



HAL
open science

Mise en œuvre du processus qualité d'une base de données médicales : application à la base MARS du Centre Hospitalier Universitaire Grenoble-Alpes

Virginie Suize

► **To cite this version:**

Virginie Suize. Mise en œuvre du processus qualité d'une base de données médicales : application à la base MARS du Centre Hospitalier Universitaire Grenoble-Alpes. Sciences pharmaceutiques. 2018. dumas-01943458

HAL Id: dumas-01943458

<https://dumas.ccsd.cnrs.fr/dumas-01943458v1>

Submitted on 3 Dec 2018

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

AVERTISSEMENT

Ce document est le fruit d'un long travail approuvé par le jury de soutenance et mis à disposition de l'ensemble de la communauté universitaire élargie.

Il n'a pas été réévalué depuis la date de soutenance.

Il est soumis à la propriété intellectuelle de l'auteur. Ceci implique une obligation de citation et de référencement lors de l'utilisation de ce document.

D'autre part, toute contrefaçon, plagiat, reproduction illicite encourt une poursuite pénale.

Contact au SID de Grenoble :
bump-theses@univ-grenoble-alpes.fr

LIENS

Code de la Propriété Intellectuelle. articles L 122. 4

Code de la Propriété Intellectuelle. articles L 335.2- L 335.10

<http://www.cfcopies.com/juridique/droit-auteur>

<http://www.culture.gouv.fr/culture/infos-pratiques/droits/protection.htm>



UNIVERSITÉ GRENOBLE ALPES

UFR DE PHARMACIE DE GRENOBLE

Année : 2018 - 2019

MISE EN ŒUVRE DU PROCESSUS QUALITÉ D'UNE BASE DE DONNÉES MÉDICALES.
Application à la base MARS du Centre Hospitalier Universitaire Grenoble-Alpes

THÈSE
PRÉSENTÉE POUR L'OBTENTION DU TITRE DE DOCTEUR EN PHARMACIE

DIPLÔME D'ÉTAT

Virginie SUIZE

[Données à caractère personnel]

THÈSE SOUTENUE PUBLIQUEMENT À LA FACULTÉ DE PHARMACIE DE GRENOBLE

Le : 29/11/2018

DEVANT LE JURY COMPOSÉ DE

Président du jury :

Docteur Marie JOYEUX-FAURE : MCF – HP2 Inserm U1042 - EFCR CHUGA

Membres :

Docteur Sébastien BAILLY : Chargé de recherche - HP2 Inserm U1042 - EFCR CHUGA
(Directeur de thèse)

Docteur Isabelle RIEU : PAST

Professeur Renaud TAMISIER : PU-PH - HP2 Inserm U1042 - EFCR CHUGA

Docteur Olivier DOUDOU : Pharmacien industriel en Assurance qualité opérationnelle

L'UFR de Pharmacie de Grenoble n'entend donner aucune approbation ni improbation aux opinions émises dans les thèses ; ces opinions sont considérées comme propres à leurs auteurs.

Doyen de la Faculté : **M. le Pr. Michel SEVE**

Vice-doyen et Directrice des Etudes :
Mme Christine DEMEILLIERS

Année 2018 - 2019

ENSEIGNANTS - CHERCHEURS

STATUT	NOM	PRENOM	LABORATOIRE
MCF	ALDEBERT	DELPHINE	TIMC-IMAG UMR 5525 CNRS, TheREx
PU-PH	ALLENET	BENOIT	TIMC-IMAG UMR 5525 CNRS, ThEMAS
PU	BAKRI	ABDELAZIZ	TIMC-IMAG UMR 5525 CNRS
MAST	BARDET	JEAN-DIDIER	TIMC-IMAG UMR 5525 CNRS, ThEMAS
MCF	BATANDIER	CECILE	LBFA – INSERM U1055
PU-PH	BEDOUC	PIERRICK	TIMC-IMAG UMR 5525 CNRS, ThEMAS
MCF	BELAIDI-CORSAT	ELISE	HP2, Inserm U1042
MAST	BELLET	BEATRICE	-
MCF	BOUCHERLE	BENJAMIN	DPM - UMR 5063 CNRS
PU	BOUMENDJEL	AHCENE	DPM – UMR 5063 CNRS
MCF	BOURGOIN	SANDRINE	IAB – CRI INSERM U823
MCF	BRETON	JEAN	LCIB – UMR E3 CEA
MCF	BRIANCON-MARJOLLET	ANNE	HP2 – INSERM U1042
PU	BURMEISTER	WILHEM	UVHCI- UMI 3265 EMBL CNRS
MCU-PH	BUSSER	BENOIT	Institute for Advanced Biosciences, UGA / Inserm U 1209 / CNRS 5309
Professeur Emerite	CALOP	JEAN	
MCF	CAVAILLES	PIERRE	TIMC-IMAG UMR 5525 CNRS
MCU-PH	CHANOINE	SEBASTIEN	CR UGA - INSERM U1209 - CNRS 5309
MCF	CHOISNARD	LUC	DPM – UMR 5063 CNRS
AHU	CHOVELON	BENOIT	DPM – UMR 5063 CNRS
PU-PH	CORNET	MURIEL	TIMC-IMAG UMR 5525 CNRS, TheREx
Professeur Emérite	DANEL	VINCENT	-
PU	DECOUT	JEAN-LUC	DPM – UMR 5063 CNRS
MCF Emérite	DELETRAZ-DELPORTE	MARTINE	LPSS – EAM 4129 LYON
MCF	DEMEILLERS	CHRISTINE	TIMC-IMAG UMR 5525 CNRS
PU-PH	DROUET	CHRISTIAN	GREPI EA7408

Mise à jour le 31/08/2018

PU	DROUET	EMMANUEL	IBS – UMR 5075 CEA CNRS HIV & virus persistants Institut de Biologie Structurale
MCF	DURMORT	CLAIRE	IBS – UMR 5075 CEA CNRS
PU-PH	FAURE	PATRICE	HP2 – INSERM U1042
MCF	FAURE-JOYEUX	MARIE	HP2 – INSERM U1042
PRCE	FITE	ANDREE	-
MCU-PH	GARNAUD	CECILE	TIMC-IMAG UMR 5525 CNRS, TheReX
PRAG	GAUCHARD	PIERRE-ALEXIS	-
MCU-PH	GERMI	RAPHAELE	IBS – UMR 5075 CEA CNRS HIV & virus persistants Institut de Biologie Structurale
MCF	GEZE	ANNABELLE	DPM – UMR 5063 CNRS
MCF	GILLY	CATHERINE	DPM – UMR 5063 CNRS
PU	GODIN-RIBUOT	DIANE	HP2 – INSERM U1042
Professeure Émérite	GRILLOT	RENEE	-
MCF Émérite	GROSSET	CATHERINE	DPM – UMR 5063 CNRS
MCF	GUIEU	VALERIE	DPM – UMR 5063 CNRS
AHU	HENNEBIQUE	AURELIE	TIMC-IMAG UMR 5525 CNRS, TheReX
MCF	HININGER-FAVIER	ISABELLE	LBFA – INSERM U1055
MCF	KHALEF	NAWEL	TIMC-IMAG UMR 5525 CNRS
MCF	KOTZKI	SYLVAIN	HP2 – UMR S1042
MCF	KRIVOBOK	SERGE	DPM – UMR 5063 CNRS
PU	LENORMAND	JEAN-LUC	TIMC-IMAG UMR 5525 CNRS, TheReX
PU	MARTIN	DONALD	TIMC-IMAG UMR 5525 CNRS
PRCE	MATTHYS	LAURENCE	-
AHU	MAZET	ROSELINE	DPM – UMR 5063 CNRS
MCF	MELO DI LIMA	CHRISTELLE	LECA – UMR CNRS 5553
AHU	MINOVES	MELANIE	HP2 – INSERM U1042
PU	MOINARD	CHRISTOPHE	LBFA - INSERM U1055
PU-PH	MOSSUZ	PASCAL	IAB – INSERM U1209
MCF	MOUHAMADOU	BELLO	LECA – UMR 5553 CNRS
MCF	NICOLLE	EDWIGE	DPM – UMR 5063 CNRS
MCF	OUKACINE	FARID	DPM – UMR 5063 CNRS
MCF	PERES	BASILE	DPM – UMR 5063 CNRS
MCF	PEUCHMAUR	MARINE	DPM – UMR 5063 CNRS
PU	PEYRIN	ERIC	DPM – UMR 5063 CNRS
AHU	PLUCHART	HELENE	TIMC-IMAG – UMR 5525 CNRS, TheMAS
MCF	RACHIDI	WALID	LCIB – UMR E3 CEA
MCF	RAVELET	CORINNE	DPM – UMR 5063 CNRS

Mise à jour le 31/08/2018

	PU	RIBUOT	CHRISTOPHE	HP2 – INSERM U1042	
	PAST	RIEU	ISABELLE	-	
-----	Professeure Émérite	ROUSSEL	ANNE-MARIE		-----
-----	PU-PH	SEVE	MICHEL	LBFA – INSERM U1055	-----
-----	MCF	SOUARD	FLORENCE	DPM – UMR 5063 CNRS	-----
-----	MCF	SPANO	MONIQUE	IBS – UMR 5075 CEA CNRS	---
AHU : Assistant Hospitalo- Universitaire	MCF	TARBOURIECH	NICOLAS	IBS – UMR 5075 CEA CNRS	
ATER : Attachés Temporaires d'Enseignement et de Recherches	MCF	VANHAVERBEKE	CECILE	DPM – UMR 5063 CNRS	
BCI : Biologie du Cancer et de l'Infection	PU	WOUESSIDDJEWE	DENIS	DPM – UMR 5063 CNRS	

AHU : Assistant Hospitalo-Universitaire
ATER : Attachés Temporaires d'Enseignement et de Recherches
BCI : Biologie du Cancer et de l'Infection
CHU : Centre Hospitalier Universitaire
CIB : Centre d'Innovation en Biologie
CRI : Centre de Recherche INSERM
CNRS : Centre National de Recherche Scientifique
DCE : Doctorants Contractuels Enseignement
DPM : Département de Pharmacochimie Moléculaire
HP2 : Hypoxie Physiopathologie Respiratoire et Cardiovasculaire
IAB : Institute for Advanced Biosciences
IBS : Institut de Biologie Structurale
LAPM : Laboratoire Adaptation et Pathogenèse des Microorganismes
LBFA : Laboratoire Bioénergétique Fondamentale et Appliquée
LCBM : Laboratoire Chimie et Biologie des Métaux
LCIB : Laboratoire de Chimie Inorganique et Biologie
LECA : Laboratoire d'Ecologie Alpine
LPSS : Laboratoire Parcours Santé Systémique
LR : Laboratoire des Radio pharmaceutiques
MAST : Maître de Conférences Associé à Temps Partiel
MCF : Maître de Conférences des Universités
MCU-PH : Maître de Conférences des Universités et Praticiens Hospitaliers
PAST : Professeur Associé à Temps Partiel
PRAG : Professeur Agrégé
PRCE : Professeur certifié affecté dans l'enseignement
PU : Professeur des Universités
PU-PH : Professeur des Universités et Praticiens Hospitaliers
SYMMES : Systèmes Moléculaires et nanoMatériaux pour l'Energie et la Santé
TIMC-IMAG : Laboratoire Technique de l'Imagerie, de la Modélisation
UMR : Unité Mixte de Recherche
UVHCI : Unit of Virus Host Cell Interactions

Mise à jour le 31/08/2018

REMERCIEMENTS

Je voudrais tout d'abord remercier grandement mon directeur de thèse, Sébastien Bailly, de m'avoir fait confiance pour travailler avec lui sur ce sujet de thèse. Ce fut un plaisir d'avoir travaillé à ses cotés car outre son appui technique, il a toujours été là pour me conseiller et me soutenir lors de l'élaboration de cette thèse.

Je souhaite également remercier Marie Joyeux-Faure, qui a accepté sans hésitation d'être la présidente de mon Jury.

Je tiens à remercier Renaud Tamisier qui a tout d'abord accepté de me faire travailler sur la base MARS mais aussi d'être dans mon jury de thèse.

Je remercie très sincèrement Isabelle Rieu, membre de mon jury de thèse, pour son soutien, ses conseils et son implication envers mon projet professionnel tout au long de mes études.

Je remercie également Olivier Doudou, ancien collègue de travail qui a su me conseiller, me suivre durant ces deux dernières années et qui me fait l'honneur d'intégrer mon jury de thèse.

Je tiens à remercier tous les utilisateurs de la base MARS qui ont eu la gentillesse de m'accorder du temps et de faire avancer le projet grâce à leur expertise. Ainsi que Fabien Yonnet, qui a supporté toute la partie statistique de ma thèse.

Je remercie l'ensemble de mes ami(e)s avec qui j'ai passé ces 5 belles années riches en souvenirs. Mais aussi mes collègues du Master PIF2P de Grenoble qui ont su m'encourager dans cette dernière ligne droite.

Je remercie du fond du cœur mes parents, sans qui mes études n'auraient pas été réalisables. Ils ont toujours su m'encourager et me soutenir dans ce que j'entreprends.

Je remercie également l'ensemble de ma famille qui m'a accompagné pendant tout mon parcours universitaire.

Et pour finir, je ne remercierai jamais assez Ludovic, mon petit ami, qui a su m'épauler et m'accompagner depuis toutes ces années. Un soutien hors-pair, merci pour tout.

TABLE DES MATIERES

I.	Introduction.....	15
II.	Contexte	21
	A. Le contexte règlementaire	21
	B. La qualité des données.....	25
	1. Généralités sur la qualité des données.....	25
	2. Mise en œuvre de la qualité des données	27
III.	La base MARS	29
	A. La collecte des informations.....	31
	B. Les informations collectées	32
	C. La structure de la base MARS	34
	1. Support informatique	34
	2. L'accès à la base	35
	3. Les fonctionnalités principales de la base MARS	36
	4. Les fonctionnalités secondaires de la base MARS	40
	D. La conformité de MARS au contexte règlementaire.....	41
	E. La qualité des données de la base MARS	43
IV.	La démarche	45
	A. Le questionnaire	45
	B. Les personnes rencontrées.....	45
V.	Les résultats de l'investigation et les actions mises en place	47
	1. Utilisation générale et les avantages/inconvénients de la base	47
	2. Problèmes de saisie des données	48
	3. Les souhaits d'évolution de la base par les utilisateurs	70

4.	Zoom sur les attentes des administrateurs : Mise en place d'indicateurs et de rapport d'activité	79
VI.	Perspectives d'évolution	85
A.	Sur la formation des utilisateurs.....	85
B.	Sur le suivi et l'adhésion des nouveaux centres adhérents	85
C.	Sur la satisfaction des utilisateurs	86
D.	Sur les problèmes soulevés lors de notre questionnaire	87
E.	Sur l'utilisation par des centres étrangers	87
F.	Sur les outils de reporting et rapport d'activité	88
G.	Sur les données d'examens	88
VII.	Conclusion	90
VIII.	Bibliographie.....	93
IX.	Annexes	98

LISTE DES FIGURES

Figure 1 - Les démarches règlementaires en fonction du type d'étude.....	22
Figure 2 - Les principaux changements suite à la mise en vigueur du RGPD	25
Figure 3 - Interaction des différents professionnels sur une base de données	26
Figure 4 - Utilisateurs de la base MARS.....	30
Figure 5 - Données à saisir dans les visites dans l'évaluation de la « fiche patient ».....	33
Figure 6 - Modalités de saisie possible dans la base MARS.....	34
Figure 7 - Organisation de la « fiche patient » - Onglet "informations générales"	37
Figure 8 - Organisation de la « fiche patient » - Autres onglets	38
Figure 9 - Alerte proposée lorsqu'un patient peut être inclus dans une étude	39
Figure 10 – Outils de reporting de la base	41
Figure 11 - Problèmes rencontrés sur les "Antécédents personnels" et les "Evènements médicaux (CIM-10)"	50
Figure 12 - Problèmes rencontrés sur la partie "Traitements"	60
Figure 13 - Arbre décisionnels des cohérences/incohérences sur l'hypertension artérielle.....	64
Figure 14 - Incohérences retrouvées et actions mises en place.....	66
Figure 15 - Alerte s'affichant à l'écran quand une incohérence HTA/traitement est détectée	67
Figure 16 - Les attentes des utilisateurs concernant la base MARS	72
Figure 17 - Etape de rédaction et d'approbation des procédures	74
Figure 18 - Répartition de l'âge des patients selon le sexe	77

Figure 19 - Nombre de patients et d'évaluations générales enregistré par mois77

LISTE DES TABLEAUX

Tableau I - Profils utilisateurs de MARS et actions associées.....	35
Tableau II - Modification des termes sur les Antécédents/événements médicaux	52
Tableau III - Items CIM-10 proposés pour qualifier un syndrome néphritique aigu	53
Tableau IV - Diversité générée par les champs libres.....	58
Tableau V - Spécialités existantes en paracétamol en France et au Canada.....	68
Tableau VI - Attentes des utilisateurs concernant les outils de reporting	71

LISTE DES ANNEXES

Annexe 1 - Questionnaire d'état des lieux de la base MARS.....	99
Annexe 2 - Réponses des utilisateurs sur les avantages et inconvénients de la base MARS.....	101
Annexe 3 - Réponses des utilisateurs concernant les problèmes d'encodage traité lors d'investigations futures.....	103
Annexe 4 – Extrait du dictionnaire CIM-10 d'aide à la saisie des évènements généraux les plus courants.....	105
Annexe 5 - Processus d'inclusion de patient dans les protocoles de recherche du CHUGA.....	107
Annexe 6 - Processus d'inclusion des patients dans la base MARS venu pour une consultation.....	108
Annexe 7 – Modèle proposé pour les procédures base MARS CHUGA.....	109
Annexe 8 – Modèle proposé pour le formulaire "d'attestation de formation" à la procédure.....	112
Annexe 9 – Modèle proposé pour le quizz d'attestation de formation	113
Annexe 10 - Modèle proposé pour le formulaire "d'état de formation des utilisateurs à la base MARS"	115
Annexe 11 - Logigramme données manquantes.....	116
Annexe 12 - Logigramme données aberrantes	117
Annexe 13 - Récapitulatif des actions réalisées et futures.....	118

ABREVIATIONS :

ANSM : Agence nationale de sécurité du médicament et des produits de santé

ARC : Attaché de recherche clinique

ATC : Classification anatomique, thérapeutique et chimique

ATCD : Antécédent

CCTIRS : Comité consultatif sur le traitement de l'information en matière de recherche dans le domaine de la santé

CHAL : Centre hospitalier Alpes Léman

CHUGA : Centre hospitalier universitaire Grenoble – Alpes

CIM-10 : Classification internationale des maladie 10^e révision

CNIL : Commission nationale de l'informatique et des libertés

CPP : Comité de protection des personnes

DCI : Dénomination Commune Internationale

DM : Donnée manquante

EFCR : Service des explorations fonctionnelles cardiorespiratoire

EPIC : Cohorte prospective européenne sur le cancer et la nutrition

ESADA : European sleep apnea database

FSAm-NPS : Nutrient profiling system of the british food Standards agency

HTA : Hypertension artérielle

MARS : Multimorbidity apnea respiratory failure sleep

MR : Méthodologie de référence

OMS : Organisation mondiale de la santé

ORL : Oto-rhino-laryngologie

OSFP : Observatoire sommeil de la fédération de pneumologie

RGPD : règlement général sur la protection des données

RIPH : Recherche impliquant des personnes humaines

SAS : Syndrome d'apnée du sommeil

SNDS : Système national des données de santé

I. Introduction

Le volume de données collectées suit une évolution exponentielle depuis plusieurs années. En 2018, on estime à 2,5 quintillions le volume d'information créée chaque jour. Par ailleurs, 90 % des données disponibles dans le monde ont été créées dans les deux dernières années, soit entre 2016 et 2018 ¹. Cette tendance se poursuivra dans les prochaines décennies ².

La recherche médicale est un des secteurs qui génère un volume important et une grande diversité de données. Les données médicales, auparavant disponibles uniquement sur papier, sont désormais collectées en permanence au chevet du patient ³, sous format numérique ou papier. Ces données peuvent être, par exemple, des données démographiques, des observations cliniques, des signaux physiologiques ou des résultats d'exams médicaux.

Cette évolution numérique est liée à un changement de la pratique médicale ⁴. En effet, de nos jours, les examens médicaux et le recueil des données en résultant peuvent être automatisés ce qui augmente considérablement le nombre de données collectées et stockées ⁴.

De plus les données médicales peuvent être complétées par un ensemble de données non structurées telles que des informations de diagnostic, des annotations provenant du personnel médical en texte libre, des fichiers texte, audio ou vidéo, etc.

L'évolution de la génération des données est accompagnée d'une évolution des méthodes d'analyse des données avec les approches dites de « Big Data ». Parmi ces approches, certaines visent à se rapprocher des conditions d'un essai clinique

randomisé. En effet, en recherche clinique interventionnelle, les essais cliniques randomisés sont considérés comme le gold standard. Ils sont mis en place pour démontrer l'efficacité ou la sécurité d'un traitement. De plus, c'est le seul moyen communément admis pour démontrer une relation causale^{5,6}.

Cependant, ils ne sont pas dénués de problèmes : ils sont difficiles à entreprendre, ils exigent souvent une grande taille d'échantillon et sont généralement assez coûteux. De plus, ils ont une faible validité externe et une application parfois limitée à la pratique clinique⁷.

Les études réalisées sur les données observationnelles, elles, ne respectent pas les conditions de randomisation propres à l'essai clinique. Cependant, si ces données ne permettent pas de conclure à des liens de causalité, les études conduites sur les bases de données, « en conditions réelles d'utilisation », apportent des informations essentielles après la commercialisation du médicament⁵. Elles permettent de recueillir différents types d'informations sur^{5,8,9} :

- La maladie
- La manière de la traiter
- Les effets-indésirables de certain traitement mais aussi les effets indésirables rares
- La population générale, etc.

De plus, plusieurs études comparatives ont montré que l'utilisation des données observationnelles peut être utilisée en complément des essais cliniques randomisés⁸.

Par exemple, parmi les méthodes récentes, certaines telles que les modèles structurels

marginiaux proposés par Robins *et al.* en 2000¹⁰, permettent d'estimer les relations de causalité possibles à partir de données observationnelles non randomisées.

Dans la littérature médicale, il y a de nombreux exemples de recherche clinique réalisée à partir de données observationnelles avec différents types de données et différents objectifs. La base OutcomeREA est une cohorte prospective démarrée en 1997 pour recueillir les données cliniques des patients en réanimation à l'entrée dans le service et tout au long de leur séjour. Actuellement cette base compte plus de 22 000 patients et plus de 200 000 journées patients. A partir des données de cette base de haute qualité il a été par exemple possible de mettre en évidence qu'un traitement empirique antifongique n'améliore pas le pronostic des patients non neutropéniques sévères en réanimation¹¹. Cette hypothèse a été confirmée par la suite avec l'essai clinique EMPIRICUS¹², montrant ainsi la complémentarité entre recherche sur données observationnelles et données expérimentales.

