | 63 Td Risser 4 | Rx du 09-01-05 : 64 Td |
| | 0 | 5 | Juillet 2000 | | 26/08/2008 | 46-52 TdLg Risser 5 (OK pas copie) | Rx du 09-06-16 : 43-52 TdLg |
| | | | | | | | |

| | | | | | | | |
|---|---|---|---|---|---|---|---|
| | 13 TLg | 5 | N/A | | 23/04/2008 | 63 Lg | |
| | **REFUSE** | 4 | Octobre 2006 | 4 | 17/12/2008 | 52 Td Risser 4 (OK Pas copie) | |
| | | | Août 2009 | | | | |
| Spondylo | Pas scoliose = OK | Pas inscrit | Juin 2006 | | 04/05/2009 | Spondylo L4-L5 et HD L5-S1 Risser 5 (OK Pas copie) | |
| Congénitale sur Hémi-vertèbre L5 | | | N/A | | 27/05/2009 | 48 Ld Risser 4 (OK Pas copie) | 09-07-21 : 49 Ld |
| | 21-10 TgTd | Pas inscrit | Mai 2006 | | 20/04/2009 | 37-69-24 TgTdLg Risser 4 (OK Pas copie) | |
| | 30 Td | 5 | Mai 2007 | | 08/06/2009 | 42-79-46 TgTdLg (OK Pas Copie) | |
| | 17-22 TgTd | 5 | N/A | | 09/04/2009 | 32-66-26 TgTdLg | |

| | | | | | | | |
|---|---|---|---|---|---|---|---|
| | | | | | | Risser 4 (OK Pas Copie) | |
| | | | | | | | |
| | 35-28 TdLg | 1 | N/A | | 25/03/2009 | 49-61 TdLg Risser 0 (OK Copie) | |
| | 1 TLd | Pas inscrit | N/A | | 09/03/2009 | 56 TLd Risser 4 (OK Pas copie) | |
| | | | | | | | |
| | | | | | | | |
| Spondylo grade 2 | OK | | N/A | | 05/05/2009 | Spondylo grade 2 | |
| Spondylo | Non | Pas inscrit | Octobre 2006 | | 21/07/2009 | Spondylo L5-S1 grade 3/5 | |
| | | | | | | | |
| | 17-9 TdLg | 4 + | Juillet 2008 | | 08/08/2008 | 65 Td Risser 0 + | Rx du 09-09-28 : 81 Td |
| | 22 Td | 3 | Mai 2008 | | 08/04/2009 | 42-66-38 TgTdLg Risser | |

| | | | | | | | |
|---|---|---|---|---|---|---|---|
| | | | | | | 3 (OK Pas copie) | |
| Cette patiente porte Spinecor, n'est pas opérée Cf commentaire | | | Non | | | | |
| | | | | | | | |
| | **REFUSE** | | | | | | |
| Évolution de la scoliose en commentaires | 19 Td | 5 | N/A | | 22/12/2010 | 27-59-30 Tg sup TdLg Risser 4 (OK Pas copie) | Rx du 11-06-08 : 30-56-29 Tg sup TdLg |
| Pas de prél. au moment S.O. | 21-16 TdLg | 5 | Mai 2007 | | 25/03/2009 | 34-55 TdTLg Risser 4 (OK Pas copie) | |
| Marfan | 36-29-10 TgTdLg | 1 + | N/A | | 22/10/2008 | 40-59-58 TgTdLg Risser 1 (OK Pas copie) | |
| | | | | | | | |