Dans un autre registre, Les études de cohortes multinationales telles que la cohorte prospective européenne sur le cancer et la nutrition (EPIC) qui comprend 471 495 adultes suivi entre 1993 et 2014, mettent en évidence des problèmes de santé publique. Récemment, il a été montré que la consommation de produits alimentaires avec un score nutrient profiling system of the British food standards agency (FSAm-NPS) élevé (qualité nutritionnelle inférieure) était associée à un risque plus important de cancer confirmant ainsi la nécessité d'informer les consommateurs sur la qualité nutritionnelle des produits alimentaires¹³.

Enfin, la mise à disposition des bases médico-administratives comme le système national des données de santé (SNDS) en France constitue une avancée considérable pour analyser et améliorer la santé de la population. Cela permet de disposer de l'ensemble des informations concernant la consommation des patients en soins de ville et à l'hôpital, mais aussi des données sur le décès et sur les complémentaires santé, et des données sociales telles que le handicap. Ces données peuvent apporter des éclairages importants sur les trajectoires de soins des patients. La fusion des bases de données cliniques avec le SNDS qui est rendue possible, permet d'aller plus loin dans la compréhension des trajectoires de soins des patients ¹⁴. A cela peuvent s'ajouter d'autres bases de données disponibles sur l'environnement des patients comme le pollen, la pollution atmosphérique, les conditions météorologiques, les indicateurs sociaux qui permettent d'enrichir la compréhension des pathologies.

Ces évolutions ont impliqué des changements réglementaires pour assurer la sécurité des données personnelles et limiter les abus possibles sur de telles données.

Pour cela, en France une évolution réglementaire a été mise en place avec la Loi Jardé qui vise à encadrer plus précisément la recherche médicale avec un volet concernant les données de santé. En parallèle, l'Union Européenne a revu sa directive sur la protection des données personnelles et l'a remplacé par une loi plus robuste : le règlement général sur la protection des données (RGPD). Cette nouvelle directive tend à mieux faire appliquer/respecter la protection des données personnelles en tant que droit fondamental de l'Homme.

Dans ce contexte, afin de constituer une base de données médicale pour le suivi prospectif des patients consultants pour un trouble du sommeil ou une pathologie respiratoire, la base de données Multimorbidity apnea respiratory failure sleep database (MARS) a été mise en place dans le service des explorations fonctionnelles cardiorespiratoires (EFCR) du centre hospitalier Grenoble- Alpes (CHUGA) au sein du laboratoire HP2, Inserm 1042 (Université Grenoble Alpes) en 2013.

Le syndrome d'apnée du sommeil (SAS), qui est la principale pathologie explorée dans le service et pour laquelle la base MARS a été constituée, correspond à des collapsus des voies aériennes supérieures, complets ou incomplets, survenant de manière répétée au cours du sommeil et qui ont quatre conséquences principales : La survenue de séquences de désaturation-réoxygénation, des épisodes transitoires d'hypercapnie, des efforts respiratoires augmentés, et une survenue de micro-éveils terminant les événements respiratoires ¹⁵⁻¹⁷. En France, 4,9% de la population générale ont déjà déclaré avoir des symptômes évocateurs d'un SAS ¹⁸.

Les conséquences cliniques d'un SAS sont la somnolence diurne, un sommeil perçu comme non réparateur, des céphalées matinales, une nycturie, et parfois des troubles de la libido, un syndrome dépressif et des troubles mnésiques ¹⁹. La somnolence diurne est associée à un nombre non négligeable d'accident de la route ^{19,20}. De plus, le SAS est un facteur de risque de survenue de maladies métaboliques et cardiovasculaires ^{15-17,19}. Il existe donc un réel enjeu de suivre prospectivement ces patients.

Le SAS est donc une pathologie multifactorielle avec des conséquences multiples et qui nécessite de recueillir de nombreuses informations.

Face à la multiplication des différentes sources d'informations, l'objectif principal est de maintenir un niveau de qualité élevé des données afin de garantir une recherche de qualité.

L'objectif de cette thèse est de mettre en œuvre un processus de qualité sur la base MARS pour contrôler, améliorer et vérifier la qualité des données saisies par les utilisateurs. Dans un premier temps nous parlerons du contexte réglementaire et qualitatif des bases de données. Pour remettre tout ceci en contexte, une présentation de la base MARS sera réalisée.

Ensuite, nous expliquerons la démarche réalisée dans cette thèse et présenterons le questionnaire associé. Un état des lieux sur les résultats et les actions mises en place sera réalisé et nous finirons avec les perspectives d'évolution de la base MARS.

II. Contexte

A. Le contexte réglementaire

La recherche clinique regroupe les études scientifiques pratiquées sur l'humain dans le domaine de la santé. Ces recherches sont indispensables pour le milieu scientifique médical car elles permettent, in fine, d'améliorer les connaissances médicales et biologiques ainsi que la prise en charge des pathologies étudiées, avec comme finalité, l'amélioration de notre santé^{21,22}.

La loi Jardé entrée en vigueur le 18 novembre 2016, définit trois catégories de recherche impliquant des personnes humaines (RIPH)^{23,24} :

- Catégorie 1 : Recherche interventionnelle comportant une intervention sur le patient, non justifiée par sa prise en charge habituelle.
- Catégorie 2 : Recherche interventionnelle à risques et contraintes minimales dont la liste est fixée par un arrêté ministériel.
- Catégorie 3 : Recherche non interventionnelle ou dite observationnelle, dans lesquelles tous les actes sont pratiqués et les produits sont utilisés de manière habituelle, sans procédure supplémentaire ou inhabituelle de diagnostic, de traitement ou de surveillance.

Ces différentes catégories permettent de répondre aux principaux axes de la recherche clinique.

En plus de définir 3 catégories de RIPH, la loi Jardé, permet de fixer le cadre réglementaire autour des études (Figure 1).

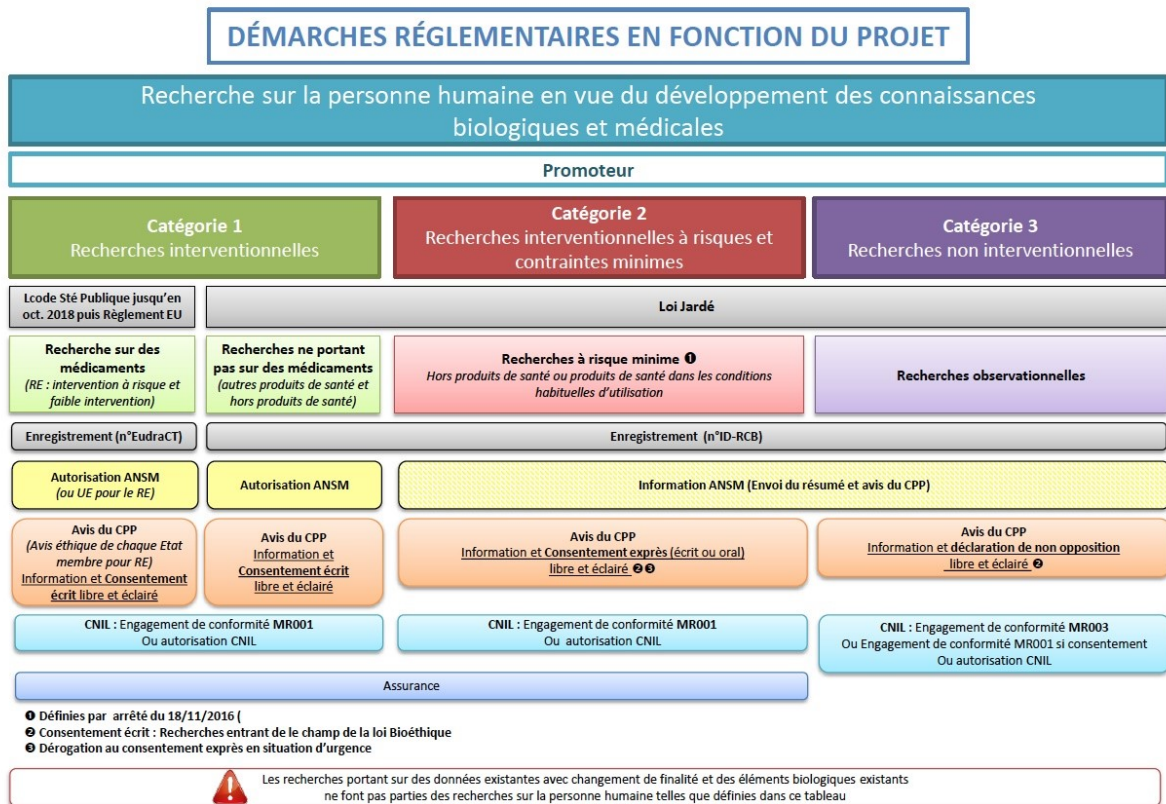


Figure 1 - Les démarches règlementaires en fonction du type d'étude

Les recherches de catégories 1 et 2 sont spécifiques et ne font pas l'objet de cette thèse. Une focalisation sera faite sur la catégorie 3, les recherches non interventionnelles observationnelles.

Pour une recherche observationnelle, différentes modalités règlementaires doivent être respectées :

- L'étude doit être enregistrée sous un numéro d'enregistrement de la recherche. Ce numéro aussi appelé « numéro ID-RCB » permet d'identifier

l'ensemble des recherches et est attribué par l'Agence nationale de sécurité du médicament et des produits de santé (ANSM)^{24,25}.

- L'ANSM doit être informée via l'envoi du résumé et l'avis du Comité de Protection des personnes (CPP)^{23,24}.
- L'avis du CPP est requis²⁶.
- Un engagement de conformité doit être délivré par La Commission nationale informatique et libertés (CNIL). Dans le cas d'une étude observationnelle, il s'agit d'un engagement de conformité Méthodologies de Référence (MR)-003^{23,24,27,28}.
- Chaque étude doit avoir un promoteur responsable des données et de leurs traitements^{23,24,26}.

Outre l'avis du CPP et de l'ANSM sur la recherche, le patient participant à l'étude doit recevoir une information claire et exhaustive au sujet de l'étude (objectif, méthodologie, résultats, etc.)²³. Cette information est délivrée par l'investigateur, le médecin ou une personne le représentant²³. Suite à la réception de ces informations, le patient doit simplement faire une déclaration de non opposition à l'étude pour pouvoir être inclus^{23,24}.

La CNIL a mis en place des MR qui encadrent le traitement de données de santé pour les trois catégories d'études définies dans le cadre de la loi Jardé²⁸. L'objectif de ces MR est de faciliter les procédures d'autorisation délivrées par la CNIL pour le traitement de données de santé qui n'identifient pas les patients. En outre les informations relatives à l'identité d'un patient directes ou indirectes, définies comme étant des données sensibles, ne doivent pas être enregistrées dans une base de

données ayant un objectif de recherche. Cependant, les initiales ou les deux premières lettres du nom peuvent tout de même être utilisées, le tout accompagné d'un numéro d'identifiant.

A ce niveau réglementaire national est venu s'ajouter une réglementation complémentaire issue de l'évolution du contexte réglementaire européen : le RGPD. Il s'appuie sur une responsabilisation des acteurs traitant des données ²⁹. Les responsables des données doivent mettre en place un certain nombre de mesures techniques et organisationnelles nécessaires au respect de la protection des données personnelles. Et ce, dès la conception du produit/service mais aussi de façon continue. Ceci implique donc que le responsable doit être en mesure de démontrer la conformité de leurs systèmes à chaque instant.

Le renforcement du contrôle des données personnelles par le citoyen est le but premier du RGPD, en se basant sur une relation de confiance et de transparence avec le responsable du traitement, ce qui constitue une plus grande souplesse. Cependant, cette souplesse est compensée par une amplification des sanctions en cas de négligence allant jusqu'à une amende de 4% du chiffre d'affaire annuel d'une entreprise ³⁰. Enfin, le RGPD prévoit une période de 3 ans pour se mettre aux normes afin de permettre aux responsables de faire face à certains termes de la loi qui restent encore flous face à des cas particuliers ³¹.

Stéphane ASTIER et Jean Philippe SOUYRIS, décrivent les principaux changements engendrés par la mise en place du RGPD ²⁹, ceux-ci étant présentés dans la Figure 2.

Obligations maintenues	Obligations renforcées	Les nouvelles obligations
<ul style="list-style-type: none"> • Respect du principe de finalité et de proportionnalité (minimisation) • Durée de conservation des données limitée • Information des personnes obligatoire 	<ul style="list-style-type: none"> • Le consentement des personnes doit être recueilli • La transmission de données doit être encadrée • La sécurité du traitement doit être assurée • Le droit des personnes doit être garanti 	<ul style="list-style-type: none"> • Les nouveaux droits doivent être garantis : Droit à l'effacement - Droit à la limitation du traitement - Droit à la portabilité des données • Une notification aux autorités et à la personne concernée doit être réalisée dans les 72 heures, lors d'une faille de sécurité • Une étude d'impact devra être réalisée • Un registre des activités de traitement doit être établi • Un data protection officer doit être désigné

Figure 2 - Les principaux changements suite à la mise en vigueur du RGPD

B. La qualité des données

1. Généralités sur la qualité des données

Pour participer au bon fonctionnement des bases de données, trois corps de métiers interagissent avec des bases de données (Figure 3). Des informaticiens s'occupant de l'interface de stockage ; des professionnels de santé (médecins, attachés de recherche clinique) ayant pour but d'enrichir la base avec des données provenant des patients et pour finir des biostatisticiens en charge de l'analyse des données de la base. L'interaction et le travail interdisciplinaire de ces trois professions sont impératifs pour garantir la qualité des données ^{32,33}.



Figure 3 - Interaction des différents professionnels sur une base de données

Cependant, malgré l'interaction pluridisciplinaire autour de cet outil, les bases de données, qu'elles soient médicales ou non peuvent être affectées par plusieurs sources d'incertitudes. Erreurs de mesure, fautes d'orthographe, abréviations, codes incorrects, données manquantes, doublons sont des sources d'altération de la qualité qui peuvent avoir des conséquences³²⁻³⁴.

Pour que ces études et les résultats soient fiables et de qualité, il est nécessaire que les données saisies soit correctes et de qualité³³.

La qualité des données n'est pas seulement réduite à une bonne saisie des données dans leurs systèmes d'information, elle permet également de gérer des données incorrectes et dupliquées³⁴.

Pour cela il est nécessaire, au préalable, de déterminer les variables à recueillir mais aussi la fréquence de recueil. Ces variables doivent être fiables, reproductibles et définies de façon précise pour que leur interprétation soit la même pour l'ensemble des investigateurs. Ces variables peuvent être saisies automatiquement par un moniteur via un import de données ou saisie manuellement par le personnel³³.

Le support informatique permettant de gérer une base de données est vivement conseillé car il permet d'optimiser l'ergonomie, de créer des alertes ou des contrôles multiples permettant de vérifier la cohérence des données. Le logiciel doit être capable de vérifier par exemple la cohérence des dates, la validité des variables quantitatives entrées comme la taille ou les valeurs biologiques par exemple, ou encore d'afficher des alertes en cas de création d'un doublon, de donner la liste de dossier incomplets, etc. ³³⁻³⁵.

Le personnel habilité à utiliser la base de données doit participer à des séances de formation, manipuler régulièrement l'outil de saisie et être informé de ses évolutions. La motivation des administrateurs est primordiale pour assurer l'animation et la mise en œuvre du système qualité ³³.

2. Mise en œuvre de la qualité des données

Avant l'enregistrement et l'archivage des données plusieurs étapes de contrôle sont nécessaires pour vérifier d'éventuelles incohérences et pouvoir les modifier en temps réel avant l'enregistrement ou l'archivage ^{33,34}.

La gestion de la qualité des données repose sur des pratiques et techniques permettant de maintenir la véracité des données via la gestion de 6 critères ³⁶ :

- La complétude : Avoir l'ensemble des informations dont on a besoin sans données manquantes.
- La validité : L'ensemble des données à disposition doivent être valides et justifiées.

- La précision : Les données saisies doivent satisfaire le degré de précision demandé.
- La cohérence : L'ensemble des données saisies ne doivent pas être incohérentes.
- La disponibilité : Dès que les données sont disponibles il est essentiel de les exporter et de les rendre visibles pour que les utilisateurs puissent les exploiter.
- L'actualisation : Permet d'avoir les données les plus récentes. Cependant les mises à jour de données suite à des ajouts, correction ou suppression doivent être tracées.

III. La base MARS

La base de données MARS, pour Multimorbidity, Apnea, Respiratory failure, Sleep a été mise en place en 2013 par le Professeur Tamisier, pour permettre le suivi prospectif à des fins de recherche des patients consultant au service EFCR du CHUGA. L'objectif principal est de fournir un outil sous la forme d'une base de données exhaustive permettant un phénotypage commun des patients explorés dans le cadre du soin courant du sommeil et pathologie respiratoire nocturne et de leur suivi prospectif. Les patients sont inclus dans la base selon deux mécanismes : soit à partir des consultations courantes soit du fait de l'inclusion dans une étude clinique. Dans les deux cas, le consentement du patient est recueilli.

Cette base permet de répondre également à deux objectifs secondaires :

- Evaluation des pratiques cliniques,
- Evaluation des soins selon le type de pathologie et/ou selon le type de prise en charge clinique.

Les populations concernées par cette base sont :

- Les patients majeurs venant consulter pour un trouble du sommeil ou une pathologie respiratoire nocturne quelle que soit la sévérité et le type de traitement mis en place.
- Les patients ayant signé la notice d'information concernant l'autorisation de l'informatisation des données (selon les recommandations de la CNIL). Les patients inaptes à comprendre la notice d'information ne peuvent donc pas être inclus.

La base MARS, comme décrite dans la Figure 4, est administrée et alimentée par un nombre important d'utilisateurs :

- Les administrateurs qui assurent la gestion de la base, suivent l'évolution de la structure, assurent la bonne adhésion des différents utilisateurs, et valident les exports à réaliser pour les analyses demandées.
- L'informaticien qui gère l'architecture et l'évolution fonctionnelle de la base.
- Les data-managers et statisticiens qui permettent de vérifier la qualité des données, leur mise en forme pour des analyses et l'analyse des données à proprement parler.
- Une fonction d'utilisateur.

Actuellement 337 comptes utilisateurs ont été créés avec différents niveaux d'accès à la base de données.

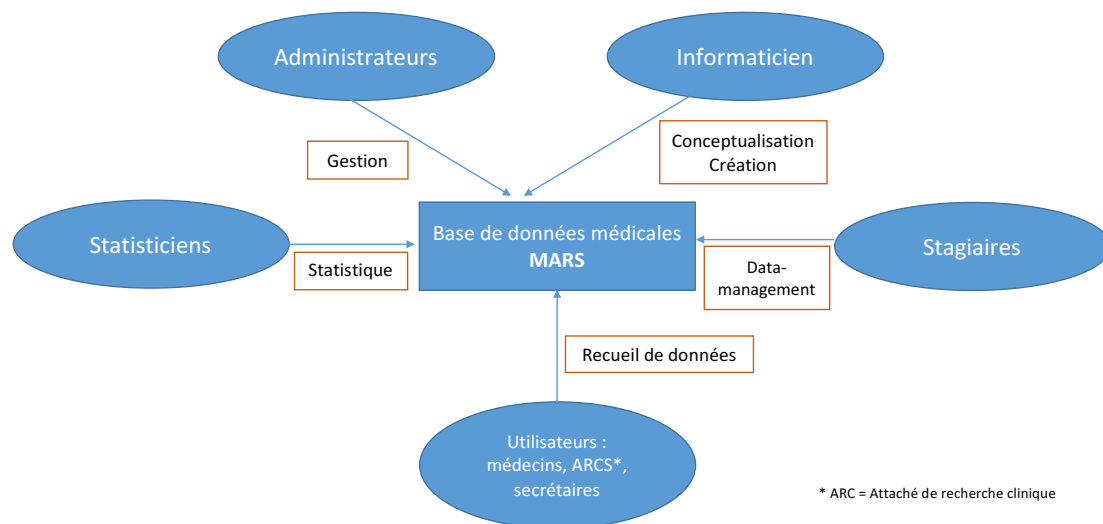


Figure 4 - Utilisateurs de la base MARS

Initialement monocentrique, l'outil a été diffusé à d'autres centres en France (Findrol, Toulouse, Lyon, Montpellier, etc.) et à l'étranger (Genève, Québec, Portugal, Belgique).

A. La collecte des informations

La collecte des informations peut se faire soit manuellement soit numériquement.

Manuellement, le patient va remplir sous format papier un questionnaire de l'évaluation générale demandant la totalité des informations présentes sur la base, lors de son arrivée dans l'établissement de consultation. Ces données seront ensuite reprises par le médecin lors de la consultation et saisies de façon manuelle dans la base par ce dernier.

Il est aussi possible d'utiliser un format numérique qui peut se présenter sous deux formes :

- Un lien informatique peut être envoyé au patient en amont de la consultation qui lui permet de remplir le questionnaire depuis son domicile.
- Une saisie sur tablette qui est fournie au patient sur place avant la consultation.

Les réponses numériques sont ensuite validées par le médecin lors de la consultation.

A la différence du questionnaire papier, les données numériques sont directement importées dans la base MARS dans la « fiche patient » préalablement créée. Cela permet de limiter la saisie manuelle qui peut être à l'origine d'erreurs de saisie.

La méthode de collecte varie entre les centres en fonction de l'organisation et des moyens financiers à disposition.

B. Les informations collectées

Les informations principales collectées pour alimenter la « fiche patient » sont dans la fiche « Evaluation générale » de l'onglet « Informations générales » après création de l'identifiant (Figure 5).

Les variables recueillies permettent de balayer l'ensemble des données cliniques/biologiques et d'examens du patient (Figure 5). Ces données sont de types qualitatives et quantitatives et sont saisies manuellement ou automatiquement via des imports (en fonction de la méthode de recueil).

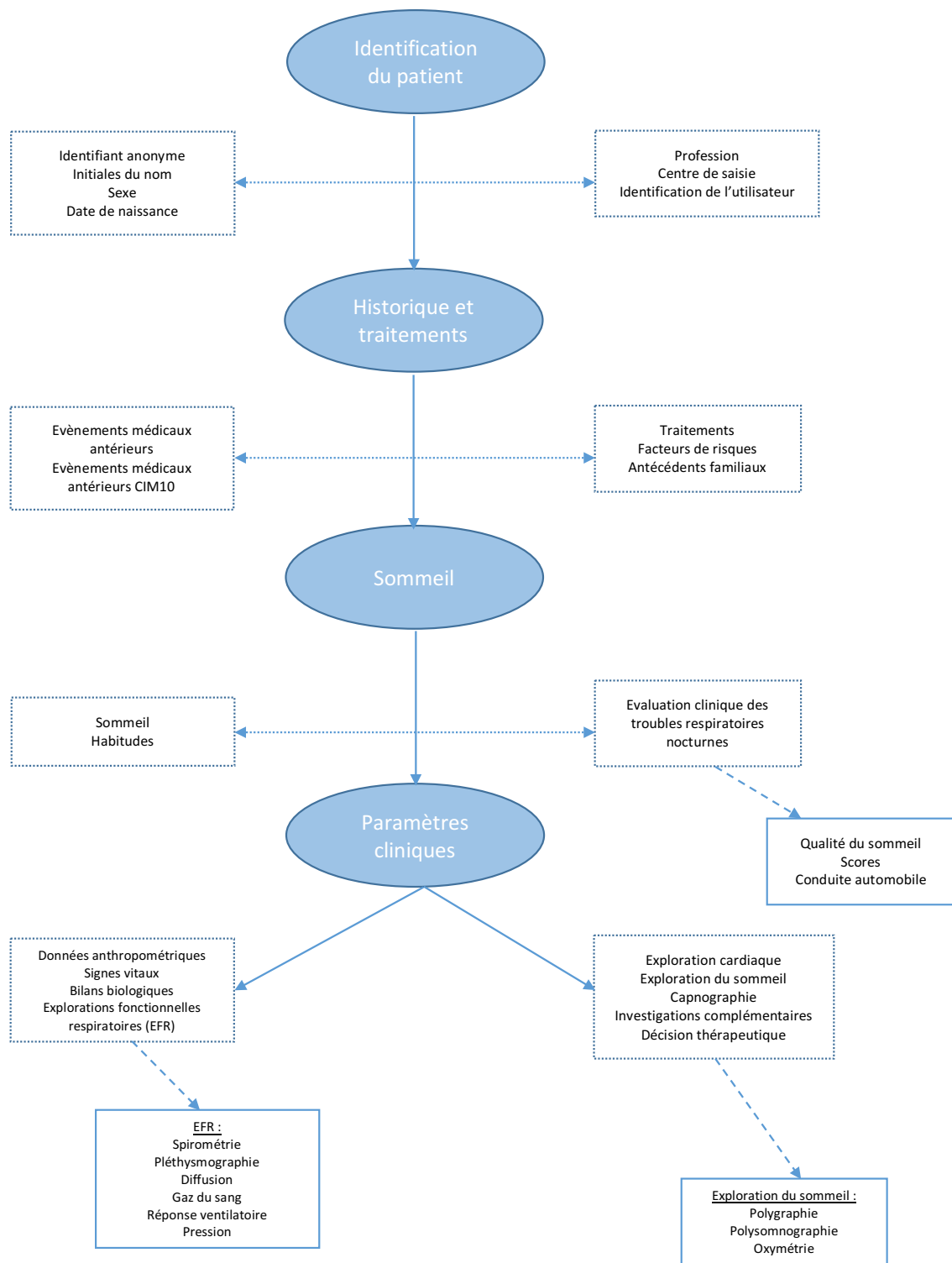


Figure 5 - Données à saisir dans les visites dans l'évaluation de la « fiche patient »

Il existe différentes façons de saisir les données dans la base (Figure 6) :

- Elles peuvent être sélectionnées via des cases à cocher (i.e. Antécédents).
- Elles peuvent être écrites manuellement par l'utilisateur et ces données peuvent dans certains cas être sélectionnées via un menu déroulant (i.e. Traitements).
- Ou encore renseignées via un curseur d'intensité allant de 0 à 10 par exemple (i.e. Trouble de la mémoire).

The figure displays three different data entry methods in the MARS system:

- Antecedents:** A form with radio buttons for 'Oui', 'Non', and 'NSP' to record medical history (e.g., HTA, HTAP, IDM).
- TRAITEMENTS:** A search interface for commercial names and ATC codes, with a dropdown menu showing search results for 'Paraceta'.
- Intensity Slider:** A visual scale from 0 to 10 for symptoms like 'Trouble de la mémoire', 'Transpiration nocturne', and 'Dyspnée', each accompanied by a 'DM' checkbox.

Figure 6 - Modalités de saisie possible dans la base MARS

C. La structure de la base MARS

1. Support informatique

La base de données MARS est une application en ligne accessible via une connexion chiffrée. Ses données sont stockées sur un serveur SQL MariaDB avec un chiffrement SSL permettant de répondre aux critères de sécurité du RGPD. Le serveur SQL est répliqué sur un autre serveur esclave, qui met à disposition une connexion sécurisée directe pour le requêtage. La base de données est sauvegardée de façon quotidienne. L'ensemble de ces opérations est effectué en continu, c'est-à-dire au fur et à mesure

du recueil des données. Le serveur qui héberge la base de données est situé dans un data center à Vitry-sur-Seine en France.

2. L'accès à la base

L'accès à la base se fait à l'aide d'un navigateur internet et d'identifiants personnels, délivrés par les administrateurs de la base. Ces accès sont utilisateurs-dépendants c'est-à-dire qu'en fonction de son statut à la connexion, un utilisateur n'aura pas accès aux mêmes informations ou ne pourra pas réaliser les mêmes opérations.

Différents accès à la base sont possibles en fonction de l'utilisation que les usagers peuvent avoir (Tableau I).

Tableau I - Profils utilisateurs de MARS et actions associées

	Super Admin	Accès graphique	Arc Superviseur	Accès restreint	Médecin Investigateur	User Admin Centre	Secrétaire
Créer des fiches patients	X	X	X	X	X	X	X
Accès à ses patients créés	X	X	X		X	X	X
Accès à l'ensemble des patients	X		X				
Création utilisateur pour son centre	X					X	
Création ou modification utilisateur tout centre	X						
Modification fiche patient	X						
Modification des bornes	X						
Accès aux graphiques de la page d'accueil		X					
Inclusion manuelle dans une étude	X		X				
Peut confirmer une donnée correcte ou aberrante					X		

3. Les fonctionnalités principales de la base MARS

On distingue trois modalités principales dans la base MARS :

- La « fiche patient » et les visites associées qui rassemblent les données cliniques du patient lors de son inclusion dans la base et au cours de son suivi.
- La proposition d'inclusion de patient dans les divers protocoles de recherche du CHUGA à partir de critères définis par l'administrateur.
- L'export des données à des fins d'analyses statistiques, de suivi qualité ou de reporting.

La création d'une « fiche patient » et des visites associées

Le module de création du patient permet de créer la « fiche patient » qui est organisée de la façon suivante :

- Le premier onglet « Informations générales » regroupe l'identification du patient, les informations relatives aux visites médicales et à l'inclusion dans les protocoles de recherche (Figure 7). Il est important de noter que les « fiches patients » sont anonymes pour assurer la confidentialité des données. Chaque fiche est identifiée via un numéro de patient sous la forme : PAXXXX, suivi des deux premières lettres du nom de famille, du mois et de l'année de naissance.

Cet onglet permet de renseigner et de suivre le dossier complet du patient. On distingue :

- L'onglet « évaluation générale » qui correspond au questionnaire principal rempli par le patient lors de sa visite dans le centre. Certaines informations nécessaires à « l'évaluation générale », telles que les examens médicaux (i.e. polysomnographie, gaz du

sang, etc.), ne sont pas toujours disponibles lors de la visite du patient, ils pourront être réalisés après la visite si besoin. Ces données seront saisies via des imports de données (source numérique) ou saisies manuellement par les médecins, les ARCS ou encore les secrétaires.

Fiche patient de PA4168 YO **/01/1995

Informations générales
Investigations complémentaires
Études
Sortie de base
Historique
Export vers BDD

Identifiant	PA4168 YO **/01/1995	Numéro visite	Date visite	Évaluation générale	Étude AT1HOME
Nom	YO.	Visite n° 1 : initiale	05/03/2018	Saisie en cours, 35 données non saisies.	Saisie en cours, 19 données non saisies.
Sexe	M	<input checked="" type="checkbox"/>			
Date de naissance	**/01/1995	Visite n° 2 : suivi	31/08/2018	Saisie en cours, 33 données non saisies.	Saisie en cours, 19 données non saisies.
Caractéristique profession		<input checked="" type="checkbox"/>			
Profession	DM				
Tags					
Centre	Centre Demo				
Saisi par	Yonnet Fabien				

Fiche patient :
Cette partie est remplie lors de la création du patient

Visites :
Réponses au questionnaire de l'évaluation générale

Étude :
Lorsque le patient est inclus dans une étude des données supplémentaires sont nécessaires

Figure 7 - Organisation de la « fiche patient » - Onglet "informations générales"

En complément de l'onglet « Informations générales », d'autres onglets sont disponibles comme (Figure 8):

- L'onglet « Investigations complémentaires ». Les visites des « fiches patient » possèdent une structure de base pour l'ensemble des patients qui peuvent être complétées par l'ajout d'examens spécifiques dans l'onglet « Investigations

complémentaires ». Ces examens peuvent être des polysomnographies, des tests de marche, des mesures de pressions artérielles, etc.

- L'onglet « Etudes » permet de voir si le patient participe ou non à une étude clinique. Il contient les informations spécifiques au protocole dans lequel un patient peut être inclus, comme par exemple : la date d'inclusion, la date de sortie, d'éventuelles déviations de protocoles, les questionnaires et les examens spécifiques à cette étude.
- L'onglet « Sortie de base » permet de renseigner la date de sortie de base du patient et ainsi de clôturer la « fiche patient ». Différentes modalités de sortie sont possibles : « fin conforme au protocole », « retrait de consentement », « décès », « décision de l'investigateur » et « autre ».
- L'onglet « Historique » permet de visualiser l'ensemble des modifications réalisées sur la « fiche patient » et d'assurer leur traçabilité.
- L'onglet « Export vers BDD (base de données) » permet de voir l'ensemble des exports vers d'autres bases partenaires réalisés pour la « fiche patient » : European Sleep Apnea Database (ESADA) et L'Observatoire Sommeil de la Fédération de Pneumologie (OSFP) réalisés.



Figure 8 - Organisation de la « fiche patient » - Autres onglets

L'inclusion de patient dans les protocoles de recherche clinique

La sélection et l'inclusion d'un patient dans un protocole peut se faire de deux façons :

- Durant une consultation. Cela est facilité par la génération d'une alerte (Figure 9) à la sauvegarde de la visite dans la « fiche patient ». En effet, si un patient répond aux critères d'inclusion, préalablement renseignés dans la base, une alerte visuelle apparaîtra et proposera l'inclusion du patient au médecin lors de la consultation.

The screenshot shows a patient record interface with a navigation bar at the top containing tabs: Informations générales, Investigations complémentaires, Études, Sortie de base, Historique, and Export vers BDD. The main content is divided into two sections. On the left is a patient information table, and on the right is a visit history table. A red box highlights a green alert message: 'Le patient correspond aux critères de l'étude Registre VAA' with a button 'Entrer dans l'étude Registre VAA'.

Identifiant	PA3135 SU **/03/1994
Nom	SU.
Sexe	F
Date de naissance	**/03/1994
Caractéristique profession	
Profession	Professions intermédiaires de l'enseignement, de la santé, de la fonction publique et assimilés
Tags	
Centre	Centre Demo
Saisi par	SUIZE Virginie

Numéro visite	Date visite	Évaluation générale
Visite n° 1 : Initiale <input checked="" type="checkbox"/>	10/10/2017	Saisie en cours, 21 données non saisies.
Visite n° 2 : suivi <input checked="" type="checkbox"/>	10/10/2017	Saisie en cours, 30 données non saisies.

Ajouter une visite de suivi N°3

Figure 9 - Alerte proposée lorsqu'un patient peut être inclus dans une étude

- A la suite de la visite d'un prestataire au domicile des patients. Cela se fait actuellement dans un protocole ciblé pour les patients diabétiques traités par une pompe à insuline (étude AT1HOME). Ces patients sont suivis par des infirmières de l'entreprise VitalAire. Si un patient répond aux critères d'inclusion de l'étude AT1HOME, l'infirmière pourra lui proposer d'être inclus dans l'étude après signature d'un consentement.

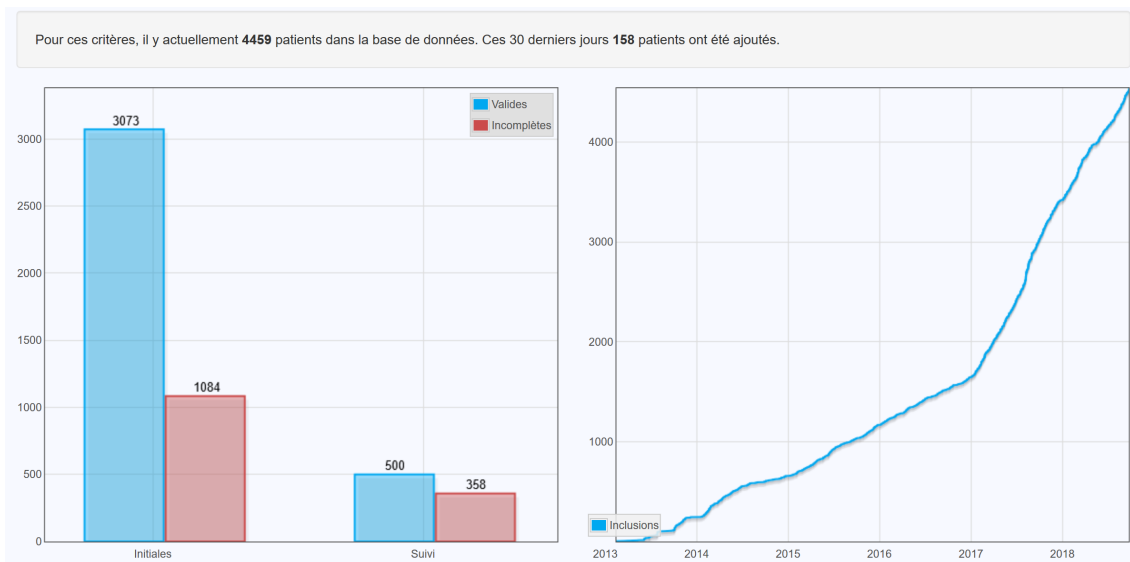
L'export des données patients

La base permet de réaliser des exports de données des patients inclus dans la base dans le but de les analyser et de les étudier pour la recherche. Ces exports peuvent être fait dans un format Excel selon plusieurs critères : selon le protocole clinique, selon le lieu d'inclusion, selon le médecin référent, etc.

4. Les fonctionnalités secondaires de la base MARS

Enfin, la base MARS possède des fonctionnalités secondaires :

- Les recherches des patients dans la base peuvent se faire selon différents critères : l'identifiant, les initiales, le mois et l'année de naissance, la personne qui a saisi le dossier, le médecin, le centre ou le protocole de recherche, etc. Ceci permet de cibler les recherches et de retrouver plus facilement des patients, 4.459 patients étant inclus dans la base au 26 septembre 2018.
- Des outils de reporting sont présents sur la page accueil de la base, après identification. Ces outils permettent de suivre certains paramètres comme présenté dans les Figure 10.



A

B

Figure 10 – Outils de reporting de la base

Graphique A : Histogramme de suivi de la complétude des visites

Graphique B : Evolution de l'inclusion des patients dans la base MARS

D. La conformité de MARS au contexte réglementaire

La collecte des données de la base MARS est conforme à la loi informatique et libertés et la mise en place du RGPD.

La déclaration CNIL pour la base MARS a fait l'objet d'un avis du Comité Consultatif sur le Traitement de l'Information en matière de Recherche dans le domaine de la Santé (CCTIRS) sur la base des objectifs médicaux et des moyens mis en place pour le recueil des données. Cette déclaration est en accord avec la MR-003 qui correspond aux traitements de données personnelles de santé ne nécessitant pas le consentement d'un patient. Une fois la demande d'avis favorable reçue, la déclaration à la CNIL a été faite et la collecte des données a été validée.

Pour ce qui est des changements induits par le RGPD, nombreux sont ceux auxquels la base de données MARS répondait déjà :

- Anonymisation des données
- Durée de stockage limitée
- Signature du formulaire de non opposition

Par ailleurs la base de données MARS est sauvegardée sur un serveur MariaDB dont la liaison est chiffrée par protocole SSL qui permet de répondre au critère de sécurité de la gestion des données personnelles. De plus les données médicales des patients sont saisies dans un formulaire sur un site web sécurisé et crypté https. L'accès à ce site est contrôlé par authentification.

Il reste donc l'aspect traçabilité du traitement des données à prendre en compte dans l'utilisation conforme de la base MARS. Pour cela, la base comporte déjà un système de recueil de chaque action effectuée, allant de la simple connexion d'un utilisateur à l'export de toutes les données dans un tableur. Par conséquent, les seules modifications à apporter à la base pour être aux normes du RGPD étaient un meilleur aperçu de la traçabilité des actions au sein de la base en cas d'audit et la mise en place d'un registre des traitements des données. Le premier point a été traité par la mise en place d'un outil retraçant toutes les modifications situées dans un intervalle de temps donné. Pour le second point, les exports de données ne se font plus directement à partir du site web lié à la base mais doivent passer par le data manager qui doit maintenir à jour ce registre en indiquant pour qui l'export est effectué, dans quel but et avec quelles données.

Il s'avère donc que la mise en place du RGPD n'a pas d'implication majeure dans la gestion de la base de données médicales MARS.

E. La qualité des données de la base MARS

La qualité des données est fondamentale, notamment lorsque ces données sont utilisées à des fins de recherche. Une erreur à la source des données entraîne un biais dans les résultats dont l'interprétation sera discutable, remettant ainsi en cause tout le travail réalisé en amont³⁷. La gestion de la qualité des données implique une prise en compte des données incorrectes, des données manquantes et des données dupliquées. Cela passe en premier lieu par la construction d'un formulaire de saisie semi-automatisé limitant les erreurs à la source, et également à différents outils permettant d'améliorer la qualité des données comme :

- La recherche de doublon lors de la création de la « fiche patient »

Lors de la création d'un nouveau patient, la base va faire une recherche de doublon à partir des deux premières lettres du nom de famille, du mois et de l'année de naissance. Cette recherche de doublon permet d'éviter de créer deux patients identiques. Si c'est le cas, une alerte visuelle apparaîtra et mentionnera les numéros de dossier des patients assimilés comme doublon pour vérifier si ces deux patients sont deux personnes distinctes ou non et de poursuivre ou non la création de la « fiche patient ».

Lors de la saisie des données dans les visites, un certain nombre d'outils qualité sont mis en place :

- Des alertes : Pour certaines variables quantitatives des alertes peuvent apparaître. Par exemple, pour la variable « Taille », une alerte apparaîtra à l'écran si la valeur saisie n'est pas comprise entre 1m et 2,2m. Suite à cette alerte, l'utilisateur pourra modifier la valeur saisie si elle est erronée ou la confirmer si cette valeur est correcte.
- Des menus déroulants : Ces menus permettent de sélectionner une information parmi un choix de propositions préenregistrées et non modifiables. Cette sélection permet d'harmoniser la saisie des données et d'éviter les abréviations et les fautes d'orthographe. Ceci a été mis en place par exemple pour les « traitements » et les « évènements médicaux antérieurs CIM-10 (Classification internationale des maladies 10^e révision) »
- Une alerte sur le nombre de données manquantes : une fois que la « fiche patient » est sauvegardée, le nombre de données manquantes est signalé sur la « fiche patient ».

Pour aller plus loin dans la gestion de la qualité et être dans une démarche d'amélioration continue, d'autres outils qualité ont été élaborés à la suite d'une analyse des besoins.

IV. La démarche

Comme la base MARS implique de multiples utilisateurs, chacun avec des attentes différentes, une étude qualitative des besoins a été réalisée. L'objectif de cette étude était de recueillir les avis sur l'utilisation de la base au quotidien, les difficultés ou les manques rencontrés de façon à dégager les propositions d'améliorations à mettre en place pour la qualité générale de l'outil.

A. Le questionnaire

Le témoignage des différents utilisateurs a été recueilli sur la base d'un questionnaire commun qui reposait sur les points essentiels de la base MARS (Annexe 1):

- L'utilisation générale de MARS : type d'utilisation (cohorte prospective ou protocole de recherche clinique) ; fréquence d'utilisation.
- Les avantages/inconvénients de la base en général.
- Les problèmes de saisie dans chaque sous partie de la base.
- Les attentes par rapport à l'outil actuel.
- Les besoins en termes d'outils de reporting et de formation.

L'analyse des informations recueillies a été transformé ensuite en plan d'action.

B. Les personnes rencontrées

D'octobre à décembre 2017, différents utilisateurs ont été sollicités pour répondre à ce questionnaire. Douze personnes de 5 centres différents ont été interrogées :

- Le CHU UCL Namur en Belgique
- L'institut universitaire de cardiologie et de pneumologie de Québec

- Le centre hospitalier Alpes-Léman (CHAL) à Findrol
- L'hôpital Pedro Hispano à Porto
- Le CHUGA à Grenoble

Ces personnes ont des fonctions différentes : six médecins, quatre attachés de recherche clinique (ARC), une secrétaire et une technicienne de laboratoire. Tous réalisent des saisies de données sur MARS soit dans le cadre de consultations soit dans le cadre de protocole de recherche.

V. Les résultats de l'investigation et les actions mises en place

1. Utilisation générale et les avantages/inconvénients de la base

D'un point de vue général, l'ensemble des utilisateurs connaissaient l'objectif ou la finalité de la base que ce soit en termes d'inclusion dans la cohorte prospective ou de protocole de recherche clinique.

Selon les intervenants, l'utilisation de la base pouvait être quotidienne, régulière voir très irrégulière par manque de temps. De façon générale, des utilisateurs grenoblois (médecins, ARCs, techniciens) de la base ont une utilisation régulière car l'utilisation de cette base est maintenant bien ancrée dans leurs pratiques. Dans les autres centres hospitaliers, l'utilisation est encore irrégulière (médecins, techniciens) car la base MARS est un nouvel outil encore difficile à utiliser comme outil à part entière. En effet, l'utilisation de la base leur semble chronophage donc l'outil est utilisé en arrière-plan. Il sera donc nécessaire par la suite, de leur expliquer le processus d'utilisation de la base et de redéfinir le rôle de chacun pour qu'ils ne voient plus cet outil comme chronophage.

Globalement, les utilisateurs trouvent que MARS est un très bon outil, synthétique, très bien organisé et intuitif. De plus, l'utilisation de la tablette dans certains centres permet de gagner du temps sur la saisie des informations. Il a été également mentionné que cet outil est un réel atout pour saisir les données des différentes visites d'un patient et de réaliser les exports vers les bases de données partenaires externes

(ESADA, OSFP). De plus, les outils qualité mis en place sur la base MARS sont très appréciés et aident les intervenants à limiter les erreurs de saisie.

Des inconvénients ont été mentionnés mais il est intéressant de noter qu'ils sont souvent en lien avec un problème d'organisation et un manque de temps dont disposent les utilisateurs. En effet, l'utilisation de la base se fait en plus de leurs fonctions quotidiennes, il peut être difficile dans un premier temps d'intégrer MARS dans leur travail quotidien mais aussi dans l'organisation d'un service entier. Il est donc intéressant et nécessaire, dans ce cas, de faire un point avec les utilisateurs et de les aider à mieux appréhender la base et son utilisation via un accompagnement individuel et la nécessité de mettre en place des actions de formation.