| | | | | | | | |
|---|---|---|---|---|---|---|---|
| | 16-18 TdLg | Pas inscrit | Mai 2003 | | 02/03/2009 | 45-42 TdLg Risser 4 | |
| | 15-15 TdLg | 3 | Décembre 2008 | | 28/09/2009 | 75-52 TdLg Risser 3 (OK Pas Copie) | |
| Prél. non faits par chirurgien | | | | | | | |
| Scoliose congénitale + Atrésie œsophage | 51-28 Tg sup. TLd | 5 | Décembre 2003 | | 04/12/2009 | 73 Tg sup. Risser 4 (OK Pas Copie) | |
| | 12 Td | 0 | Non | | 12/05/2009 | 50 TLd Risser 0 (OK Pas copie) | Rx du 09-12-07 : 52 TLd |
| | | | | | | | |
| | 20-17 TgTLd | 5 | Juin 2007 | | 14/04/2009 | 57-57 TgTLd Risser 5 (OK Pas Copie) | Rx du 10-01-08 : 50-56 TgLd |
| Mère refuse X 2 | | | | | | | |
| | 28-21 TdLg | 4 | Juillet 2009 | | 29/06/2009 | 74-66 TdLg Risser 3 | Rx : 2010-02-12 64-71 TdLg |

| | | | | | | | |
|---|---|---|---|---|---|---|---|
| | | | | | | (OK pas Copie) | |
| | | | | | | | |
| Paralysie cérébrale | 44 TLd (assise) | 5 | ? | | 23/06/2009 | 105 TLd (OK pas de Copie) | Rx du 10-01-15 : au-dessus de 125 TLd |
| | 0 | Pas inscrit | Mars 2006 | | 04/01/2010 | Spondylo L5-S1 + Spondylolise L5 | |
| | 14 TLd | 5 | Mai 2004 | | 08/08/2008 | 47 Td (OK Pas Copie)<br>Risser 5 | Rx du 10-01-18 : 46 Td |
| | 24-31 TdLg + 1 cm G | 5 | Novembre 2006 | | 18/11/2008 | 86 TLg Risser 4 (OK Pas Copie) | Rx du 10-01-22 : 50-89 TdTLg |
| | 11 Lg | Pas inscrit | N/A | | 28/07/2009 | Spondylo L5 bil. | Rx du 2010-01-07 : 11-18 TdLg |
| | 8 Tg | 4 | Août 2005 | | 09/02/2010 | 43 TLg Risser 4 (OK Pas copie) | Rx du 10-01-25 : 48 TLg |
| | | | | | | | |
| Enfant handicapé, mère | **REFUSE** | | | | | | |

| | | | | | | | |
|---|---|---|---|---|---|---|---|
| refuse | | | | | | | |
| Exérèse de matériel en Février 2011 | 17-31-15 Tg sup.TdLg | 5 | Juin 2001 | | 09/11/2007 | 27-56-30 Tg sup.TdLg Risser 5 (OK Copie) | Rx du 10-02-01 : 27-56-29 Tg sup.TdLg |
| | 10-14-12 TgTdLg | 4 | Avril 2009 | | 23/06/2009 | 28-58-21 TgTdLg (OK Pas Copie) Risser 0 | Rx du 10-02-01 : 27-62-30 TgTdLg Risser 3 |
| | | | | | | | |
| | | | | | | | |
| | 8-9 TdTLg | 4 | Août 2006 | | 09/04/2009 | 36-58 TdTLg Risser 4 (OK Pas copie) | Rx du 10-02-05 : 31-55 Td TLg |
| | | | Mars 2004 | | 08/08/2008 | 50 Td Risser 5 | Rx du 10-02-05 : 50 Td |
| | 16-23-21 TgTdLg | 3 | Non | | 18/01/2010 | 37-81-47 TgTdLg Risser 3 (Pas Copie) | OK |