D'autres inconvénients ponctuels ont été mentionnés comme : le manque de champs libres, le manque d'informations sur l'examen Oto-Rhino-Laryngologie (ORL), des problèmes d'interfaces de la base, etc. Ces remarques ont été prises en compte et ont fait l'objet de modifications ou d'améliorations lorsque cela était justifié et nécessaire. Les réponses des utilisateurs au questionnaire sont mentionnées dans l'Annexe 2. Les justifications de la mise en place d'action ou non des administrateurs sur certaines de ces remarques sont présentées dans l'Annexe 13.

2. Problèmes de saisie des données

Dans cette thèse, seules les réponses sur les parties des « Evènements médicaux personnels » et « Traitements » seront développées. En effet, ces sections posaient le plus de problème et demandaient un travail prioritaire.

Cependant, les problèmes sur la saisie des données dans la base ont été recueillis pour l'ensemble des sous parties de l'évaluation générale.

L'ensemble des problèmes de saisie, hors « Evènements médicaux personnels » et « Traitements », sont présentés en Annexe 3 et les justifications des administrateurs sur la mise en place ou non d'actions dans l'Annexe 13. Certaines de ces remarques ont été traitées, d'autres sont en cours de discussion ou en projets futurs.

La sous partie « Evènements médicaux personnels » est découpée en deux parties :

1. La partie « Antécédents personnels » qui contient une liste non exhaustive à cocher (« oui » ou « non ») des antécédents généraux. Cette liste regroupe une liste d'antécédents les plus courants pour faciliter et avoir une saisie simple et rapide.
2. La partie « Evènements médicaux CIM10 » qui est une zone de champs libres reliée à la CIM-10 permettant de rajouter des antécédents/événements médicaux selon la classification internationale des maladies, 10^e révision (CIM-10), qui ne serait pas présent dans la partie précédente « Antécédents personnels ». C'est une classification statistique non exclusivement médicale codant notamment les maladies, signes, symptômes, circonstances sociales et causes externes de maladies ou de blessures, publiée par l'Organisation mondiale de la santé (OMS)³⁸.

La sous partie « Traitements » permet de renseigner l'ensemble des traitements du patient via un menu déroulant basé sur la classification anatomique, thérapeutique et chimique (ATC) des médicaments. Ceux-ci sont divisés en différents groupes selon

l'organe ou le système sur lequel ils agissent ou leurs caractéristiques chimiques ou thérapeutiques.

a) Zoom sur les problèmes au niveau des « Evènements médicaux personnels »

Il existe deux types de problèmes qui peuvent être identifiés lors de la saisie (Figure 11) :

- Des problèmes majeurs qui affectent de façon plus importante à la qualité des données ;
- Des problèmes mineurs, liés à l'utilisation.

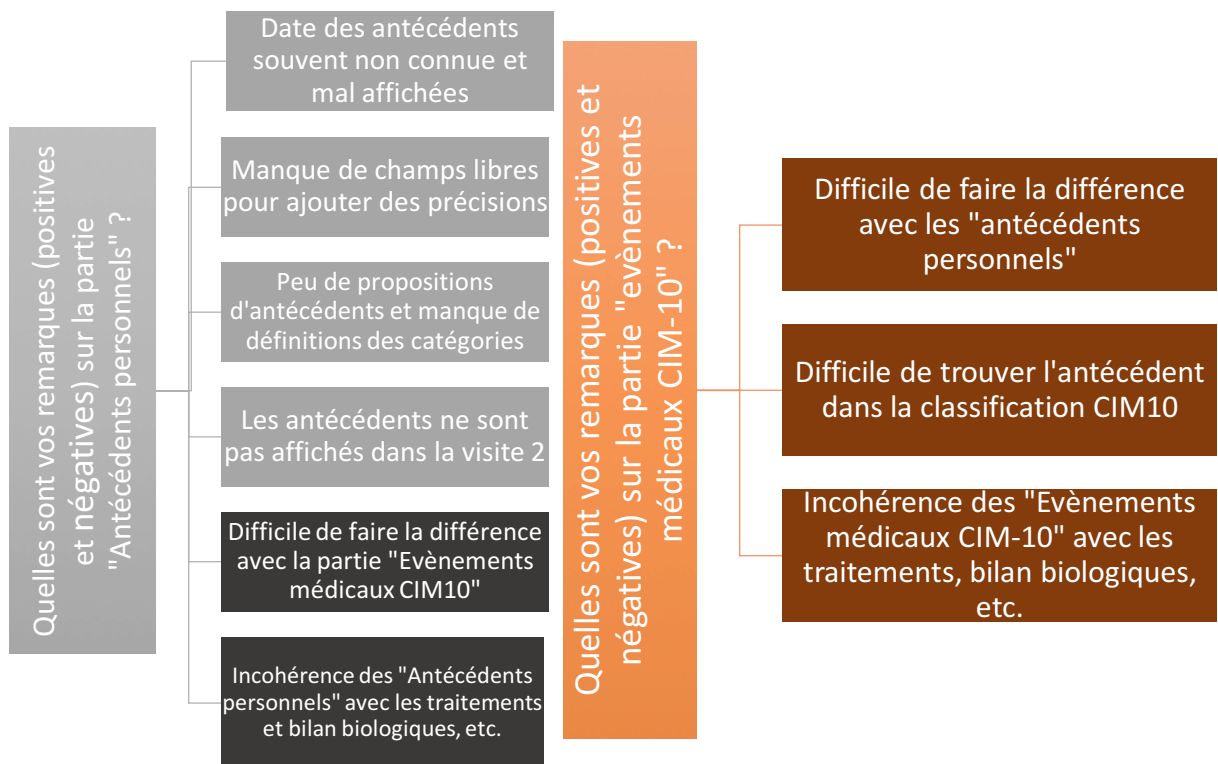


Figure 11 - Problèmes rencontrés sur les "Antécédents personnels" et les "Evènements médicaux (CIM-10)"

Légende : Couleurs claires : problème mineurs ; Couleurs foncées : Problèmes majeurs

Les problèmes majeurs :

Difficultés de différencier les « Antécédents personnels » des « Evènements médicaux CIM-10 »

Problématique :

La compréhension des parties « Antécédents personnels » et « Evènements médicaux (CIM10) » n'était pas la même selon les utilisateurs (médecins, ARC, secrétaires).

Par exemple en Belgique, la partie « Antécédents personnels » était utilisée pour renseigner les antécédents des patients sur le long terme alors que la CIM10 était utilisée pour renseigner un événement médical spontané apparu à un instant « t » comme par exemple une hospitalisation. Si l'antécédent n'était pas présent dans la partie « Antécédents personnels », il n'était pas rajouté avec la CIM-10.

Alors qu'à Grenoble, ces deux parties étaient utilisées correctement. Les items de la CIM-10 venaient compléter les antécédents du patient non présent dans la partie « Antécédents personnels ».

Cette interprétation différente entre les acteurs est un problème majeur d'harmonisation de saisie des données. Il est donc primordial de former et de structurer l'information pour éviter ce type de confusion.

Action apportée :

Pour pallier à cette difficulté et faciliter la compréhension, il a été décidé de modifier les noms de ces deux parties (Tableau II).

Tableau II - Modification des termes sur les Antécédents/événements médicaux

Visite initiale		Visites de suivi	
Avant modification	Après modification	Avant modification	Après modification
Antécédents personnels	Événement médicaux antérieurs	Événements médicaux	Nouvel événement médical depuis la dernière visite
Événement médicaux CIM-10	Événement médicaux antérieurs CIM-10		

Dans la visite initiale, il a été décidé de mettre un seul terme : « Événements médicaux antérieurs » pour éviter toute confusion.

Pour les visites de suivi, la partie « antécédents personnels » n’existait pas, seul la CIM-10 était présente. Les termes ont cependant été modifiés pour faciliter la compréhension. Il a été précisé que les antécédents à mentionner devaient être les nouveaux antécédents apparus depuis la visite précédente.

Difficile de trouver l'antécédent dans la classification CIM-10

Problématique :

La CIM-10 est constituée d’un ensemble ordonné de 14 400 codes différents adaptés à des situations médicales précises. De ce fait, il peut être difficile de trouver rapidement l’item que l’on souhaite renseigner dans cette classification, à fortiori en l’absence de formation médicale.

En effet, l’item peut être soit trop précis, comme dans le cas de la glomérulonéphrite aigue (Tableau III). Soit il présente des dénominations particulières difficiles à trouver dans la classification. Comme par exemple pour l’item CIM-10 correspondant à une

allergie aux pénicillines. Pour renseigner ce type d'allergie il faut écrire correctement l'ensemble de l'item « Effets indésirables lors d'usage thérapeutique de pénicillines ».

Tableau III - Items CIM-10 proposés pour qualifier un syndrome néphritique aigu

Codes possibles pour coder une syndrome néphritique aigu via la CIM-10	Signification
N00.0	- Syndrome néphritique aigu avec anomalies glomérulaires mineures
N00.1	- Syndrome néphritique aigu avec lésions glomérulaires segmentaires et focales
N00.2	- Syndrome néphritique aigu avec glomérulonéphrite membraneuse diffuse
N00.3	- Syndrome néphritique aigu avec Glomérulonéphrite proliférative mésangiale diffuse
N00.4	- Syndrome néphritique aigu avec glomérulonéphrite proliférative endocapillaire diffuse
N00.5	- Syndrome néphritique aigu avec glomérulonéphrite mésangio-capillaire diffuse
N00.6	- Syndrome néphritique aigu avec maladie à dépôt dense
N00.7	- Syndrome néphritique aigu avec glomérulonéphrite diffuse en croissant
N00.8	- Syndrome néphritique aigu avec autres modifications morphologiques
N00.9	- Syndrome néphritique aigu

Pour pallier à ce problème les professionnels utilisent le moteur de recherche internet pour trouver l'item correspondant à leur recherche. Ceci est donc pour eux une perte de temps et une source d'erreur ou d'imprécision possible. Ceci d'autant plus que quand le moteur de recherche internet n'apporte pas de réponse, l'item est

sélectionné arbitrairement avec plus ou moins de précision sur la pathologie, sans lien avec l'antécédent du patient.

Action apportée :

Pour faciliter la recherche d'événement indésirable dans la CIM-10, un dictionnaire d'aide à la saisie CIM-10 a été réalisé (Annexe 4).

Ce dictionnaire a été réalisé à partir des événements indésirables les plus fréquemment renseignés dans la base à partir de requêtes informatiques. Ensuite, les antécédents/événements ont été associés aux codes CIM-10 correspondants pour en faciliter la saisie et limiter les recherches d'informations complémentaires sur les moteurs de recherches internet.

Cependant, il est nécessaire de noter que ce dictionnaire n'est pas exhaustif et ne répondra pas à l'ensemble des attentes des utilisateurs dans certains cas bien précis.

Par la suite, ce dictionnaire pourra être amélioré et complété si son utilisation est bénéfique. De plus des dictionnaires des variables CIM-10 propres à chaque étude pourront être créés, pour être encore plus précis.

Les problèmes mineurs :

La date des antécédents est souvent inconnue et le format n'est pas adapté

Problématique :

Dans le cas des pathologies chroniques, comme le SAS plusieurs facteurs peuvent être à l'origine comme l'obésité, les modifications craniofaciales, l'altération de la fonction musculaire des voies aériennes supérieures, la neuropathie pharyngée, etc. ^{15,17}. De plus, l'apparition SAS peut à son tour être à l'origine de différents événements comme une dysfonction endothéliale, un stress oxydatif, une inflammation systémique pouvant impacter le système cardio-vasculaire à moyen-long terme ou pouvant induire des perturbations métaboliques ^{15,17}.

En effet, les études existantes confirment une relation causale entre apnée du sommeil et hypertension, maladies coronariennes, arythmie, insuffisance cardiaque et accidents vasculaires cérébraux ^{15,17}. Il existe donc une chronologie bien définie dans la survenue des comorbidités cardiovasculaires.

La connaissance des dates de survenue des événements médicaux antérieurs à une consultation permet de retracer la trajectoire de santé des patients et peut être pris en compte dans l'analyse des données à des fins de recherche. Cependant, le mode de recueil des dates des événements médicaux antérieurs dans la base, notamment le fait de devoir renseigner la data complète, était à l'origine de données manquantes. En effet, les patients ne se souviennent que rarement de la date précise de survenue d'un antécédent médical, certains pouvant remonter à plusieurs années.

De plus, en terme fonctionnel, la saisie de la date de survenue des antécédents n'est possible que lors de la sauvegarde de la visite. Ceci est une perte de temps car l'utilisateur doit revenir sur la visite pour saisir la date d'apparition de l'antécédent.

Action proposée :

Normalement le questionnaire est envoyé au patient en amont de la visite pour qu'il puisse le remplir et le médecin récupère l'ensemble des informations lors de la consultation. Cependant cet envoi n'est pas pratiqué par l'ensemble des médecins, c'est pour cela que dans certain cas le questionnaire est rempli dans la salle d'attente avant la visite.

Lors de la prise de rendez-vous, il serait intéressant de sensibiliser le patient sur l'importance du remplissage du questionnaire en amont de la visite pour récupérer l'ensemble des données et notamment les dates des antécédents. Dans le cas de figure où il n'est pas possible pour le patient d'avoir cette information, il pourrait être intéressant de faire jouer le réseau professionnel hôpital-libéral pour la récupérer.

Pour ce qui est du problème d'affichage de la date des antécédents, ceci est dû à une mauvaise programmation de l'action dans le système informatique de la base. Cette erreur a été corrigée directement par l'informaticien.

Manque de champs libres pour ajouter des précisions

Problématique :

Sur la partie « Antécédents personnels », il a été mentionné que l'ajout de champs libres apporterait plus de précision sur les données des patients. En effet, sur cette partie, il est seulement possible de sélectionner, via une case à cocher, la présence ou non d'une liste d'antécédent. Par exemple, il suffit de cocher l'antécédent « Diabète de type I » s'il est présent. Mais il n'existe pas d'autres champs pour renseigner d'autres détails sur l'antécédent « Diabète de type I ».

Action proposée :

Il a été décidé de ne pas rajouter plus de champs libres dans la base. Cette décision s'explique par le fait que la base de données MARS, ne doit pas se substituer au dossier médical des patients et qu'il existe déjà un champ libre, en fin de dossier patient permettant de mentionner des observations annexes si besoin. Il n'est donc pas nécessaire d'en ajouter.

De plus, à partir d'un même terme, il est possible de générer différentes entrées similaires dans un champ libre, mais qui devront faire chacune l'objet d'un traitement spécifique pour uniformiser les données (Tableau IV). Cela engendre du temps de traitement supplémentaire et un risque d'erreur possible.

Tableau IV - Diversité générée par les champs libres

Idée initiale	Différentes possibilités d'écriture
<p style="text-align: center;">Pratique le vélo en compétition</p>	<p style="text-align: center;">→</p> <p style="text-align: center;">Pratique le vélo en compét</p> <p style="text-align: center;">Pratique le velo en compétition</p> <p style="text-align: center;">Pratique le velo en competition</p> <p style="text-align: center;">Vélo : Compétition</p> <p style="text-align: center;">PRATIQUE LE VELO EN COMPETITION</p>

Peu de propositions d'antécédents et manque de définition des catégories

Problématique :

Selon certains utilisateurs, la liste des antécédents mentionnée dans la partie « Antécédents personnels » n'est pas assez exhaustive et précise.

De plus, certains antécédents à cocher peuvent regrouper différentes pathologies comme dans le cas des « Maladies neurologiques » et des « Maladies psychologiques ». Dans certains cas, cette absence de précisions peut être à l'origine d'une incompréhension quant aux données à saisir et de ce fait entraîne une baisse de la qualité des données.

Action proposée :

Cette liste proposée dans la partie « Antécédents personnels » est déjà conséquente, et son intérêt serait perdu si elle devenait trop exhaustive. En effet, il serait plus long de chercher l'antécédent en question dans la liste que dans la CIM-10.

Pour pallier au problème de regroupement de certains antécédents dans un groupe de pathologies, chaque regroupement sera mentionné de façon explicite dans le système documentaire et l'information sera transmise aux utilisateurs. Par exemple, pour le cas des « Maladies neurologiques » il sera mentionné que pour tout patient atteint de maladie d'Alzheimer, maladie de Parkinson, d'épilepsie, de sclérose en plaque, etc. il faudra cocher l'antécédent « Maladie neurologique ».

Les antécédents ne sont pas affichés dans la visite 2

Problématique :

Les antécédents saisis dans l'évaluation générale de la visite initiale (ou visite 1) ne sont pas reportés dans la visite 2. Ceci est une contrainte supplémentaire, car l'utilisateur est obligé d'ouvrir à nouveau le formulaire de visite initiale pour voir l'ensemble des antécédents du patient.

Action proposée :

Le non report des antécédents dans les visites de suivi est normal et voulu. En effet, La deuxième visite permet de mentionner les modifications et les changements de données concernant le patient par rapport à la visite précédente. Elle n'a pas la fonction de résumer la visite précédente et de ce fait de se substituer au dossier médical du patient. De plus, pour que l'utilisateur comprenne que seuls les nouveaux antécédents doivent être renseignés dans les visites de suivi, le titre de cette partie a été modifié pour qu'il n'y ait pas de divergence entre les utilisateurs (Tableau II).

b) Zoom sur les Traitements

Après la partie sur les antécédents, la partie « Traitements » de l'évaluation générale posait problème (Figure 12). De la même façon des problèmes majeurs et mineurs ont été identifiés.

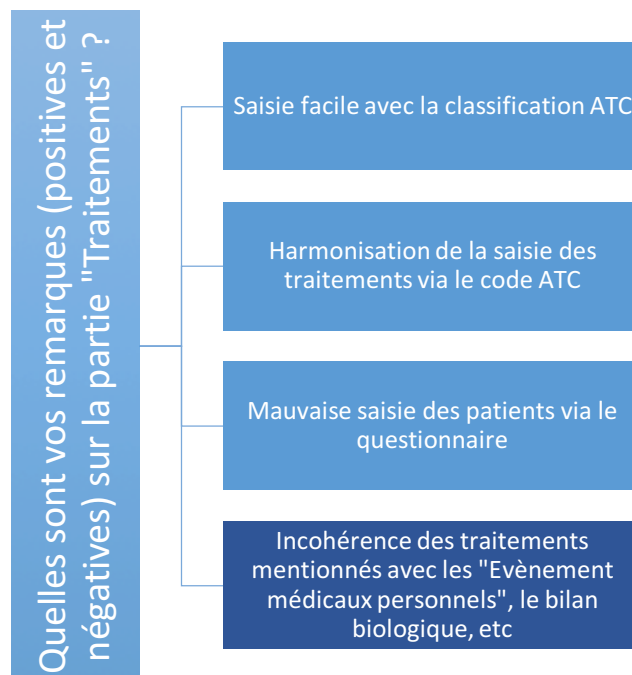


Figure 12 - Problèmes rencontrés sur la partie "Traitements"

Légende : Couleurs claires : problème mineurs ; Couleurs foncées : Problèmes majeurs

Problème majeur :

Incohérence des données saisies dans les parties des "Evénements médicaux personnels", des « Traitements », des bilans biologiques, etc.

Problématique :

Il s'agit d'un problème commun (Cf. Figure 11 et Figure 12) à la partie « Traitements » et « Evénements médicaux personnels » et entraîne un manque de cohérence de

certaines données saisies dans la base. En effet, aucun moyen actuel ne permet de faire le lien entre les différentes données saisies dans la base et d'attester de leur cohérence. A l'heure actuelle, aucune vérification de cohérence n'est faite entre les données (par exemple entre les événements médicaux personnels, les traitements et même les données biologiques du patient). C'est à dire qu'un patient peut être saisi avec :

- Un antécédent d'hypertension artérielle, des valeurs de pression artérielle diastolique et systolique haute et aucun traitement antihypertenseur,
- Un traitement antihypertenseur mais aucun antécédent d'hypertension artérielle et des valeurs de tension normales,
- Un patient hypertendu, une pression artérielle normalisée et aucun traitement.

Sauf dans certains cas exceptionnels (prise en charge du patient en attente ou particulière), ces incohérences non justifiées ne sont pas un gage de la qualité des données saisies. Par la suite, lors d'analyse des données, cela peut entraîner un biais dans l'interprétation des résultats.

Action proposée :

Pour cette partie, il a été décidé de se focaliser sur les événements médicaux et les traitements et plus précisément sur l'antécédent hypertension artérielle (HTA) et sa prise en charge médicamenteuse associée. Ce choix a été fait selon plusieurs critères :

- Se focaliser sur un antécédent pour étudier l'incohérence avec précision et déployer de façon satisfaisante des actions. Si l'alerte fonctionne et montre un intérêt, ceci sera ensuite généralisé aux autres antécédents.

- Se focaliser sur l'HTA car l'HTA est un antécédent avec un réel enjeu de santé publique. De plus, le SAS peut être à l'origine de l'HTA, donc il était concret de se focaliser sur celui-ci dans le cadre de la base MARS car un certain nombre de patient sont atteints d'HTA.

Pour mener à bien ceci, la première étape a été de lister les classes ATC correspondantes reliées à une HTA :

- Diurétiques : Classe ATC C03
- Bêtabloquants : Classe ATC C07
- Inhibiteurs calciques : Classe ATC C08
- Inhibiteurs de l'enzyme de conversion : Classes ATC C09 A et B
- Antagonistes de l'angiotensine II : Classes ATC C09 C et D

La deuxième étape était la réalisation d'un état des lieux du taux d'incohérences dans la base. Pour cela plusieurs conditions ont été établies :

- Un antécédent d'HTA est associé à au moins un traitement antihypertenseur appartenant à une des classes ATC ci-dessus : Les données sont cohérentes.
- Un antécédent HTA sans traitement antihypertenseur peut être une incohérence. En effet, si les règles hygiéno-diététiques dans la prise en charge de l'HTA ne sont pas suffisantes alors un traitement antihypertenseur doit être instauré.
- Un antécédent d'HTA non listé et sans traitement associé est considéré comme une cohérence.

- Un antécédent d'HTA non listé avec la présence d'un traitement antihypertenseur est considéré comme une incohérence si le patient ne présente pas d'autres antécédents cardiovasculaires. En effet, les antihypertenseurs peuvent être utilisés dans d'autres pathologies cardiaques en dehors de l'hypertension comme dans le cas d'une insuffisance cardiaque, maladies coronariennes, angine de poitrine, etc. Ce cas particulier ne peut donc pas être considéré comme une incohérence.

Le taux d'incohérence entre ces relations est donné par le rapport entre le nombre d'incohérences sur le nombre total de patient pris en compte.

Autre élément à prendre en compte, chaque traitement médical est référencé par son code ATC. Cependant il faut faire attention car la classe ATC regroupe plusieurs variables pour un même code. Par exemple : Le code ATC d'une molécule regroupe toutes les formes possibles de cette molécule comprenant des dosages différents, des formes pharmaceutique (i.e. solution, comprimés), etc. Le détail des cohérences et incohérences est présenté dans la Figure 13 ci-dessous :

Évaluation générale de visite initiale Cohérence ATCD HTA et TTT

HTA : Antécédent d'Hypertension Artérielle
 CARD : Antécédent Cardiovasculaire
 TTT : Traitement anti-hypertenseur
 ATC : Code ATC lié à un traitement

+ : Présence
 - : Absence

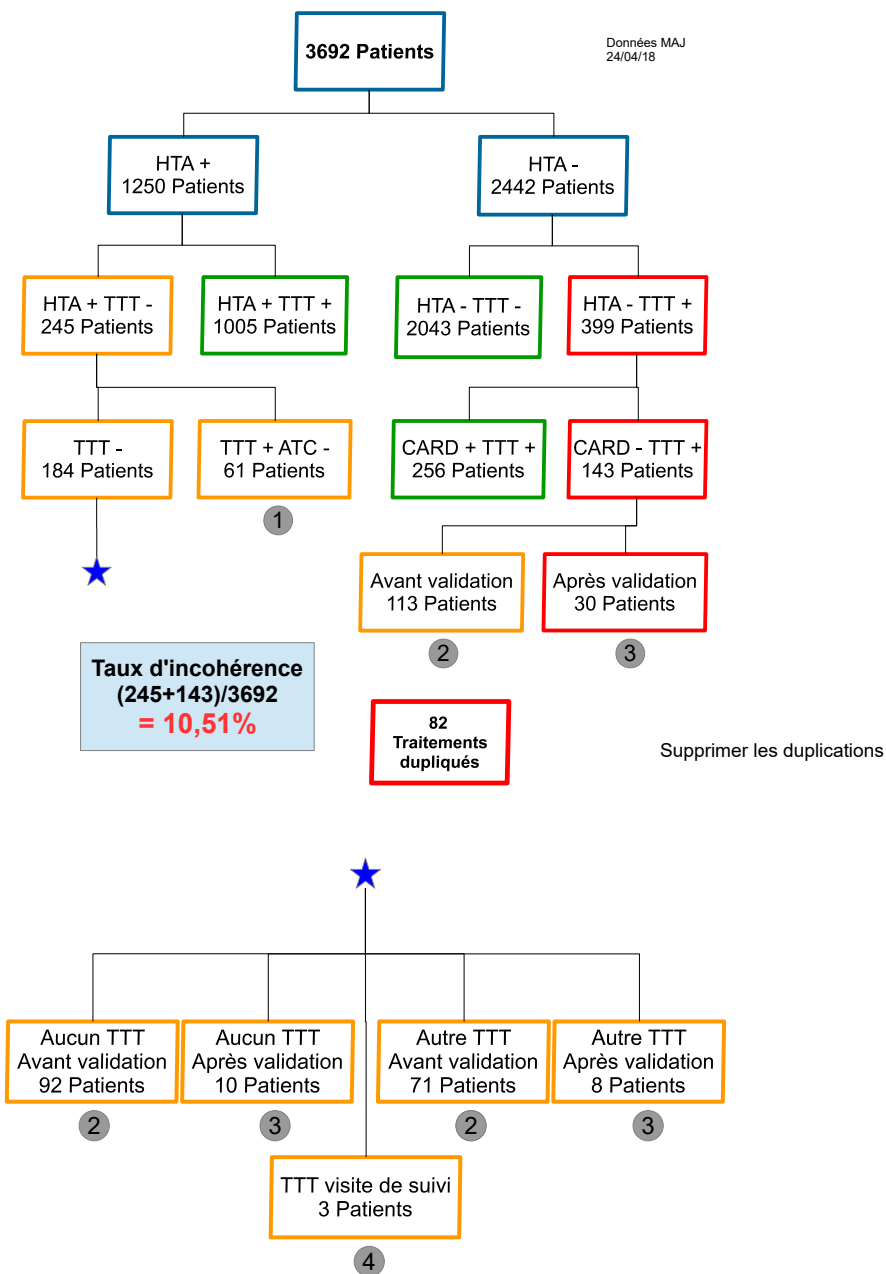


Figure 13 - Arbre décisionnels des cohérences/incohérences sur l'hypertension artérielle

Les cases représentées en bleu montrent une information ni cohérente, ni incohérente. Celles en vert représentent les cohérences et en rouge les incohérences. Enfin la case en orange représente l'incohérence qui peut être résolue.

C'est en suivant la définition précédemment décrite que des premiers résultats ont été présentés au 24 avril 2018 concernant ce critère de qualité.

Nous avons pu mettre en évidence un taux d'incohérence général de 10,51% sur l'ensemble de la base MARS pour l'HTA.

Avec cette analyse, il est possible de mettre en évidence différentes incohérences :

- Incohérence de traitement : Antécédent présent sans traitement associé ou antécédent absent avec un traitement antihypertenseur (sans antécédent cardiovasculaire). Ceci concerne 327 patients sur 3692. La vérification des évaluations générales par le médecin permet de faire diminuer le nombre d'incohérences à 48 (initialement : 327). En effet, les médecins référents sont amenés à vérifier les « fiches patient ». Suite à cette validation, normalement, aucune incohérence ne doit être retrouvée dans la « fiche patient ».
- Incohérence due à la classe ATC : Présence d'un antécédent, d'un traitement mais la classe ATC n'est pas renseignée. Ceci concerne 61 évaluations générales.
- Dans 82 évaluations générales il a été retrouvé que les traitements étaient dupliqués.
- Et pour finir, dans 3 fiches patients, les traitements étaient mentionnés dans la visite de suivi et non dans la visite initiale. Pour ces 3 patients, il suffit de basculer les traitements de la visite de suivi vers la visite initiale.

Ce taux est assez important et il est intéressant de noter que ce dernier ne correspond qu'à l'association antécédent HTA/Traitement antihypertenseur. Il est donc très fortement probable que d'autres incohérences de ce type existent dans la base MARS.

Pour faire diminuer ce taux général d'incohérence, il a été décidé de mettre en place diverses actions (Figure 14) :

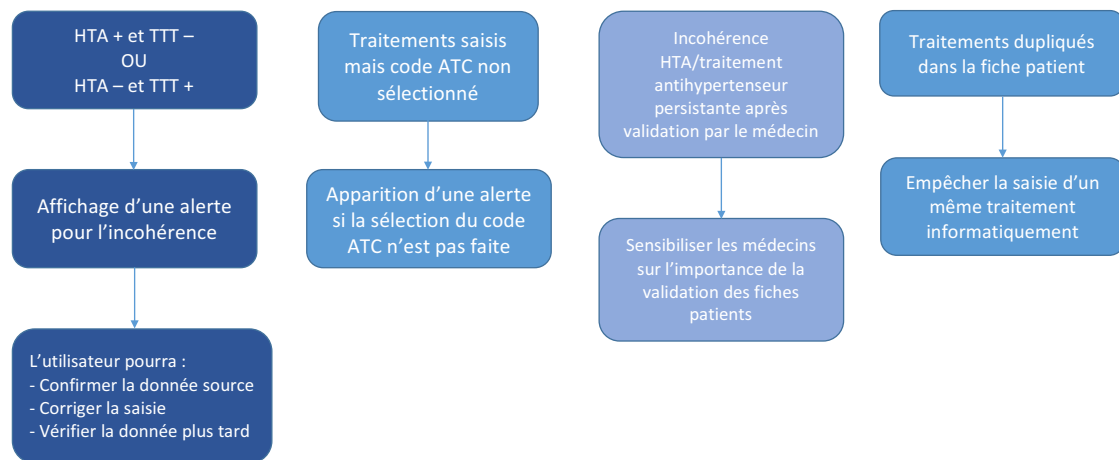


Figure 14 - Incohérences retrouvées et actions mises en place

Il est important de noter que l'alerte d'incohérence (Figure 15) a été déployée dans un premier temps sur un petit groupe d'utilisateurs pour qu'ils puissent se familiariser avec l'outil et faire remonter des remarques si nécessaire. Au bout de quelques semaines, après résolution des différents problèmes constatés et confirmation de l'intérêt, l'alerte sera déployée sur l'ensemble des utilisateurs de la base.

Attention donnée aberrante x

La présence d'HTA n'est pas cohérente avec les traitements. (Traitements :
Aucun)

Conforme donnée source

Donnée en cours de vérification

Corriger

Figure 15 - Alerte s'affichant à l'écran quand une incohérence HTA/traitement est détectée

Problèmes mineurs :

Saisie facile avec la classification ATC et harmonisation de la saisie des traitements via le code ATC

Problématique :

L'utilisation de la classification ATC pour les traitements permet de faciliter et d'harmoniser la saisie des données. Cependant, cela reste limité aux centres français. En effet, pour les autres établissements, les noms des spécialités peuvent être différentes (Tableau V), et la correspondance ne se fait pas avec le menu déroulant proposé. Il était nécessaire de compléter cette partie avec la dénomination commune internationale (DCI) ou le princeps, la voie d'administration, le dosage, etc. Il n'existait pas d'harmonisation sur cette donnée.

Cependant, de façon très minoritaire, il existe encore des dossiers patients dans la base MARS avec des traitements non harmonisés sur la classification ATC (519 au 12/07/2018).

Tableau V - Spécialités existantes en paracétamol en France et au Canada

France	Canada
Paracétamol Doliprane®	Paracetamol Tylernol®
Paracétamol Efferalgan®	Paracetamol Atazol®

Action proposée :

Pour pallier à ce problème, il sera nécessaire de saisir le nom du traitement via la DCI. En effet, les DCI sont les mêmes dans le monde entier. Cette saisie via la DCI permettra de s'affranchir de la divergence des noms commerciaux, des dosages et des formes pharmaceutiques dans la base MARS.

Ce qui est important c'est de connaître la molécule prise par le patient. La formulation galénique ou le dosage peuvent avoir de l'importance dans des études plus ciblées, mais dans le cas d'une cohorte prospective, ces informations ne sont pas indispensables. De plus, quel que soit le conditionnement ou le dosage, le code ATC sera le même.

En conclusion, dans un premier temps la saisie via la DCI sera demandée mais dans l'idéal il faudrait que la saisie du traitement puisse se faire selon différents paramètres soit la DCI, soit la classe ATC soit le nom commercial. La saisie d'un de ces trois paramètres permettra de sélectionner automatiquement les deux autres paramètres

et la saisie sera donc harmonisée. Cette deuxième proposition peut être une perspective d'évolution de la base.

Mauvaise saisie des traitements par les patients via le questionnaire

Problématique :

Lors de la saisie sur le questionnaire informatique, il n'est pas précisé qu'il faut mentionner un traitement par ligne. De ce fait, les patients ont tendance à écrire l'ensemble des traitements sur la même ligne, ce qui rend impossible la reconnaissance des médicaments avec la classification ATC et le professionnel est obligé de retranscrire les traitements manuellement sur la base. Ceci est un problème qui peut être considéré comme mineur d'un point de vue de gestion mais aussi comme majeur car cette tâche peut être chronophage pour le médecin par la suite.

Action proposée :

Pour pallier à ce problème, il sera précisé de ne mentionner qu'un seul médicament par ligne sur le questionnaire rempli par les patients. Cette simple précision permettra de faire la présélection ATC automatiquement et éviter la réorganisation des traitements par les médecins par la suite.

3. Les souhaits d'évolution de la base par les utilisateurs

Pour finir, les principales attentes des utilisateurs vis-à-vis de la base concernent principalement l'optimisation de l'utilisation de la base et la mise en place de nouveaux outils pour aller plus loin dans l'exploitation des données (Figure 16).

Plusieurs pistes de développement ont été proposées :

- Des modes d'emploi pour les diverses professions : administrateurs, médecins, ARC, secrétaires, techniciens de laboratoire. Ces manuels permettront d'avoir, sur un support écrit et formalisé l'ensemble des actions réalisées par les différents utilisateurs.
- Des informations ponctuelles donnant une explication des modifications informatiques réalisées sur la base. En effet, à l'heure actuelle, les utilisateurs ne sont pas avertis des modifications faites, ils les découvrent lorsqu'ils se connectent de nouveau sur la base.

De plus, les utilisateurs aimeraient que la base propose des outils supplémentaires pour faciliter leur travail et apporter une plus-value à leur pratique. En mettant en place par exemple divers outils de reporting (Tableau VI), aider à l'inclusion dans les protocoles pour aller plus loin dans l'utilisation.

Tableau VI - Attentes des utilisateurs concernant les outils de reporting

Utilisateurs	Attentes
Médecins	- Graphique sur l'évolution des : <ul style="list-style-type: none"> • Scores (i.e. Epworth, dépression, etc.) • Bilan biologiques (i.e. Glycémie, cholestérol) • Signes vitaux (i.e. Pression artérielle) • Données anthropométriques (i.e. taille-poids) - Compte rendu de visite
ARC	-Graphique sur le suivi des inclusions -Export type pour les demandes d'exams -Synthèse de la visite
Administrateurs	-Récapitulatif du nombre d'inclusion dans la base et dans les protocoles -Evolution du nombre d'export fait vers ESADA et l'OSFP -Mise en place d'indicateurs qualité
Secrétaires/Techniciens	<i>Pour le moment ces outils ne leur sont pas utiles</i>

Les centres étrangers aimeraient que la base soit adaptée à leur utilisation en modifiant la langue et les unités.

Les administrateurs de la base souhaiteraient que le processus d'utilisation soit clair et bien compris de tous pour éviter les quiproquos et que chacun soit conscient des tâches qu'il a à réaliser.

De plus, ils aimeraient que des rapports périodiques d'indicateurs/d'outils de reporting soient envoyés aux responsables des centres pour voir l'évolution et suivre les marqueurs qualité (taux de données manquantes, taux de données aberrantes, taux d'incohérences) de la base pour pouvoir apporter des réponses concrètes en cas de dysfonctionnement constaté (écart important d'un mois sur l'autre, taux élevé de données manquantes, etc.).



Figure 16 - Les attentes des utilisateurs concernant la base MARS

Ajouter une partie sur l'exploration ORL

Suite à la demande des professionnels quant à la création d'une partie sur « l'exploration ORL », nous leur avons demandé de nous indiquer leurs attentes et de nous proposer les items qu'ils aimeraient voir apparaître.

A l'heure actuelle, les professionnels travaillent sur ce point pour créer une partie ORL précise et objective pour que l'interprétation et la cotation des échelles soit la même pour tout le monde.

De plus, cette possibilité de créer de nouvelles variables va être étendue de façon plus générale à la base pour pouvoir rajouter facilement une variable nécessaire. Cet outil permettra d'adapter l'utilisation de la base en fonction des besoins des utilisateurs.

Mise en place de supports documentaires

Un système documentaire est déployé et sera directement intégré à la base MARS pour être accessible par tous, à tout moment. En cas de doute, l'utilisateur pourra dans un premier temps se référer au système documentaire et si cela n'est pas suffisant il pourra bien évidemment contacter l'équipe administrative pour de plus amples précisions.

Le système documentaire va s'articuler avec deux types de documents :

- Des processus
- Des procédures

Les processus seront au nombre de 2, un pour l'inclusion dans les protocoles de recherche et un pour l'inclusion dans la cohorte prospective. Ils permettront de représenter de façon visuelle l'ensemble des différentes étapes de saisie des informations, les acteurs et les indicateurs associés.

Le fait de mettre au clair et de documenter les processus permet de clarifier l'ensemble des étapes et les tâches des différents acteurs.

Les procédures seront en support des processus pour apporter davantage de détails sur les différentes étapes.

Par exemple, pour la partie d'inclusion dans les protocoles de recherche, les subtilités et les cas particuliers de chaque protocole seront précisés dans les procédures.

Les deux processus (Annexe 5 et Annexe 6) ont été réalisés avec une équipe pluridisciplinaire (médecin, ARC, techniciens) afin de prendre en compte l'ensemble des acteurs et des tâches. Suite à différentes discussions, des améliorations ont été apportées pour être au plus proche des besoins.

Les procédures seront rédigées à l'aide d'un modèle commun (Annexe 7) et approuvées selon les étapes décrites dans la Figure 17 pour que l'ensemble des procédures de la base MARS soit sous le même format.

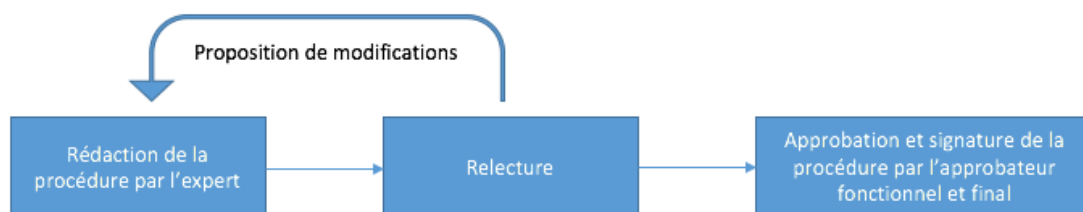


Figure 17 - Etape de rédaction et d'approbation des procédures

Pour que ce système documentaire soit efficace, il est nécessaire de mettre en place un système de suivi de formation via la réalisation d'un formulaire « d'attestation de formation » à la procédure (Annexe 8). En effet, pour que l'utilisation de la base MARS soit faite correctement il est essentiel de mettre en place une traçabilité des formations des utilisateurs.

Pour ce fait, à chaque création et modification de procédure, l'utilisateur devra lire la procédure et attester de sa formation via la signature d'un formulaire « d'attestation de formation » attaché à la procédure. De plus, l'utilisateur devra réaliser un test

(Annexe 9) pour évaluer la qualité de sa formation. Ce test sera corrigé et validé par l'approbateur final de la procédure associée. Si le pourcentage de réponses justes est inférieur à celui autorisé (80%), l'utilisateur se verra dans l'obligation de se réaliser à nouveau la lecture de la procédure et le test de formation.

De plus, un administrateur de la base contrôlera et actualisera le formulaire « d'état de formation des utilisateurs » de la base MARS de façon périodique, tous les 3 mois (Annexe 10). Ce formulaire permettra de vérifier l'état de formation de l'ensemble des utilisateurs de la base MARS à chaque procédure.

Mise en place d'outils de reporting :

Ces outils permettent de faire des représentations graphiques sous différentes formes (i.e. camembert, histogramme, courbes, etc.) pour les utilisateurs avec plusieurs objectifs :

- Montrer de façon visuelle certaines données qui sont saisies dans la base
- Montrer l'utilité de la base
- Suivre l'activité et donc sa performance
- Donner des informations sur la qualité des données et démontrer l'importance de la bonne saisie
- Valoriser le travail des utilisateurs

Dans le cadre de la base MARS, il peut être intéressant de déployer ces outils de reporting sur différentes pages de la base :

- Sur la page d'accueil pour avoir une vision globale des informations saisies. Différents outils de reporting peuvent être judicieux : Une pyramide des âges en fonction du sexe (Figure 18), la représentation sous forme de boîte des

différents antécédents les plus saisis, l'évolution sous forme de courbes des différents indicateurs, l'évolution du nombre d'inclusion dans la base ou par protocole, le nombre de création de « fiche patient » et de visites (Figure 19), etc.

- Sur les visites des patients. Deux types d'outils peuvent être envisagés : un reporting de suivi de la qualité des données et un reporting sur l'évolution des données patients.

Pour le suivi de la qualité, il peut être intéressant dans un premier temps de calculer et de montrer l'évolution des taux de données manquantes/aberrantes et incohérentes dans chaque « fiche patient » en fonction du temps. Ce suivi permettrait de sensibiliser les utilisateurs à la bonne saisie des données mais aussi de valoriser leur travail dans le temps.

Le reporting pourrait aussi être utilisé dans les « fiche patient » pour suivre l'évolution de certains scores/bilan biologiques par exemple. Cependant, ce type de suivi clinique ne rentre pas dans l'objectif de la base MARS. La base ne doit pas se substituer aux dossiers cliniques des patients des centres hospitaliers.

Répartition de l'âge des patients selon le sexe

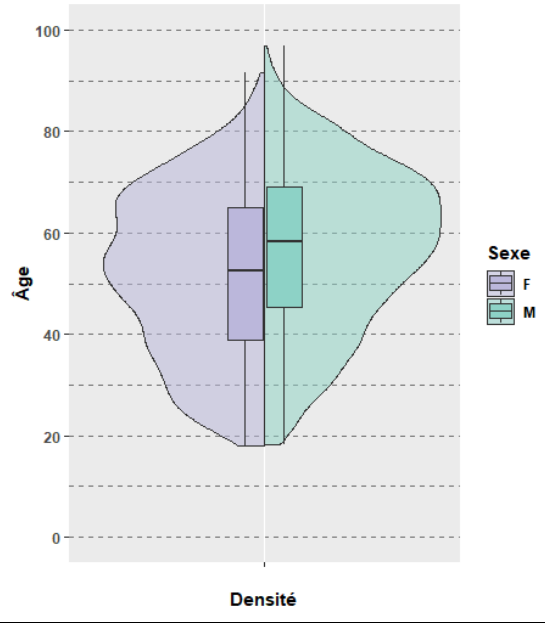


Figure 18 - Répartition de l'âge des patients selon le sexe

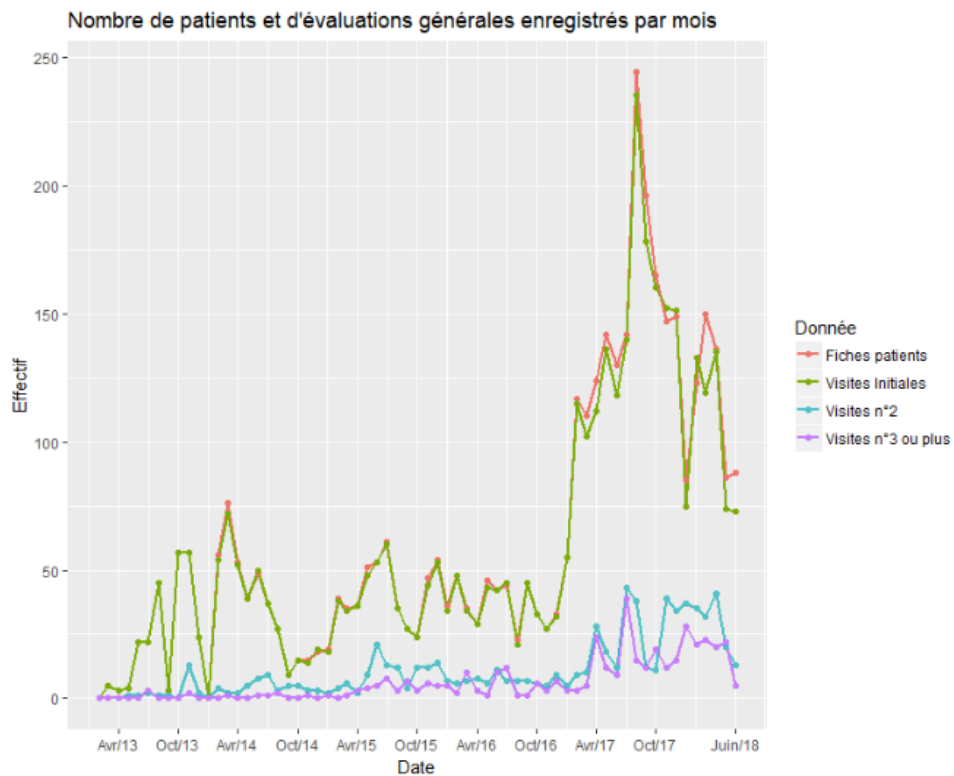


Figure 19 - Nombre de patients et d'évaluations générales enregistré par mois

Aider à l'inclusion dans les protocoles cliniques :

Pour aider à l'inclusion des patients dans les protocoles cliniques des alertes vont être mis en place dans les dossiers patients. En effet, lors de la saisie, si plusieurs critères d'inclusions sont respectés, une alerte s'affichera pour prévenir le médecin que le patient est susceptible d'être inclus dans un protocole clinique.