| | | | | | | | |
|---|---|---|---|---|---|---|---|
| | 16-2 TgTd | 4 | Novembre 2007 | | 29/04/2009 | 51 Td Risser 4 | Rx du 10-02-12 : 53 Td |
| | 13-13 TgLd | 4 | Octobre 2006 | | 11/03/2009 | 69 Td Risser Risser 4 (OK Pas Copie) | Rx du 10-02-22 : 68 Td |
| | 15 Td | Pas inscrit | N/A | | 31/08/2009 | 56 Td (OK Pas Copie) | Rx du 10-02-03 : 30-58 TgTd Risser 4 |
| | Minime lombaire | | Avril 2009 | | 04/11/2009 | 54 TLg Risser 4 | Rx du 10-02-05 : 50 TLg |
| | | | | | | | |
| | | | | | | | |
| | 8-9 TdLg | 5 | Novembre 2006 | | 14/04/2009 | 20-48-65-33 Tg sup.TdTLgLd Risser 5 | Rx du 10-03-16 : 21-50-67-31 Tg supTdTLgLd |
| | 20-27 TdLg | 5 | Janvier 2004 | | | | Rx 10-03-26 : 36-60-41 TgTdLg |
| Projet Guoruey 1 | | | | | | | |
| Projet Guoruey 2 Prél. refait le 11-07-14 | 20 TLg | 5 | Décembre 2007 | | 09/12/2009 | 29-50 TdTLg Risser 4 (OK Pas Copie) | Rx 11-12-07 : 29-53 TdTLg |

| | | | | | | | |
|---|---|---|---|---|---|---|---|
| Projet Guoruey 15 (prél. refait pour Guoruey avec 3 tubes lavande) | | | | | | | |
| Projet Jumelle fait le 10-03-08 + Refait le 10-08-09 avec tubes bleus | 18-21 Tg sup.Td | 5 | Juillet 2008 | | 12/08/2009 | 52-32 TdLg Risser 3 (OK Pas Copie) | Rx 10-08-09 : 31-57-31 Tg sup.TdLg Risser 4 |
| | | | | | | | |
| | | | | | | | |
| | | | | | | | |
| | | | | | | | |
| | | | | | | | |
| | | | | | | | |
| | | | | | | | |
| | | | | | | | |
| Projet Guoruey 28 | | | Juin 2010 | | | | |
| Projet Guoruey 29 | | | N/A | | | | |
| Projet Guoruey 31 | | | Novembre 2009 | | | | |

| | | | | | | | |
|---|---|---|---|---|---|---|---|
| Projet Guoruey 34 | 12 Td | 4 | Janvier 2010 | | 26/08/2010 | 66-33 TdLg Risser 1 (OK Pas Copie) | Rx 11-02-28 : 36-76-38 Tg supTdLg |
| Ce patient n'est pas opéré Cf commentaire | | | N/A | | | | |
| Projet Guoruey 36 | | | Avril 2010 | | | Questionnaire environnemental rempli | |
| | | | | | | | |
| | 25-29 Tg supTd | 5 | Novembre 2008 | | 16/08/2010 | 25-49 TgTLd Risser 4 | Rx du 11-02-07 : 27-52 Tg TLd |
| | | | | | | | |
| | | | | | | | |
| | 14-14-18 Tg sup. TdLg | 1 | Non | | 13/12/2010 | 73-54 TdLg Risser 0 (OK Pas copie) | Rx du 11-02-22 : 32-82-57 TgTdLg |
| Projet Guoruey 39, opérée le 13-06-11 | | | Septembre 2008 | | | Questionnaire environnemental rempli | |

| | | | | | | | |
|---|---|---|---|---|---|---|---|
| | 19 Lg | 5 | Juin 2007 | | 07/09/2010 | 41 TLg Risser 4 | Rx du 11-01-19 : 43 TLg Risser 4 |
| | 16-18-16 TgTdLg | 5 | Novembre 2008 | | 30/11/2010 | 27-57-45 Tg sup.TdLg Risser 4 | Rx du 11-02-01 : 29-58-46 Tg sup.TdLg |
| | 15 Lg | 4 | Août 2008 | | 19/01/2011 | 57-38 TdLg Risser 4 (OK Pas Copie) | Rx du 11-04-11 : 59-38 TdLg Risser 4 |
| Scoliose congénitale | 19 Lg | 3 | NA | | 18/01/2011 | 57 Lg Risser 3 (OK Pas Copie) | Rx du 11-05-03 : 55 Lg Risser 3 |
| | 15-15 TdLg | 4 + | Mai 2009 | | 30/09/2010 | 54-35 TdLg Risser 4 | Rx du 11-03-24 : 55-35 TdLg |
| | 13-11 Tg supTLd | 4 | Mars 2005 | | 01/03/2010 | 39 Td Risser 4 | Rx du 11-05-09 : 43 Td |
| Cypho-Scoliose post Tuberculose vertébrale | 15-14 TdLg | 5 | N/A | | 24/02/2011 | 50-37 TdLg Risser 5 | Rx du 11-03-30 : 45-33 TdLg |
| | 9 Td | 5 | N/A | | 17/02/2011 | 24-50-36 Tg supTdLg Risser 3 (OK Pas Copie) | Rx du 11-07-04 : 25-56-41 Tg supTdLg |

| | | | | | | | |
|---|---|---|---|---|---|---|---|
| | 16-16 TdLg | 4 | Mars 2009 | | 20/12/2010 | 55-41 TdLg Risser 3 + (OK Pas copie) | Rx du 11-06-23 : 25-58-41 Tg supTdLg |
| | 0 | 3 + | Avril 2010 | | 15/09/2010 | 47-25 TdLg Risser 3 (OK Copie) | Rx du 11-07-18 : 61-34 TdLg |
| | 20-21 TdLg | 2 + | Décembre 2010 | | 30/01/2011 | 60-40 TdLg Risser 2 (OK Pas copie) | Rx du 11-07-18 : 28-69-41 Tg supTdLg |
| | 17 Lg | 4 | Mars 2008 | | 23/02/2011 | 47 TLg Risser 3 (OK Pas Copie) | Rx du 11-06-16 : 51 TLg |
| | 7-13 TdLg | 4 | Novembre 2009 | | 20/04/2011 | 26-45-31 Tg supTdLg Risser 4 (Pas Copie) | Rx du 11-06-27 : 23-45-31 Tg supTdLg |
| | | | | | | | |
| | | | | | | | |
| | 13-11-9 Tg supTdLg | 4 | Janvier 2010 | | 2010/10/07 | 41-29 TdLg Risser 3 | Rx du 11-09-22 : 53-41 TdLg |

| | | | | | | | |
|---|---|---|---|---|---|---|---|
| | | | | | | (OK Pas Copie) | |
| | 20-19 TdLg | 4 | Juin 2008 | | 28/02/2011 | 60-44 TdLg Risser 4 (OK Pas Copie) | Rx du 11-09-19 : 64-44 TdLg |
| Refuse le prélèvement post-op. | **REFUSE** | | Décembre 2009 | | 02/02/2011 | 63 Td Risser 3 | Rx du 11-06-15 : 67 Td Risser 4 |
| | 8-8 TdTLg | 5 | Juillet 2006 | | 06/07/2011 | 11-39 TdTLg Risser 4 (Ok Pas copie) | Rx du 11-10-25 : 11-38 TdTLg |
| | 17-21 Tg supTd | 5 | N/A | | 11/03/2011 | 50 Td Risser 4 + (OK Pas Copie) | Rx du 11-11-01 : 56 Td |
| | 17 Td | 4 | Octobre 2008 | | 04/07/2011 | 26-36 TdLg Tisser 4 (OK Pas copie) | Rx du 11-09-07 : 27-40 TdLg |
| Pradder Willi | 18-25 TdTLg | 5 | Juin 2007 | | 27/04/2011 | 27-50 TdTLg Risser 5 (OK Pas Copie) | Rx du 11-09-20 : 36-50 TdTLg |
| | 17-10 TdLg | 5 | Mars 2009 | | 06/05/2011 | 53 TLg Risser 4 | Rx du 11-11-21 : 60 TLg |