Par ailleurs, dans le système documentaire, l'ensemble des critères d'inclusion seront mentionnés dans les procédures pour que l'ensemble des professionnels puissent avoir l'information facilement sur les critères d'inclusion/exclusion et sur le descriptif de l'étude.

Faciliter l'accès aux newsletters de modifications :

Après réflexion, des « pop-up » s'afficheront sur la page d'accueil (après identification) contenant l'ensemble des modifications faites sur la base depuis la dernière connexion de l'utilisateur.

Un « pop-up » est une fenêtre secondaire qui s'affiche, sans avoir été sollicitée par l'utilisateur devant la fenêtre de navigation principale lorsqu'on navigue sur Internet.

Cette proposition permettra de proposer des alertes avec plus de visibilité et l'utilisateur pourra plus se concentrer sur ces modifications car il sera déjà engagé dans une démarche d'utilisation de la base.

Adapter la base à l'international :

Pour s'affranchir des problèmes de langage et unités différentes, lors de la connexion sur la plateforme en ligne le changement de ces paramètres se fera automatiquement en fonction du centre sélectionné par les administrateurs. Les futures personnes

portugaises qui saisiront pourront utiliser l'interface automatiquement en portugais et avec les unités de ce centre, sans avoir à faire de modification.

Ce choix automatique via l'identification permettra de s'affranchir d'erreurs pouvant être faites par exemple lors de la sélection manuelle des unités par les utilisateurs lors de la saisie ou lors de la conversion des unités.

A noter que lorsqu'un export de la base sera fait, les unités seront converties automatiquement vers un système commun, à savoir les unités françaises. Cette conversion automatique permettra de ne pas avoir des valeurs exportées avec des unités différentes, permettant ainsi de simplifier le travail de data-management et d'éviter des erreurs de conversion manuelles seront évitées.

4. Zoom sur les attentes des administrateurs: Mise en place d'indicateurs et de rapport d'activité

Mise en place d'indicateurs

Il est essentiel pour l'amélioration de la qualité des données de mettre en place une démarche de suivi par le biais d'indicateurs. Ces indicateurs permettront aux utilisateurs et aux administrateurs d'avoir une idée de la qualité des données dans la base. Les administrateurs pourront, de ce fait, encourager leurs utilisateurs si les résultats sont bons mais aussi mettre en place des actions curatives, correctives et préventives si les résultats ne sont pas satisfaisants.

Dans un premier temps, 3 indicateurs ont été identifiés :

- Taux de données manquantes

Les données manquantes sont les données pour lesquelles il n'y a pas de valeur enregistrée alors que celle-ci aurait dû être saisie. Pour calculer ce taux, il est nécessaire de définir au préalable le lien entre les variables. En effet de nombreuses variables dépendent directement de la valeur d'une autre. Par exemple si le patient n'a pas réalisé de polysomnographie, les résultats de cet examen ne peuvent être renseignés dans la base et il ne faut par conséquent pas les compter comme des données manquantes.

Dans les questionnaires de la base MARS, quand une donnée est manquante, l'utilisateur a la possibilité de cocher une case « DM (Donnée manquante) ». Cette information est enregistrée dans la base. Il y a alors deux calculs différents du taux de données manquantes.

Le premier considère que si l'utilisateur a coché la case « DM » la donnée n'est pas considérée comme manquante car elle est justifiée par l'utilisateur. Il peut arriver que l'information demandée ne soit pas pertinente pour le patient et par conséquent qu'il y ait une raison valable à l'absence de valeur. Comme par exemple, les problèmes d'érection si la patiente est une femme. Cependant si la donnée doit être présente mais que l'utilisateur veut exprimer le fait qu'elle est manquante en cochant cette case, la donnée n'entre pas dans le calcul du taux ce qui fausse celui-ci.

Le second calcul considère qu'une donnée est manquante à partir du moment où

l'absence de valeur est identifiée. Le fait que l'utilisateur coche la case « DM » n'a alors plus d'impact sur le calcul du taux. Ce second calcul peut alors être interprété d'une façon différente. Il représente non seulement le taux de données manquantes mais aussi la pertinence des informations demandées. En effet plus il y aura d'informations demandées non pertinentes avec le cas du patient plus ce taux sera important.

Par ailleurs, certaines variables ont une valeur par défaut qui peut être considérée comme une valeur correcte et par conséquent ces variables ne peuvent pas être comptées dans le calcul du taux de données manquantes. C'est le cas par exemple des antécédents médicaux qui sont enregistrés à « Non » par défaut. On ne peut alors pas savoir si le patient est réellement sain ou si l'antécédent n'a pas été signalé. Le logigramme de l'Annexe 11 présente les étapes du calcul du taux de données manquantes.

Le raisonnement de ce logigramme est de compter le nombre de données présentes et de données manquantes sur l'ensemble des variables considérées puis de calculer le taux en faisant le quotient du nombre de données manquantes sur le nombre de données manquantes plus présentes.

Selon ces étapes, en juillet 2018, les résultats sont de 13.3% de données manquantes suivant le premier calcul (en ne comptant pas les cases « DM » cochées comme des données manquantes). Le taux passe à 20% si on considère une valeur manquante, case « DM » cochée ou non, soit 160.808 données manquantes sur 805.115 que l'on devrait avoir. Par conséquent ces taux posent la question de la pertinence de certaines

données demandées mais surtout de la raison de l'absence de celles qui devraient être présentes.

- Taux de données aberrantes

Les données aberrantes sont les données qui ne sont pas justifiées et qui sont différentes de celles attendues. Ceci peut être le cas quand une donnée n'est pas au bon format (pourcentage) ou si elle sort des bornes de normalité. Pour certaines variables de la base, des bornes sont mentionnées (i.e. poids, taille) ce qui permet d'alerter l'utilisateur via une alerte que la donnée saisie est aberrante. L'utilisateur peut bien évidemment confirmer la valeur saisie si celle-ci est correcte. Par exemple : la taille est exprimée en centimètre. Or il est possible que sur des données papiers, par erreur, la taille soit reportée en mètre. Ainsi si l'utilisateur saisit 1,71 dans la base, une alerte pour donnée aberrante apparaît aussitôt, invitant l'utilisateur à corriger sa saisie, en l'occurrence, entrer 171 cm.

Pour calculer ce taux, il a été nécessaire de borner le maximum de variables numérique. Cependant un compromis est à faire au moment de la déclaration des bornes. En effet une borne trop éloignée des valeurs normales de la variable peut accepter des valeurs qui sont aberrantes et une valeur trop proche de ces valeurs peut entraîner une apparition trop récurrente de l'alerte, une lassitude de l'utilisateur à répondre à la notification et une vérification moins efficace. Ce travail peut aussi être réalisé sur les variables quantitatives en précisant l'ensemble des modalités possibles.

Le calcul du taux ne peut s'effectuer que sur les variables dont la saisie des données est libre et dont les bornes sont définies au préalable. Pour une variable dont la saisie n'est pas libre, comme une donnée à saisir sur un curseur par exemple, il n'est pas

possible de saisir une valeur aberrante et par conséquent la prise en compte de cette variable dans le calcul du taux fausserait celui-ci. Ceci est aussi le cas pour les variables dont les bornes ne sont pas définies.

Par ailleurs certaines valeurs peuvent être en dehors des bornes et être correctes pour des cas particuliers. C'est pour cette raison que la base MARS alerte l'utilisateur dans ce cas et qu'elle lui laisse la possibilité de confirmer ou non l'exactitude de la saisie.

Le calcul de données aberrantes réalisé ne prend en compte que les variables pour lesquelles des bornes ont été définies ce qui montre l'importance de la mise en place de bornes pour un maximum de variables numériques.

Le calcul du taux de données aberrantes suivant cette méthode (Annexe 12) a montré un taux de 7.38% de données aberrantes dans la base, en juillet 2018. Ce taux représente aussi le taux de données qui ne pourront pas être utilisées lors d'une analyse statistique faisant appel aux variables concernées et souligne le fait que des mesures doivent être prises pour diminuer le nombre de valeurs aberrantes enregistrées et corriger les anciennes.

- Taux d'incohérence :

Le dernier indicateur auquel on peut s'intéresser est celui du taux d'incohérence entre les variables. Par exemple si pour un patient une maladie est déclarée dans ses antécédents médicaux mais qu'il n'y a pas de traitements saisis contre cette maladie, on peut se poser des questions sur l'exactitude des données saisies. Ce critère de cohérence ne peut pas être appliqué à l'ensemble des variables de la base mais

l'objectif était de l'étendre au fur et à mesure sur les variables les plus pertinentes pour cet indicateur.

Dès à présent, cet indicateur peut être calculé pour l'antécédent HTA et les traitements antihypertenseurs (Cf. Zoom sur les Traitements). Ce calcul a permis de mettre en place des actions préventives pour éviter la survenue de cette incohérence. Via le déploiement de ces actions, il va être possible de mesurer l'impact de celles-ci sur l'évolution du taux d'incohérence HTA/traitement. Si le taux d'incohérence diminue dans le taux, cela voudra dire que les actions sont efficaces.

Editer des rapports d'activités aux administrateurs :

Suite à la mise en place de ces différents indicateurs et outils de reporting, il est intéressant de réaliser un rapport d'activité aux administrateurs pour qu'ils puissent avoir une idée chiffrée et visuelle de l'activité de la base MARS. Ce rapport peut être édité mensuellement et envoyé à chaque administrateur.

Il peut contenir différentes informations :

- Les indicateurs de qualité : Taux de données manquantes, taux de données aberrantes et taux d'incohérence. Ces taux permettront aux administrateurs de suivre la qualité des données de la base et de mettre en place des actions adaptées : d'encouragements ou curatives/correctives/préventives.
- Les outils de reporting : Pour suivre par exemple le taux d'inclusion dans la base reflétant le bon fonctionnement des différents processus de la base. Et ainsi mettre en place des actions si nécessaires. D'autres outils de reporting peuvent être intégrés à ce rapport pour avoir une vision d'ensemble des informations dans la base.

Un tableau récapitulatif des différentes actions mises en place ou non pour cette thèse est présenté en Annexe 13.

VI. Perspectives d'évolution

Les perspectives d'évolutions pour améliorer la qualité des données de la base MARS peuvent être nombreuses et de différents types :

A. Sur la formation des utilisateurs

Le système documentaire de la base MARS est en cours de déploiement et il faudra absolument que l'ensemble des procédures soient mise à jour pour garantir la bonne formation des utilisateurs et la transmission du savoir-faire. Ce système documentaire ne doit pas rester statique mais doit être en perpétuelle évolution. La gestion de celui-ci peut sembler chronophage mais l'utilisation de la base sera meilleure par la suite et les administrateurs seront gagnant sur la qualité des données.

B. Sur le suivi et l'adhésion des nouveaux centres adhérents

Il est important que les nouveaux centres adhérents soient formés dès leurs arrivées et qu'ils aient le droit à un suivi plus important. En effet, pour certain centre, en étant à des milliers de kilomètres, il peut être difficile de poser ses questions et d'avoir une aide directe de personne maîtrisant l'outil.

C'est pour cela qu'un « Guide de la base MARS » pourrait être crée et donné à chaque nouveau centre et nouveaux utilisateurs.

Ce guide aurait plusieurs objectifs, par exemple :

- Présenter la base MARS et quelques chiffres ;
- Présenter les diverses fonctionnalités de la base (fiche patient, exports, inclusions dans les protocoles, etc.)
- Présenter le système documentaire en place et sensibiliser à la nécessité du suivi de formation ;
- Présenter les divers outils utilisables : Questionnaires numériques/papiers, etc.

Ce guide pourrait être utilisé dans deux cas de figure : premièrement, pour présenter la base aux centres voulant y adhérer, apparenté à un support de communication. Et deuxièmement, aux nouveaux centres adhérents pour que les utilisateurs puissent faire connaissance avec ce nouvel outil sans commencer directement par le système documentaire.

C. Sur la satisfaction des utilisateurs

Pour cette thèse un état des lieux a été réalisé sur les avantages/inconvénients et attentes des utilisateurs. Le questionnaire réalisé a permis de mettre en avant les problèmes les plus sensibles pour pouvoir les améliorer en priorité.

Il pourrait être intéressant, périodiquement, de réaliser un petit questionnaire informatique de quelques questions pour avoir, de la même façon, l'avis des utilisateurs sur l'utilisation de la base et faire remonter ainsi de nouveaux problèmes ou même de nouvelles attentes. Ce questionnaire de satisfaction rentre toujours dans le cadre d'une démarche d'amélioration continue de l'utilisation de la base et de la qualité des données.

D. Sur les problèmes soulevés lors de notre questionnaire

L'ensemble des problèmes rencontrés n'ont pas pu être tous traités lors de cette thèse. Cependant, ceci ne veut pas dire que ces derniers ne seront pas exploités. S'ils ont du sens, ils seront bien évidemment traités dans les mois à venir.

De plus, certaines actions ont ciblé uniquement certaines parties/variables comme l'alerte de cohérence HTA/traitement qui a permis d'améliorer la qualité des données sur ces deux variables. Il sera nécessaire de déployer cette alerte sur les autres antécédents pour palier à d'autres incohérences de ce type.

Cette alerte de cohérence pourrait aussi dans le futur, être reliée avec le bilan biologique pour permettre une cohérence entre l'antécédent, le bilan biologique et les traitements associés.

E. Sur l'utilisation par des centres étrangers

Les paramétrages d'utilisation des centres étrangers se mettront en place à partir du moment où les administrateurs des centres auront sélectionné la langue et les unités adapté à leurs pays dans la partie administrative de la base MARS.

Actuellement la base a été adaptée à l'utilisation portugaise pour le centre de Porto. Il se peut que dans les mois/années à venir, la base soit utilisée par d'autres centres.

Il serait intéressant de prévoir, également, une traduction de la base en anglais qui est la langue internationale. En plus de cette traduction de la base, il me paraît nécessaire, aussi, d'insérer une CIM-10 en anglais pour faciliter la recherche d'antécédents par les centres étrangers. La confusion entre les langues sera moins importante et pourra limiter des erreurs dans la saisie des données.

F. Sur les outils de reporting et rapport d'activité

Certains outils de reporting, indicateurs qualité et rapport d'activité ont été calculés ou créés. Un certain nombre de décisions doivent encore être pris dans les semaines à venir, notamment sur le choix des outils de reporting les plus pertinents et l'organisation du rapport d'activité.

Ces divers outils peuvent, bien sûr, évoluer en fonction de la pertinence d'utilisation et des retours des utilisateurs.

Il est important de noter que certains indicateurs qualité n'ont pas pu être calculés sur l'ensemble des données du fait d'un manque de temps. Il serait donc intéressant de recalculer ces indicateurs en prenant en compte l'ensemble des données de la base pour avoir un indicateur plus proche de la réalité (i.e. données aberrantes : connaître l'ensemble des bornes des variables). De plus, si un intérêt a été retrouvé suite à la mise en place de l'alerte d'incohérence HTA/traitements, les actions devront être mises en place sur les autres antécédents pour limiter le taux d'incohérence.

G. Sur les données d'examens

Dans la « fiche patient » beaucoup de résultats d'examens (bilan biologique, exploration fonctionnelle respiratoire/cardiaque/sommeil, etc.) sont à saisir. Actuellement la majorité de ces données sont saisies manuellement (à l'exception de l'exploration du sommeil pour le centre grenoblois).

En effet, via le questionnaire, il a été mis en évidence que ces résultats d'examens n'étaient pas toujours mis à jour par manque de temps. Dans un premier temps, il faut bien sûr sensibiliser le personnel sur l'importance d'avoir ces données dans les « fiches patients » mais il serait aussi intéressant de faire des exports informatiques des

données directement depuis le logiciel médical du centre hospitalier. Ceci demande beaucoup de travail informatique mais permettrait de s'abstenir des erreurs de retranscription possibles.

VII. Conclusion

THÈSE SOUTENUE PAR : Virginie SUIZE

TITRE :

MISE EN ŒUVRE DU PROCESSUS QUALITÉ D'UNE BASE DE DONNÉES MÉDICALES.
APPLICATION À LA BASE MARS DU CENTRE HOSPITALIER UNIVERSITAIRE GRENOBLE-ALPES

CONCLUSION :

Bien que dans la hiérarchie des niveaux de preuve, la recherche expérimentale reste le niveau le plus élevé, et l'essai clinique randomisé contrôlé la référence, les recherches sur les bases de données observationnelles sont maintenant reconnues comme pouvant apporter des réponses complètes à des questions cliniques complexes. Cela a été rendu possible principalement par l'évolution des méthodes de gestion des bases de données qui a conduit à une structuration de l'information médicale.

Cependant, il n'est pas suffisant d'avoir uniquement des données médicales structurées, mais il importe que ces données soient de qualité de façon à éviter les biais liés aux sources d'informations qui présentent un problème majeur dans la littérature scientifique internationale.

Ainsi la mise en œuvre d'une démarche qualité continue doit permettre de garantir la qualité de ces données, et cela, dès la première étape de saisie de l'information. Les objectifs d'une telle démarche sont de proposer aux utilisateurs divers outils pour harmoniser les pratiques de saisie, de mettre en place des actions de formation, et d'assurer le suivi continu de la qualité par l'édition et la publication de différents rapports d'indicateurs qualité et d'outils de reporting. Un tel système doit permettre un

accompagnement des utilisateurs dans un environnement essentiellement pluridisciplinaire, avec des utilisations et des attentes différentes. Avoir à disposition un système documentaire écrit permet d'assurer la transmission du savoir-faire entre les utilisateurs et prendre en compte l'évolution du personnel.

Depuis septembre 2017, j'ai pris en charge la mise en place du processus de qualité pour la base de données MARS dans un contexte réglementaire évolutif. La gestion de l'ensemble de ce projet a pu se faire grâce à la mobilisation d'acteurs de différentes disciplines (informaticiens, médecins, pharmaciens, biostatisticiens, datamanager et ARCs). Cela a permis de faire évoluer la base pour s'inscrire dans le contexte du RGPD, pour répondre aux exigences de qualité attendues par les utilisateurs et les administrateurs. A ce jour, la majorité des outils sont en place et pour les derniers, l'objectif est de finaliser le déploiement pour le premier semestre 2019. Au-delà du déploiement des outils, la dynamique d'un processus qualité n'a de sens que si les acteurs sont impliqués et mobilisés au quotidien. C'est pour cela que j'ai proposé d'accompagner, dans les semaines à venir, la mise en place des démarches de rédaction de procédures et des actions de formations qui vont se déployer dans les différents centres. En effet, le management de la qualité passe d'abord par le management des acteurs du système qualité, et donc passe par l'animation permanente de ce système pour pouvoir le faire évoluer avec les pratiques.

En ce sens, le pharmacien peut avoir différents rôles dans la gestion d'une base de données médicale. Le premier est qu'il est un acteur pivot entre les différents acteurs de saisie de l'information médicale, jouant le rôle d'interface entre des connaissances médicales et l'interface de saisie. Le deuxième est le contrôle et l'expertise sur les données médicales (traitements, antécédents, etc.), en complément du médecin, permettant la mise en œuvre d'une démarche qualité adaptée aux données. L'application de tout ce processus pour

améliorer la qualité de la base MARS en est une bonne illustration.

VU ET PERMIS D'IMPRIMER

Grenoble, le : 15 octobre 2018

LE DOYEN



Pr. Michel SEVE
et par délégation

Le Doyen de Pharmacie
Pr. Michel SEVE

LE PRÉSIDENT DE LA THÈSE



Dr. Marie JOYEUX-FAURE

VIII. Bibliographie

1. ConsoGlobe. Informations publiées dans le monde sur le net (en Gigaoctets). Available at: <https://www.planetoscope.com/Internet-/1523-informations-publiees-dans-le-monde-sur-le-net-en-gigaoctets-.html>. (Accessed: 14th September 2018)
2. Ffoulkes P. InsideBIGDATA guide to the intelligent use of big data on an industrial scale. (2017).
3. Bailly S, Meyfroidt G, Timsit JF. What's new in ICU in 2050: big data and machine learning. *Intensive Care med* **44**, 1524–1527 (2017).
4. Vincent J-L. Critical care - where have we been and where are we going? *Critical Care*. 2013;17(Suppl 1):S2.
5. Black N. Why we need observational studies to evaluate the effectiveness of health care. *BMJ* **312**, 1215–1218 (1996).
6. Friedrich K. Port, F. K. Port 2000 *Kidney international* Role of observational studies.pdf. *Kidney international* **57**, 3–6 (2000).
7. Angus DC. Fusing randomized trials with big data: The key to self-learning health care systems? *JAMA* **314**, 767–768 (2015).
8. Fox Trial Finder. Comprendre les essais cliniques. Available at: <https://foxtrialfinder.michaeljfox.org/fr/understanding-clinical-trials/clinical-trials-101/>. (Accessed: 23th September 2018).
9. Port, F. K. Role of observational studies versus clinical trials in ESRD research. *Kidney International* **57**, S3–S6 (2000).
10. Robins JM, Hernan MA, Brumback B, (2000) Marginal structural models and causal inference in epidemiology. *Epidemiology* 11: 550-560.
11. Bailly S, Bouadma L, Azoulay E, Orgeas MG, Adrie C, Souweine B, Schwe - bel

C, Maubon D, Hamidfar - Roy R, Darmon M, Wol M, Cornet M, Timsit JF. Failure of empirical systemic antifungal therapy in mechanically ventilated critically ill patients. *Am J Respir Crit Care Med* **191**, 1139–46 (2015).