| | | | | | | | |
|---|---|---|---|---|---|---|---|
| | | | | | | (OK Pas Copie) | |
| | 15-20 TgTLd | 3 | NA | | 23/06/2011 | 34-70 TgTLd Risser 3 (OK Pas Copie) | Rx du 11-11-25 : 34-75 TgTLd |
| | 14 Td | 0 + | Février 2011 | | 19/05/2011 | 39-52-36 TgTdLg Risser 0 (OK Pas Copie) | Rx du 11-10-09 : 46-57-42 TgTdLg |
| | 14 Ld | 5 | Janvier 2008 | | 06/07/2011 | 26-37 Tg sup.Td Risser 4 (Ok Pas Copie) | Rx du 11-12-02 : 20-32 Tg sup.Td |
| | 16-20-9 Tg supTdTLg | 4 | Novembre 2009 | | 04/07/2011 | 32-61-46 Tg supTdTLg Risser 4 (OK Pas Copie) | Rx du 11-12-09 : 32-60-43 TgTdTLg |
| | 12-12 TdLg | 0 | Décembre 2011 | | 02/08/2011 | 32-75-45 Tg supTdLg Risser 0 (OK Pas Copie) | Rx du 11-10-17 : 40-85-48 Tg sup.TdLg |

| | | | | | | | |
|---|---|---|---|---|---|---|---|
| | 14-10 TdLg | 0 | Non | | 13/06/2011 | 31-53-32 TgTdLg Risser 0 (OK Pas Copie) | Rx du 11-12-13 : 37-61-35 TgTdLg |
| | | | | | | | |
| | 20-27-20 Tg supTdLg | 1 | Août 2012 | | 24/10/2011 | 72-71 TdLg Risser 1 (OK Pas Copie) | Rx du 12-01-11 : 33-74-70 Tg sup.TdLg |
| | 14-36 TdTLg | 5 | Avril 2007 | | 31/03/2010 | 51-46 TdTLg Risser 4 (OK Pas Copie) | Rx du 12-01-20 : 52-46 TdTLg |
| | 12-12 TdTLg | 4 | Mai 2009 | | 21/09/2011 | 16-41-49 Tg supTdTLg (OK Pas Copie) | Rx du 12-01-05 : 19-40-47 Tg supTdTLg |
| | 15-21 TdLg | 5 | Octobre 2008 | | 27/07/2011 | 68-49 TdLg Risser 4 + (OK Pas Copie) | Rx du 12-02-06 : 70-50 TdLg |
| | 16-10 Tg sup.Td | 5 | Juillet 2004 | | 06/06/2011 | 50 Td Risser Risser 5 (OK Pas Copie) | Rx du 12-03-06 : 27-53-26 Tg supTdLg |

| | | | | | | | |
|---|---|---|---|---|---|---|---|
| | 22-25 TdLg | 4 | Juin 2007 | | 12/10/2011 | 28-42 TdTLg Risser 3 (OK Pas Copie) | Rx du 12-01-18 : 30-49 TdTLg |
| | 17-22-9 Tg sup.TdLg | 3 | Octobre 2011 | | 14/07/2011 | 32-54-22 Tg sup.TdLg Risser 0 (OK Pas Copie) | Rx du 12-04-24 : 44-81-42 Tg sup.TdLg |
| | 10 Td | 4 | Janvier 2009 | | 05/10/2011 | 49 Td Risser 2 (OK Pas Copie) | Rx du 12-05-07 : 58 Td |
| | 20 TLg | 3 | Avril 2009 | | 05/10/2011 | 50-38 TdLg Risser 3 (OK Pas Copie) | Rx du 12-05-03 : 50-39 TdLg |
| Vient pour 1 fois seulement (Amie de Alexis ….. ????? Chercheur) | | | Non | | | | |
| | 23-19-25 Tg sup.TdLg | 5 | Mai 2007 | | 07/07/2009 | 51-48 TdLg Risser 4 (OK Pas Copie) | Rx du 12-05-09 : 57-48 TdLg |

| | | | | | | | |
|---|---|---|---|---|---|---|---|
| | 15-14 TdLg | 5 | Décembre 2006 | | 29/08/2011 | 26-38 TdTLg Risser Risser 0 | Rx du 12-05-01 : 31-42 TdTLg |
| Projet Marianne | 23 Td | 0 | Non | | 23/05/2012 | 60-30 TdLg Risser 0 | Rx du 12-08-08 : 65-32 TdLg |
| | 20-11 TdLg | 3 | Décembre 2009 | | 08/09/2011 | 49-43 TdTLg Risser 3 (OK Pas Copie) | Rx du 12-05-02 : 54-51 TdTLg |
| | 22-26 Tg supTd | 5 | Février 2007 | | 31/03/2011 | 37-54-28 Tg supTdLg Risser 5 (OK Pas Copie) | Rx du 12-06-18 : 39-54-29 Tg supTdLg |
| | | | Mai 2007 | | 21/03/2011 | 46-47 TdLg Risser 5 | Rx du 12-06-22 : 40-51 TdLg |
| Rx du 09-09-21 : 1 ere Visite : 49-68 TdTLg | 41-30 TdTLg | 0 | Non | | 05/04/2012 | 61-68 TdTLg Risser 0 (Pas Copie) | Rx du 12-06-11 : 61-70 TdTLg |
| Projet de Marianne | | | | | | | |
| Projet Marianne | 26-22 TdLg | 4 | N/A | | | | |
| Syndrome Loeys-Dietz | 49-48 TdLg | 5 | ??????? | | 27/02/2012 | 85-84 TdLg | Rx du 12-03-23 : 85-84 |

| | | | | | | | |
|---|---|---|---|---|---|---|---|
| + Spondylo grade 2 opéré 08-01-30 + | | | | | | Risser 5 (Ok Pas Copie) | TdLg idem |
| | 22-28 TdLg | 5 | N/A | | 16/01/2012 | 58-41 TdLg Risser 5 (OK Pas Copie) | Rx du 12-07-03 : 55-39 TdLg |
| | **REFUSE** | | Janvier 2009 | | 09/12/2009 | 51-40 TdLg Risser 4 | Rx du 12-06-12 : 60-41 TdLg |
| | 17-10 Tg sup.Td | 4 + | Décembre 2009 | | 05/12/2011 | 45-22 TdLg Risser 4 | Rx du 12-06-18 : 45-21 TdLg |
| | | | | | | | |
| Projet Marianne | | | | | | | |
| | **REFUSE** | | Avril 2008 | | 22/06/2011 | 46 Td Risser 4 | Rx du 12-06-18 : 50 Td |
| | 0 | 1 | Juin 2012 | | 12/02/2012 | 57-34 TdLg Risser 0 | Rx du 12-07-23 : 67-36 TdLg |
| | 15-20 TdLg | 5 | Mai 2009 | | 17/02/2011 | 47-48 TdLg Risser 2 (OK Pas Copie) | Rx du 12-02-27 : 47-37 TdLg |
| | 8-19 TdTLg | 4 + | Juin 2009 | | 27/01/2011 | 47-34 TdLg Risser 4 | Rx du 12-05-31 : 50-39 TdLg |

| | | | | | | | |
|---|---|---|---|---|---|---|---|
| | | | | | | (OK Pas Copie) | |
| | 16-13 Tg sup.Td | 5 | Mars 2007 | | 04/01/2010 | 26-54-35 Tg sup.TdLg Risser 4 + | Rx du 12-06-06 : 32-59-34 Tg sup.TdLg |
| | 19-12 TdLg | 5 | Décembre 2008 | | 2011/09/21 | 58-43 TdLg Risser 4 + (OK Pas Copie) | Rx du 12-09-07 : 58-48 TdLg |
| | 25-24-5 Tg supTdLg | 4 | Juillet 2011 | | 04/04/2012 | 69-57 TdLg Risser 4 | Rx du 12-08-23 : 77-57 TdLg |
| | 19-12 TdLg | 5 | Novembre 2007 | | 2010/01/22 | 50-49 TdLg Risser 5 (OK Pas Copie) | Rx du 12-09-17 : 53-49 TdLg |
| | **REFUSE** | | Novembre 2011 | | 04/04/2011 | 24-49-42 Tg sup.TdLg Risser 4 | Rx du 12-10-09 : 24-57-48 Tg sup.TdLg |
| Congénitale (Hémivertèbre L1) | 18-26 TdTLg | 0 + | Avril 2013 | | 23/08/2012 | 46 TLg Risser 0 | Rx du 12-10-29 : 48 TLg |
| | **Mère REFUSE** | | | | | | |
| | 22-21 TdLg | 5 | N/A | | 01/12/2011 | 63-54 TdLg | Rx du 12-11-05 : 66-54 |