12. Timsit JF, Azoulay E, Schwebel C, Charles PE, Cornet M, Souweine B, Klouche K, Jaber S, Trouillet JL, Bruneel F, Argaud L, Cousson J, Meziani F, Gruson D, Paris A, Darmon M, Garrouste - Orgeas M, Navellou JC, Foucrier A, Allaouchiche B, Das V, Gangneux JP, Ruckly S, Maubon D, Jullien V, Wol M. Empirical micafungin treatment and survival without invasive fungal infection in adults with ICU - acquired sepsis, candida colonization, and multiple organ failure: the EMPIRICUS randomized clinical trial. *JAMA* 316:1555-1564. *JAMA* **316**, 1555–1564 (2016).

13. Deschasaux M, Huybrechts I, Murphy N, Julia C, Hercberg S, Srouf B, Kesse-Guyot E, Latino-Martel P, Biessy C, Casagrande C, Jenab M, Ward H, Weiderpass E, [...], Touvier M. Nutritional quality of food as represented by the FSAm-NPS nutrient profiling system underlying the Nutri-Score label and cancer risk in Europe: Results from the EPIC prospective cohort study. *PLOS Medicine* (2018).

14. Système national des données de santé. Qu'est-ce que le SNDS ? Available at: <https://www.snds.gouv.fr/SNDS/Qu-est-ce-que-le-SNDS>. (Accessed: 13th September 2018)

15. Lévy P, Kohler M, McNicholas W, Barbé F, McEvoy R, Somers V, Lavie L, Pépin JL. Obstructive sleep apnoea syndrome. *Nature Reviews Disease* **1**, 1–20

16. Lévy P, Pépin JL, Demattéis M, Viot-Blanc V, Bettega G, Tamisier R. Physiopathologie du SAS et de ses conséquences. Available at: http://www.sfrms-sommeil.org/documents/DIU2008/DIU2008-Levy-PhysiopathSAS_csq.pdf. (Accessed: 13th September 2018)

17. Javaheri S, Barbe F, Campos-Rodriguez F, Dempsey J, & Khayat R, Javaheri S, Malhotra A, Martinez-Garcia M, Mehra R, Pack A, Polotsky V, Redline S, Somers V. Sleep Apnea Types, Mechanisms, and Clinical Cardiovascular Consequences. *Journal of the american college of cardiology* 841–58 (2017).
18. Fuhrman, C., Fleury, B., Nguyễn, X.-L. & Delmas, M.-C. Symptoms of sleep apnea syndrome: High prevalence and underdiagnosis in the French population. *Sleep Medicine* **13**, 852–858 (2012).
19. Knauert, M., Naik, S., Gillespie, M. B. & Kryger, M. Clinical consequences and economic costs of untreated obstructive sleep apnea syndrome. *World J Otorhinolaryngol Head Neck Surg* **1**, 17–27 (2015).
20. Sassani A1, Findley LJ, Kryger M, Goldlust E, George C, Davidson TM. Reducing motor-vehicle collisions, costs, and fatalities by treating obstructive sleep apnea syndrome. *Sleep* **27**, 453–8
21. Inserm. La recherche clinique. Available at: <https://www.inserm.fr/recherche-inserm/recherche-clinique>. (Accessed: 17th September 2018)
22. Groupement interregional de recherche clinique et d'innovation. Qu'est ce que la recherche clinique ? Available at: <https://www.girci-soom.fr/content/quest-ce-que-la-recherche-clinique>. (Accessed: 9th September 2018)
23. DEMERVILLE. Nouvel encadrement des recherches médicales - Loi Jardé. Available at: https://www.recherchecliniquepariscentre.fr/wp-content/uploads/2017/02/jarde_formation-janvier2017_LDemerville.pdf. (Accessed: 18th September 2018).
24. QUINTIN. La recherche impliquant la personne humaine. Available at: <https://www.recherchecliniquepariscentre.fr/wp-content/uploads/2017/03/DIU-CP->

- Classification-des-essais-17-02-2017-C.-Quintin.pdf. (Accessed: 18th September 2018).
25. Agence nationale de sécurité du médicament et des produits de santé. Essais cliniques. Available at: [http://ansm.sante.fr/Activites/Essais-cliniques/Obtenir-un-numero-d-enregistrement/\(offset\)/4](http://ansm.sante.fr/Activites/Essais-cliniques/Obtenir-un-numero-d-enregistrement/(offset)/4). (Accessed: 23th September 2018).
26. QUINTIN. Les intervenants de la Recherche Clinique. Available at: https://www.recherchecliniquepariscentre.fr/wp-content/uploads/2012/01/coursDIU-INFTEC_Les-intervenants-de-la-RC_2012_CQuintin.pdf. (Accessed : 16th September.
27. CNIL. MR-003 Méthodologie de référence. Available at: <https://www.cnil.fr/sites/default/files/atoms/files/mr-003.pdf>. (Accessed: 17th September 2018).
28. CNIL. Ce qui change avec les nouvelles méthodologies. Available at: <https://www.cnil.fr/fr/recherches-dans-le-domaine-de-la-sante-ce-qui-change-avec-les-nouvelles-methodologies-de-reference>. (Accessed: 14th September 2018)
29. ASTIER, P. S. & Cour, A. à la. #RGPD : vos obligations avant/après. *HAAS Avocats* (2017). Available at: <https://www.haas-avocats.com/actualite-juridique/rgpd-vos-obligations-avantapres/>. (Accessed: 17th September 2018)
30. LegalPlace. Quelles sanctions en cas de non respect du RGPD ? (2018). Available at: <https://www.legalplace.fr/guides/rgpd-sanction/>. (Accessed: 17th September 2018)
31. Edition Francis Lefebvre. RGPD : quels impacts pour les entreprises et leurs salariés à partir du 25 mai 2018 ? Available at: <https://www.eff.fr/actualites/social/contrat-de-travail/details.html?ref=ui-f5a39cbb-c7e3-4583-a218-9114d6280a4d>. (Accessed: 17th September 2018)
32. Choong Ho Lee. Medical Big data : promise and challenges. *Kidney Research and Clinical Practice*, 2017, 36:3-11.

33. Timsit JF, Nakache D, Alberti C, Metais E. JF Timsit. La Recherche sur les bases de données en réanimation. *Réanimation*, 2006, 15(4):313-9. *Réanimation* **15**, 313–319 (2006).
34. AUBIN. Indicateurs de performance & Qualité des données : Vers une démarche industrielle dans un grand Hôpital Français. Available at: <https://hal.archives-ouvertes.fr/hal-01492623/document>. (Accessed: 13th September 2018).
35. Le webzine de la HAS. Qualité des soins et bases de données - Epithor, Epicard. Available at: https://www.has-sante.fr/portail/jcms/c_1122438/fr/qualite-des-soins-et-bases-de-donnees-epithor-epicard. (Accessed: 1st October 2018).
36. Listflow. Data Quality - Définition & bonnes pratiques en marketing B2B. (2018). Available at: <https://listflow.io/fr/blog/data-quality-definition/>. (Accessed: 28th September 2018)
37. Hilson D. Les données d'entrée et de sortie du Management des Risques. Available at : <http://www.risk-doctor.com/pdf-briefings/risk-doctor05f.pdf>. (Accessed : 14 october 2018).
38. Ministère des affaires sociales et de la santé. Classification statistique internationale des maladies et des problèmes de santé connexes. Available at : https://www.atih.sante.fr/sites/default/files/public/content/2665/cim10_2015_final_0.pdf. (Accessed :14th October 2018).

IX. Annexes

Annexe 1 - Questionnaire d'état des lieux de la base MARS.....	99
Annexe 2 - Réponses des utilisateurs sur les avantages et inconvénients de la base MARS.....	101
Annexe 3 - Réponses des utilisateurs concernant les problèmes d'encodage traité lors d'investigations futures.....	103
Annexe 4 – Extrait du dictionnaire CIM-10 d'aide à la saisie des évènements généraux les plus courants.....	105
Annexe 5 - Processus d'inclusion de patient dans les protocoles de recherche du CHUGA.....	107
Annexe 6 - Processus d'inclusion des patients dans la base MARS venu pour une consultation.....	108
Annexe 7 – Modèle proposé pour les procédures base MARS CHUGA.....	109
Annexe 8 – Modèle proposé pour le formulaire "d'attestation de formation" à la procédure.....	112
Annexe 9 – Modèle proposé pour le quizz d'attestation de formation	113
Annexe 10 - Modèle proposé pour le formulaire "d'état de formation des utilisateurs à la base MARS"	115
Annexe 11 - Logigramme données manquantes.....	116
Annexe 12 - Logigramme données aberrantes	117
Annexe 13 - Récapitulatif des actions réalisées et futures.....	118

Annexe 1 - Questionnaire d'état des lieux de la base MARS

QUESTIONNAIRE BASE MARS :

Nom/Prénom :

Questionnaire :

1. Tout d'abord, savez-vous à quoi sert cette base ?

2. Quelle est votre utilisation de la base MARS ? L'utilisez vous pour le suivi clinique ou de protocole ? L'utilisez-vous régulièrement ?

3. Qu'attendez-vous de cette base/outil ?

4. Quels sont les avantages généraux de cette base/outil ?

5. Quels sont les inconvénients généraux de la base ?

6. Si nous reprenons point par point les différents éléments de la base MARS, pouvez-vous me dire les différents problèmes (ou points positifs) que vous rencontrez ?
 - Antécédents personnels
 - Evènements médicaux (CIM10)
 - Facteurs de risques
 - Traitements
 - Antécédents familiaux

- Sommeil (sommeil, habitudes, évaluation clinique des troubles respiratoires nocturnes)
- Paramètres cliniques (données anthropométriques, signes vitaux, bilan biologique, EFR, exploration cardiaque, exploration du sommeil et décision thérapeutique)

7. Quels sont les points d'amélioration que vous attendez ?

8. Etes vous satisfait des différents outils mis en place pour éviter les erreurs de saisie ? (Alertes, etc.). Aimeriez-vous d'autres outils à votre disposition ?

9. Utilisez vous les outils de reporting ? Qu'attendez-vous de ces outils ? (Informations récoltées, organisation, etc.). Quels outils vous sembleraient pertinent ?

10. Sentez vous le besoin d'être formé à l'utilisation de la base ou de recevoir plus d'informations ? Seriez-vous prêt à accorder du temps pour suivre des formations ?

Conclusion :



Annexe 2 - Réponses des utilisateurs sur les avantages et inconvénients de la base MARS

Questions	Réponses des utilisateurs	Propositions d'améliorations par les utilisateurs
Quels sont les avantages généraux de cette base/outil ?	Outil synthétique pour le dossier patient (+++ Traitement et antécédents)	∅
	Bonne structure de la base	
	Aide à la dictée des comptes rendus médicaux	
	Très bon questionnaire	
	Facile d'utilisation	
	Vision d'ensemble des patients	
	Possibilité d'export pour projet multicentrique : OSFP, ESADA	
	Utilisation tablette : gain de temps	
Quels sont les inconvénients généraux de la base ?	Manque de champs libres sur les différentes parties de la base	Ajouter Champs libre tout au long de la base
	Peu d'informations sur les « troubles ventilatoires »	Ajouter plus d'items à ce sujet
	Echelle de score parfois subjective : Ex : Syndrome de Jambes Sans Repos	Préciser les échelles à utiliser
	Méthodologie pour encoder la base : revenir sur la visite	∅
	Pas de vérification de la saisie : pas de vérification de la qualité des données	∅
	Pas de possibilité de poser des questions	Zone libre dans chaque rubrique pour poser sa question
	Impossible de retrouver ses propres patients dans la base	∅
	Problème de connexion pour les exports. Ex : ESADA	∅
	Remplissage de l'état civil long et fastidieux	∅
	Reformulation avec patient	∅

	Difficile de faire remplir le questionnaire aux patients	∅
	Impossible de supprimer une « fiche patient » erronée	∅
	L'ensemble des fonctionnalités de la base n'est pas connu	Formation à la base
	Cohérence de certaines données : Antécédents et traitements	∅
	Mauvais import de donnée sur le « PLM pendant le TTS »	∅
	Langue différentes dans les pays adhérents	Traduction de la base

Annexe 3 - Réponses des utilisateurs concernant les problèmes d'encodage traité lors d'investigations futures

Questions	Réponses des utilisateurs	Propositions d'améliorations par les utilisateurs
Facteurs de risques	Le sport n'est pas un facteur de risque	Le déplacer
	Obligé de cocher Facteur de risque « Oui » si pratique de sport et sans alcool/tabac : illogique	∅
	Manque notion de fréquence hebdomadaire pour le sport	Ajouter une notion de fréquence hebdomadaire pour le sport. Intégrer un vrai questionnaire pour le sport.
	Alcool : Manque notion de fréquence par jour	Rajouter une notion de fréquence par jour ?
Antécédents familiaux	Mauvaise organisation	Les rafraichir dans la visite 2. Les déplacer dans les facteurs de risques.
Sommeil (sommeil, habitudes, évaluation clinique des troubles respiratoires nocturnes)	Termes compliqués pour le patient dans le questionnaire papier	Changer les termes
	Manque notion sur les troubles de la concentration	Le rajouter
	Troubles libido/érection : Peu abordé	Utile ?
	Motif de consultation	Avoir la possibilité de sélectionner plusieurs causes ?
	Non harmonisation des scores	Faire un résumé sur les score à utiliser
	Données très peu saisies (Portugal)	∅
	Echelle de dépression de Beck	La rajouter

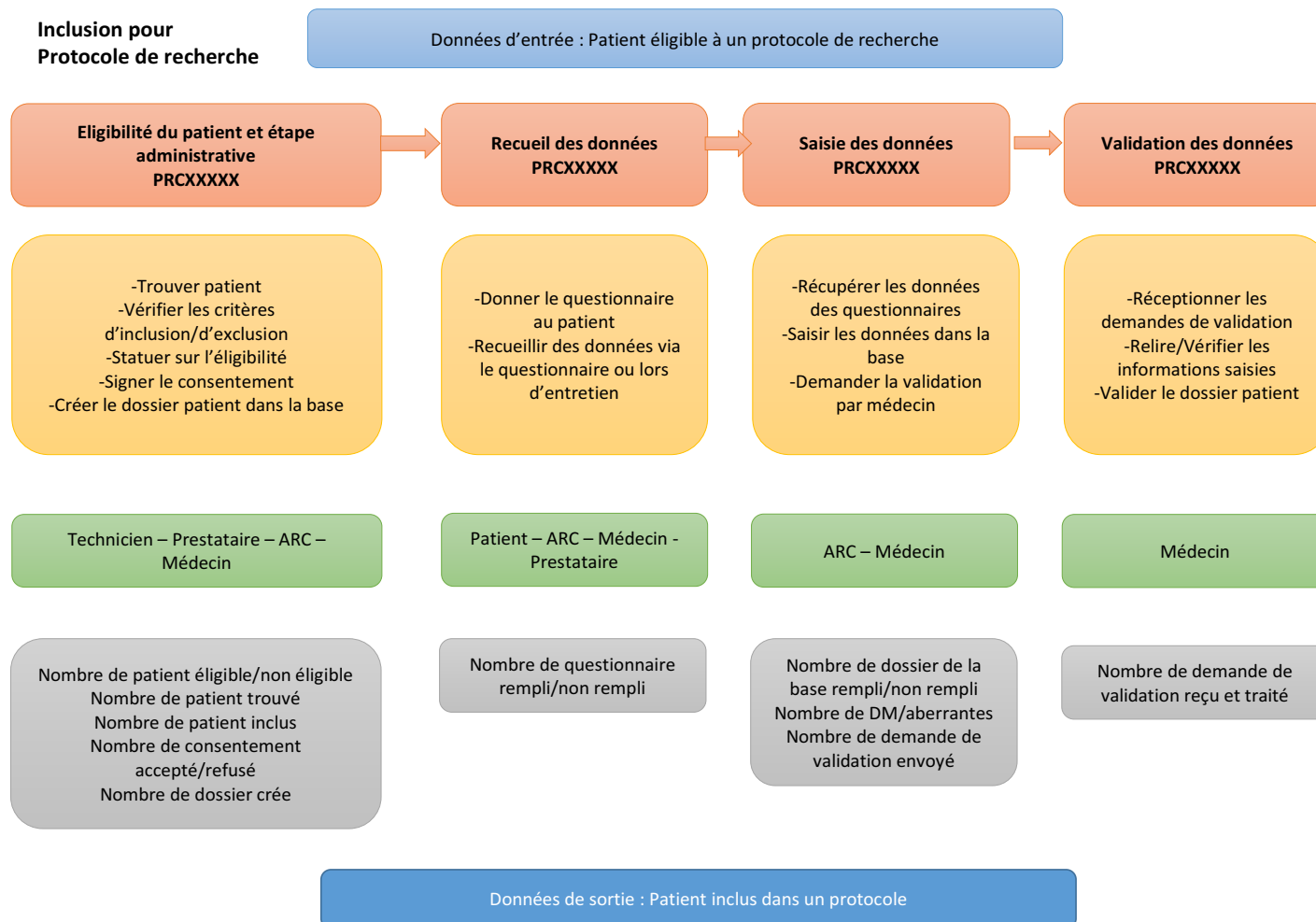
	Habitudes	Rajouter heure de réveil ? Mettre des plages horaires ?
	Réarrangement de certaines parties	∅
	Scores : Différence entre interprétation patient et médecin	∅
	Partie B ne donne pas envie d'être remplie	∅
	Permis : mauvaise connaissances des permis	Reprendre et redéfinir cette partie
Paramètres cliniques (données anthropométriques, signes vitaux, bilan biologique, EFR, exploration cardiaque, exploration du sommeil et décision thérapeutique)	A partir de la partie des « signes vitaux » : Peu exploité par manque de temps	Faire des importation des résultats
	« Profil maxilo-facial semble être un facteur SAS » : Peu adapté, et très subjectif du médecin	Mettre des items plus spécifiques
	Bilan biologique : Non encodé pour le moment	∅
	Exploration cardiaque : Peu renseigné	∅
	Bilan biologique : Pas les mêmes unités : mg/dl	Convertir ou sélectionner l'unité ?
	Données très peu saisies (Portugal + CHAL)	∅

Dictionnaire CIM-10 :

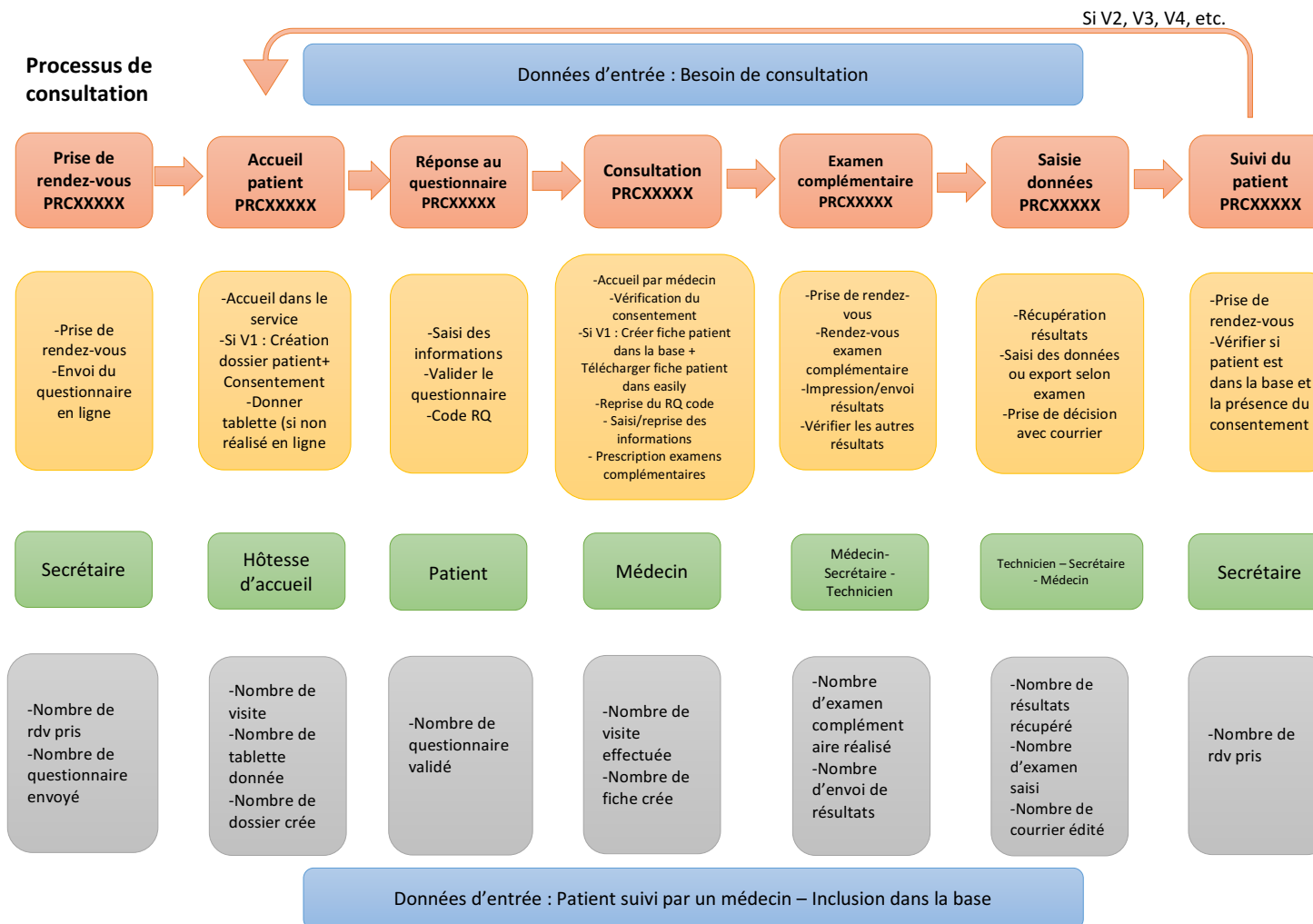
Maladies endocriniennes, nutritionnelles et métaboliques	
Hypothyroïdie	E039
Hyperthyroïdie	E059
Thyroïdite auto-immune	E063
Diabète de type I	E109
Diabète de type II	E119
Pied diabétique	E145
Obésité	E669
Hypercholestérolémie	E780
Hypertriglycéridémie	E781
Hyperuricémie	E790
Maladies de l'appareil circulatoire	
HTA	I10
Hypotension	I959
Cardiopathie ischémique	I259
HTAP	I270
Angine de poitrine	I209
Infarctus du myocarde	I219
Embolie pulmonaire	I269
BAV	I443
FA	I48
AVC	I64
Insuffisance coronaire	I251
Insuffisance cardiaque	I509
Insuffisance mitrale	I340
Insuffisance aortique	I351
Insuffisance veineuse	I872
Fibrillation et flutter auriculaire	I48
Trouble du rythme	I499
Hypertrophie du ventricule G	I517
Artériopathie	I702
Athérosclérose	I709
Anévrisme	I729

Syndrome de Raynaud	I730
Embolie et thrombose	I749
Phlébite et thrombophlébite	I809
Varice ulcérée	I830
Hémorroïdes	I849
Maladies de l'appareil respiratoire	
BPCO	J440
Rhinite allergique	J304
Asthme	J459
Œdème pulmonaire	J81
Pneumothorax	J939
SAS	G473
Insuffisance respiratoire	J969
Maladie pulmonaire restrictive	J961
Maladies du système nerveux	
SAS	G473
Neuropathie du système nerveux autonome au cours de maladies endocriniennes et métaboliques	G990
Maladies de l'appareil génito-urinaire	
Insuffisance rénale aiguë	N179
Colique néphrétique	N23
Insuffisance rénale chronique	N189
Insuffisance rénale	N19
Goutte	N228
Affection prostatique	N419
Hyperplasie de la prostate	N40
Affection de la vessie	N327
Pyélonéphrite	N111
Néphropathie diabétique	N083
Symptômes et signes relatifs aux appareils circulatoire et respiratoire	
Tachycardie	R000
Bradycardie	R001
Palpitation	R002

Annexe 5 - Processus d'inclusion de patient dans les protocoles de recherche du CHUGA



Annexe 6 - Processus d'inclusion des patients dans la base MARS venu pour une consultation



Annexe 7 – Modèle proposé pour les procédures base MARS CHUGA



PROCÉDURE BASE MARS PRC XXXXX

Titre de la procédure	
Service	Choisissez un élément.
Référentiel	

Rôle / Département	Les dates et signatures sont imprimées à la fin du document
Auteur	Choisissez un élément.
Approbateur fonctionnel	Choisissez un élément. Choisissez un élément.
Approbateur Final	Choisissez un élément.