| | | | | | | | |
|---|---|---|---|---|---|---|---|
| | | | | | | Risser 4 + | TdLg |
| | | | | | | | |
| | 25-22 TdLg | 5 | Avril 2009 | | 01/09/2011 | 22-43 TdTLg Risser 4 + | Rx du 12-06-07 : 21-44 TdTLg |
| | 23-10 TdLg | 4 + | Juin 2011 | | 01/08/2012 | 37-82-61 Tg sup.TdLg Risser 4 | Rx du 12-12-07 : 37-86-67 Tg sup.TdLg |
| | 13 Ld | 4 | Juin 2009 | | 27/09/2011 | 38 TLd Risser 3 | Rx du 12-06-11 : 39 TLd |

# ANNEXE IV

## Exemple du code source de l'application

```
<div class="form-group">
                                        @Html.LabelFor(model => model.FamilyHistory,
new { @class = "control-label col-md-4" })
                                        <div class="col-md-8">
                                            @Html.EditorFor(model =>
model.FamilyHistory)

                                            @Html.ValidationMessageFor(model =>
model.FamilyHistory)
                                        </div>
                                    </div>

                                    <div class="form-group">
                                        @Html.LabelFor(model => model.Medication, new
{ @class = "control-label col-md-4" })
                                        <div class="col-md-8">
                                            @Html.EditorFor(model => model.Medication)
                                            @Html.ValidationMessageFor(model =>
model.Medication)
                                        </div>
                                    </div>
                                    <div class="form-group">
                                        @Html.LabelFor(model => model.DomesticAnimals,
new { @class = "control-label col-md-4" })
                                        <div class="col-md-8">
                                            @Html.EditorFor(model =>
model.DomesticAnimals)
                                            @Html.ValidationMessageFor(model =>
model.DomesticAnimals)
                                        </div>
                                    </div>
                                    <div class="form-group">
                                        @Html.LabelFor(model => model.SchoolYear, new
{ @class = "control-label col-md-4" })
                                        <div class="col-md-8">
                                            @Html.EditorFor(model => model.SchoolYear)
                                            @Html.ValidationMessageFor(model =>
model.SchoolYear)
                                        </div>
                                    </div>
                                </div>
                            </div>
```

## ANNEXE V

## Validation des champs et des données par le client

### Validation des transformations

| Champ | Validation des Fonctionnalités | Validation des Données |
|---|---|---|
| **Participants** | | |
| Liste des participants | ✓ | ✓ |
| Ajouter un participant | ✓ | ✓ |
| Éditer un participant | ✓ | ✓ |
| Éditer les détails du participant | ✓ | ✓ |
| Ajouter un rendez-vous | ✓ | ✓ |
| Éditer un rendez-vous | ✓ | ✓ |
| Éditer un prélèvement du participant | ✓ | ✓ |
| **Corset** | | |
| Ajouter un corset | ✓ | ✓ |
| Éditer un corset | ✓ | ✓ |
| Supprimer un corset | ✓ | ✓ |
| **Laboratoire** | | |
| Liste des prélèvements | ✓ | ✓ |
| Éditer un prélèvement | ✓ | ✓ |
| Administration | ✓ | ✓ |
| Liste des utilisateurs | ✓ | ✓ |
| Ajouter un utilisateur | ✓ | ✓ |
| Éditer un utilisateur | ✓ | ✓ |
| List de données de l'application | ✓ | ✓ |
| Conditions | ✓ | ✓ |
| Diagnostique | ✓ | ✓ |
| État d'échantillon | ✓ | ✓ |

| | | |
|---|---|---|
| **Groupe** | | |
| Groupe ethnique | ✓ | ✓ |
| Rôle de famille | ✓ | ✓ |
| Type d'échantillon | ✓ | ✓ |
| Type de chirurgie | ✓ | ✓ |
| Type de Cobb | ✓ | ✓ |
| Type de Corset | ✓ | ✓ |
| **Changer de langue** | ✓ | ✓ |
| Français | ✓ | ✓ |
| Anglais | ✓ | ✓ |

Remarque :.....