Historique / Document History

Version	Modifications
«v1»	Création

Table des matières / Contents

1	Objectif	2
2	Vocabulaire Spécifique	2
3	Descriptif de la Procédure	2
	3.1 Domaine d'application.....	2
	3.2 Description	2
4	Cas Particuliers (Interdictions, Exceptions).....	2
5	Responsabilités	2
6	Références	2
	6.1 Références Normatives et Légales	2
	6.2 Autres Références	2
7	Autres Rubriques	2
8	Documentation de Suivi	3
	8.1 Enregistrement des Données	3
9	Annexes au Descriptif de la Procédure.....	3

1 Objectif

2 Vocabulaire Spécifique

Vocabulaire	Définition

3 Descriptif de la Procédure

3.1 Domaine d'application

3.2 Description

4 Cas Particuliers (Interdictions, Exceptions)

(Cas particulier entre les V1/V2 ou entre les protocoles de recherches)

5 Responsabilités

(Responsabilités à chaque étapes)

6 Références

6.1 Références Normatives et Légales

(Référence à des référentiel/normes)

6.2 Autres Références

(Référence à d'autres procédure liées)

7 Autres Rubriques




8 Documentation de Suivi

8.1 Enregistrement des Données


(Si données enregistrées préciser le logiciel de traitement : Par exemple pour les indicateurs : dans la base MARS ou Excel)

9 Annexes au Descriptif de la Procédure

Annexe 8 – Modèle proposé pour le formulaire "d'attestation de formation" à la procédure

			
Attestation de formation Base MARS :			
Numéro de procédure : PRC			
Nom	Prénom	Date de formation	Signature

Annexe 9 – Modèle proposé pour le quizz d'attestation de formation

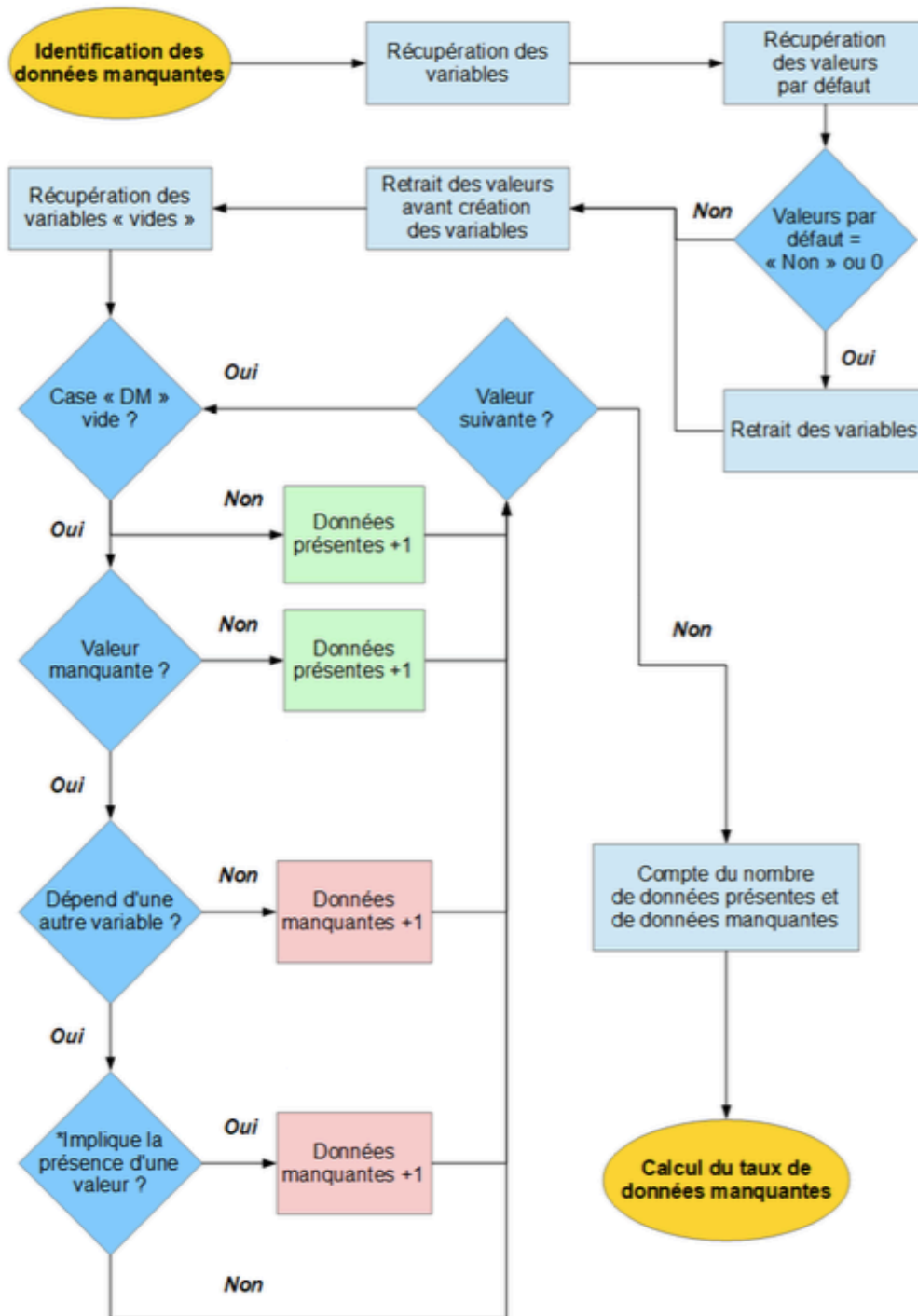
		
QCM attestant la formation à la Base MARS :		
Numéro de procédure associée : PRC		
Nom/Prénom de l'utilisateur :		
Questions :	Réponse du collaborateur	Réponse du formateur
Cochez les réponses exactes :		
1...	<input type="checkbox"/>	<input type="checkbox"/>
2...	<input type="checkbox"/>	<input type="checkbox"/>
3...	<input type="checkbox"/>	<input type="checkbox"/>
4...	<input type="checkbox"/>	<input type="checkbox"/>
Cochez les réponses exactes :		
1...	<input type="checkbox"/>	<input type="checkbox"/>
2...	<input type="checkbox"/>	<input type="checkbox"/>
3...	<input type="checkbox"/>	<input type="checkbox"/>
4...	<input type="checkbox"/>	<input type="checkbox"/>
Cochez les réponses exactes :		
1...	<input type="checkbox"/>	<input type="checkbox"/>
2...	<input type="checkbox"/>	<input type="checkbox"/>
3...	<input type="checkbox"/>	<input type="checkbox"/>
4...	<input type="checkbox"/>	<input type="checkbox"/>
Cochez les réponses exactes :		
1...	<input type="checkbox"/>	<input type="checkbox"/>
2...	<input type="checkbox"/>	<input type="checkbox"/>
3...	<input type="checkbox"/>	<input type="checkbox"/>
4...	<input type="checkbox"/>	<input type="checkbox"/>
Signature de l'utilisateur		
Date :		

Questions :	Réponse du collaborateur	Réponse du formateur
Cochez les réponses exactes :		
1...	<input type="checkbox"/>	<input type="checkbox"/>
2...	<input type="checkbox"/>	<input type="checkbox"/>
3...	<input type="checkbox"/>	<input type="checkbox"/>
4...	<input type="checkbox"/>	<input type="checkbox"/>
Cochez les réponses exactes :		
1...	<input type="checkbox"/>	<input type="checkbox"/>
2...	<input type="checkbox"/>	<input type="checkbox"/>
3...	<input type="checkbox"/>	<input type="checkbox"/>
4...	<input type="checkbox"/>	<input type="checkbox"/>
Cochez les réponses exactes :		
1...	<input type="checkbox"/>	<input type="checkbox"/>
2...	<input type="checkbox"/>	<input type="checkbox"/>
3...	<input type="checkbox"/>	<input type="checkbox"/>
4...	<input type="checkbox"/>	<input type="checkbox"/>
Signature de l'utilisateur :		
Date :		

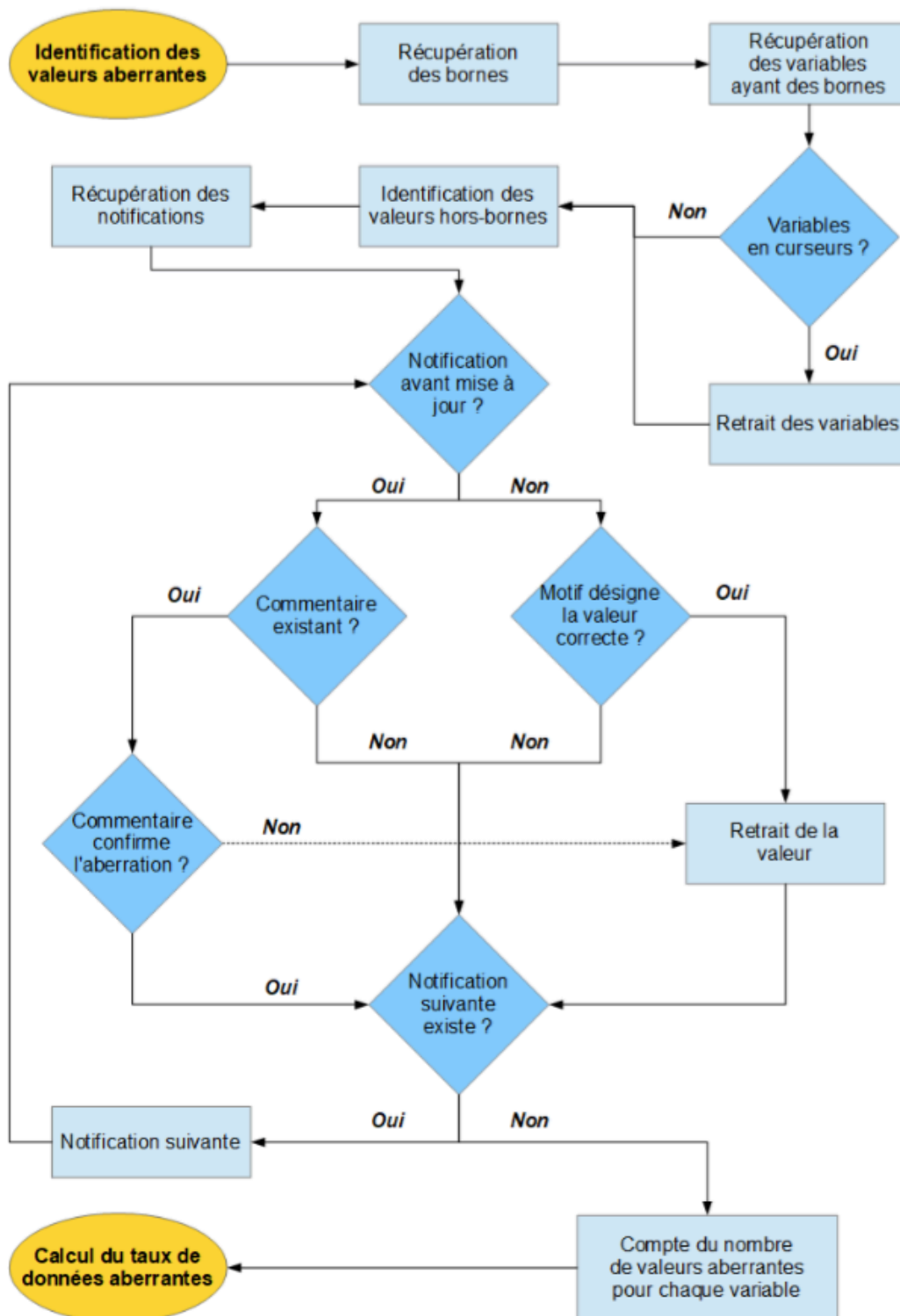
Nombre de réponse juste :
Pourcentage de réponse juste :
Validation de la formation : Oui / Non

Nom-prénom du formateur :
Signature du formateur :

Annexe 11 - Logigramme données manquantes



Annexe 12 - Logigramme données aberrantes



Annexe 13 - Récapitulatif des actions réalisées et futures

	Avis des utilisateurs	Etat des lieux : Réalisé (R)- Futur (F) Non réalisé (NR)	Justification des administrateurs
Les problèmes liés à l'utilisation de la base	Manque de champs libres sur les différentes parties de la base	NR	Un champ libre est disponible à la fin du questionnaire. Multiplier les champs libres est à l'origine d'un travail de data management conséquent. Le dossier patient de la base MARS reste un outil de travail et ne substitue pas au dossier médical du patient des centres hospitaliers
	Peu d'informations sur les « troubles ventilatoires »	F	
	Echelle de score parfois subjectives	NR	L'ensemble des scores doivent être évalué avec les échelles internationales pour avoir une harmonisation des données
	Méthodologie pour encoder la base chronophage : revenir sur la visite à chaque fois qu'il faut entrer de nouvelles données	NR	Il est nécessaire de revenir à chaque fois sur la visite pour saisir de nouvelles données patient
	Pas de vérification sur la qualité de la saisie	NR	Possibilité de demander au médecin de vérifier la « fiche patient ». En faire la demande via la base
	Pas de possibilité de poser des questions	F	Possibilité de poser les questions aux administrateurs par mail. Dans le futur une plateforme sur la base MARS sera à disposition pour rédiger et envoyer directement les questions aux personnes habilitées
	Impossibilité de retrouver ses	NR	Recherche de ses propres patients possibles. Ceci sera

	propres patients dans la base		expliquer dans le système documentaire mis en place.
	Problème de connexion pour les exports Ex : ESADA	F	
	Remplissage de l'état civil long et fastidieux	F	Améliorer l'organisation du service pour avoir une meilleure ergonomie. Rentrer en contact avec les personnes concernées
	Reformulation obligatoire avec le patient → Chronophage	NR	La reformulation avec le patient de l'ensemble des données saisies dans le questionnaire est essentielle et nécessaire pour vérifier l'exactitude des données saisie par le patient
	Difficile de faire remplir le questionnaire aux patients	F	Sensibiliser le patient de l'importance de la réponse au questionnaire lors de la prise de rendez-vous
	Impossibilité de supprimer des « fiches patient » erronées	NR	Seuls les administrateurs peuvent effectuer cette action. Ceci permet d'éviter les suppressions involontaires et la perte d'information
	Certaines fonctionnalités de la base sont non connues	R	Un système documentaire va être mis en place avec un suivi de formation
	Incohérence entre les données : Antécédents – traitements – bilans biologiques	R pour l'HTA F pour le reste	Mise en place d'alerte de cohérence lors de la saisie des données.
	Date des antécédents (ATCD) non connue par le personnel	R	Demander ce type d'information aux patients et/ou médecin traitant et/ou pharmacien en amont
	La date de survenue de l'ATCD s'affiche une fois que la visite à été	R	Modification sur le système informatique

sauvegardé		
Certaines catégories d'ATCD ne sont pas définies. Ex : Maladies neurologiques	R	Catégories explicitées dans le système documentaires
Peu de proposition d'ATCD prédéfinies	NR	Cette partie permet une présélection rapide des ATCD les plus fréquents. Augmenter cette liste n'aura plus d'intérêt
Parfois difficile de trouver les ATCD que l'on veut dans la CIM-10	R	Création d'un dictionnaire d'aide à la saisie CIM-10
Les ATCD ne s'affichent pas dans les visites de suivis	NR	Le non report des ATCD est normal. Les ATCD qu'il faut mentionner dans les visites de suivis sont les « nouveaux » ATCD apparus depuis la visite précédente.
Différence entre la CIM-10 et les antécédents non explicités	R	Redéfinition de ces deux catégories sur la base et dans le système de formation
Le questionnaire sur le sport n'est pas bien placé dans la base. Ce n'est pas un facteur de risque	F	Déplacer cette catégorie Intégrer un vrai questionnaire officiel sur le sport
Manque notion de fréquence hebdomadaire pour le sport	F	
Tabac : Tabac en cours → obligé de cocher la fin Calcul du PA : pas de prise en compte sevrage	F	Modification informatique
Mauvais remplissage des traitements sur le	R	Précision de la démarche à suivre sur le questionnaire

questionnaire numérique par les patients		
Menu déroulant de sélection des traitements non obligatoire	R	Démarche à suivre précisée dans le système documentaire
Non harmonisation de la saisie des traitements	R	Harmonisation via le code ATC - DCI
Manque informations sur l'examen ORL	R	Réalisation d'une partie dans la base sur l'examen ORL
Termes compliqués du questionnaire papier pour le patient	R	Modification des termes compliqués pour le patient dans le questionnaire papier
Manque trouble de la concentration	F	
Motif de consultation unique	F	Modification sur le système informatique. Possibilité de sélectionner plusieurs motifs de consultation
Echelle de dépression de Beck	F	Ajouter cette échelle sur la base MARS
Manques d'informations sur les Habitudes (heure de réveil, Mettre des plages horaires, etc.)	F	
Permis : Mauvaise connaissances des permis	F	Reprendre et redéfinir cette partie
Unités différentes d'un pays à l'autre	R	Sélection de l'unité par le responsable de chaque centre pour les profils utilisateur de son centre
Mauvais import de donnée sur le « PLM pendant le TTS »	F	Vérifier l'export et réaliser la modification
Scores	NR	Harmonisation avec échelles

			internationales
	Langues différentes dans les pays adhérents	R pour le Portugal F pour les autre langue	Traduction de la base
	Manque notion de fréquence hebdomadaire pour le sport	F	
	Utilité de l'information sur la libido/érection	NR	Information pouvant être utilise
	Données peu saisies de certaines parties	F	Sensibiliser à l'importance de la saisie de l'ensemble des informations Redéfinir les processus
	Réarrangement de certaines partie	F	Etudier cas par cas les parties
Les attentes des utilisateurs	Mise en place d'un mode d'emploi	R	Mise en place d'un système documentaire avec suivi de formation
	Export des visites pour faire des demandes d'examens	F	
	Edition d'un compte rendu de consultation automatiques	F	
	Outils de reporting avec Graphiques : paramètres bio, scores, inclusion etc.	R	Mise en place de différents outils de reporting en fonction des besoins
	Rapport d'utilisation d'un rapport mensuel pour les administrateurs	R	Mise en place d'un rapport activité de la base MARS
	Bonne formation à l'utilisation initiale	F	Mise en place d'un système documentaire avec suivi de formation

	Création d'une newsletter pour les modifications de la base	R	Mise en place de pop-up sur chaque compte utilisateur avec les dernières modifications depuis la dernière connexion
	Aider à l'inclusion dans les protocoles cliniques	R	Affichage d'un pop-up lors de la saisie lorsque le patient présente plusieurs critères d'inclusion Description des différents protocoles d'études dans le système documentaire
	Redéfinition des processus d'utilisation de la base	R	Réalisation d'un système documentaire reprenant les différents processus



Serment de Galien



« Je jure en présence des Maîtres de la Faculté, des Conseillers de l'Ordre des Pharmaciens et de mes condisciples :

D'honorer ceux qui m'ont instruit(e) dans les préceptes de mon art et de leur témoigner ma reconnaissance en restant fidèle à leur enseignement.

D'exercer, dans l'intérêt de la santé publique, ma profession avec conscience et de respecter non seulement la législation en vigueur, mais aussi les règles de l'honneur, de la probité et du désintéressement.

De ne jamais oublier ma responsabilité et mes devoirs envers le malade et sa dignité humaine ; en aucun cas, je ne consentirai à utiliser mes connaissances et mon état pour corrompre les mœurs et favoriser des actes criminels.

Que les hommes m'accordent leur estime si je suis fidèle à mes promesses. Que je sois couvert(e) d'opprobre et méprisé(e) de mes confrères si j'y manque ».

THÈSE SOUTENUE PAR : Virginie SUIZE

TITRE :

MISE EN ŒUVRE DU PROCESSUS QUALITÉ D'UNE BASE DE DONNÉES MÉDICALES.
APPLICATION À LA BASE MARS DU CENTRE HOSPITALIER UNIVERSITAIRE GRENOBLE-ALPES

RESUME :

Le volume de données collectées suit une évolution exponentielle depuis plusieurs années et cette tendance se poursuivra dans les prochaines décennies. La recherche médicale est un des secteurs qui génère un volume important et une grande diversité de données, notamment dû aux changements des pratiques médicales. Cette évolution est accompagnée d'un changement de paradigme dans les méthodes d'analyse des données avec les approches « Big Data ». Plusieurs études comparatives ont montré que l'utilisation des données observationnelles, via les bases de données médicales, peuvent être utilisées en complément aux essais cliniques randomisés.

Cependant, pour assurer la publication de résultats de qualité et éviter l'apparition de biais, il est essentiel d'avoir des données structurées et de qualité.

Dans ce contexte, afin de constituer une base de données médicale pour le suivi prospectif des patients consultants pour un trouble du sommeil ou une pathologie respiratoire, la base de données Multimorbidity apnea respiratory failure sleep database (MARS) a été mise en place dans le service des explorations fonctionnelles cardiorespiratoire du centre hospitalier Grenoble- Alpes.

Pour garantir la qualité des résultats de recherche sur cette base, un processus qualité a été déployé via la mise en place de divers outils tels que :

- Des alertes de saisie,
- Un système documentaire pour assurer la formation des utilisateurs,
- Des outils de reporting pour mettre en valeurs certaines données importantes,
- Un rapport de suivi aux administrateurs permettant de suivre l'activité et la qualité des données via des indicateurs.

MOTS CLÉS : MARS - Qualité - Règlementation - Documentation - Reporting - Indicateurs

FILIÈRE : Pharmacie - Filière Industrie