.....

.....

.....

.....

Nom du responsable du stage : Anita Franco

Date : 2016-08-19

Signature :.....

## LISTE DE RÉFÉRENCES BIBLIOGRAPHIQUES


[1] ST-Laurent Christian, 2015, rapport de maitrise en génie, concentration technologie de l'information : *Prototype d'information de recherche clinique du laboratoire de génétique moléculaire des maladies/malformations musculosquelettiques du chu Sainte-Justine – Itération 1 (Base de données et portail web)*, École de technologie supérieur, 15p.

[2] Architectures d'intégration de données [En Ligne ]
https://depinfo.u-cergy.fr/~vodislav/Master/IED/fichiers/integration.pdf, consulté le 2016-05-27

[3] Qu'est-ce que un ETL? [En Ligne ]
http://www.geomarketing.ca/2014/11/21/quest-ce-quun-etl/,consulté le 2016-04-23

[4] ERRAKI Mehdi, Hadoop augmente la chaîne décisionnelle [En Ligne ]
https://time4bigdata.wordpress.com/, consulté le 2016-04-30

[5] Alimentation des données [En Ligne ]
http://pentaho-tutorial.blogspot.ca/2008/01/premire-tape-dune-solution-complte-de.html, consulté le 2016-06-04

[6] Purnima Bindal et Purnima Khurana, 2015, « ETL Life Cycle », (IJCSIT) International Journal of Computer Science and Information Technologies, Vol. 6 (2).

[7] SQL Server Integration Services [En Ligne ]
https://msdn.microsoft.com/en-us/library/ms141026.aspx, consulté le 2016-06-15

[8] Intégrer tous les types de données sur des plateformes mainframe et distribuées [En Ligne],http://www-03.ibm.com/software/products/fr/ibminfodata, consulté le 2016-06-17

[9] Oracle Data Integrator [En Ligne ]
http://www.oracle.com/technetwork/middleware/dataintegrator/overview/index.html,consulté le 2016-06-17

[10] RESOURCES / SAS PRODUCTS & SOLUTIONS, [En Ligne ]
http://support.sas.com/software/products/etls/, consulté le 2016-06-17

[11] Data Integration, Le moyen le plus rapide et économique pour connecter les données, [En Ligne ] ,https://fr.talend.com/products/data-integration, consulté le 2016-06-17

[12] Welcome to Scriptella ETL Project, [En Ligne ]
http://scriptella.org/, consulté le 2016-06-17

[13] KETL - Kinetic ETL , [En Ligne ]
https://www.openhub.net/p/ketl, consulté le 2016-06-17

[14] Data Integration – Kettle , [En Ligne ]
http://community.pentaho.com/projects/data-integration/, consulté le 2016-06-17

[15] JasperSoft, [En Ligne ]
http://www.open-source-guide.com/Solutions/Applications/Decisionnel-suite/Jaspersoft consulté le 2016-06-18

[16] Les ETL Open Source : Une réelle alternative aux solutions propriétaires,[En Ligne] http://business-intelligence.developpez.com/tutoriels/etl-open-source/?page=Pourquoi, consulté le 2016-06-24

[17] Magic Quadrant for Data Integration Tools [En Ligne ]
http://www.primenumerics.com/White%20Papers/Magic%20Quadrant%20for%20Data%20Integration%20Tools.pdf, consulté le 2016-07-04

[18] Ranjith Katragadda , Sreenivas Sremath Tirumala et David Nandigam,2015, «ETL tools for Data Warehousing: An empirical study of Open Source Talend Studio versus Microsoft SSIS»,Conference: ICWISCE'2015 International Conference on Web Information System and Computing Education, The 2nd World Congress on Computer Applications and Information Systems (WCCAIS'2015).