



BNDMR

Banque Nationale de Données
Maladies Rares



Orientations Stratégiques de la Banque Nationale de Données Maladies Rares (BNDMR) 2020

Janvier 2020

Sommaire

| | |
|---|-----------|
| Sommaire | 2 |
| Glossaire | 3 |
| Introduction | 4 |
| 1. Contexte..... | 4 |
| 2. Objectifs..... | 4 |
| Axe n°1 : BaMaRa | 5 |
| 1. Evolutions de BaMaRa..... | 5 |
| 1.1. <i>Maintenance corrective et évolutive</i> | 5 |
| 1.2. <i>Modules d'informatique décisionnelle et d'administration</i> | 5 |
| 1.3. <i>Mise à jour des référentiels</i> | 5 |
| 1.4. <i>Evolutions techniques</i> | 5 |
| 2. Certification « hébergeur de données de santé » (HDS) de l'AP-HP..... | 6 |
| Axe n°2 : Interopérabilité | 6 |
| 1. Suivi de l'Appel à Projets interopérabilité..... | 6 |
| 2. Connecteur – Cadre d'interopérabilité..... | 7 |
| 3. Interopérabilité avec d'autres bases de données..... | 7 |
| Axe n°3 : Déploiement de BaMaRa | 8 |
| 1. Prérequis : Migration des données de CEMARA..... | 8 |
| 1.1. <i>Correction des données de CEMARA et migration vers BaMaRa</i> | 8 |
| 1.2. <i>Formation des référents filières</i> | 8 |
| 2. Méthodologie de déploiement et calendrier..... | 9 |
| 2.1. <i>Méthodologie de déploiement</i> | 9 |
| 2.2. <i>Calendrier</i> | 9 |
| Axe n°4 : La Banque Nationale de Données Maladies Rares | 9 |
| 1. Mise en œuvre de l'autorisation..... | 9 |
| 2. Développements informatiques..... | 10 |
| 3. Comité Scientifique BNDMR..... | 10 |
| 4. Valorisation des données..... | 11 |
| Axe n°5 : L'appariement BNDMR - SNDS | 11 |
| Axe n°6 : L'Intelligence Artificielle au service du projet BNDMR | 12 |
| 1. Algorithme de structuration des comptes-rendus médicaux..... | 12 |
| 2. Algorithme de dé-identification de textes libres..... | 12 |
| 3. Alerte sur incidences suspectes..... | 13 |
| 4. Détection de doublons potentiels..... | 13 |
| Axe n°7 : Dissémination | 13 |
| Axe n°8 : Travaux européens | 14 |
| 1. RD Code..... | 14 |
| 2. EJP Rare Diseases..... | 14 |
| 3. Registres ERN..... | 14 |

| | |
|---|-----------|
| 4. European University Hospital Alliance (EUHA) | 15 |
| Axe n° 9 : recueil complémentaire Foetopathologie..... | 15 |
| Axe n° 10 : Patients sans diagnostic | 15 |
| 1. Pilote sur l'errance et l'impasse diagnostique avec la filière FILNEMUS | 16 |
| 2. Déploiement national inter-filières..... | 16 |
| Axe n° 11 : Autres actions du PNMR 3..... | 16 |
| 1. Action 3.2 : nouveaux entrepôts de données interopérables..... | 16 |
| 2. Action 4.3 : générer des connaissances en vie réelle | 17 |
| 3. Action 6.1 : groupe de coordination de l'innovation sur les maladies rares | 17 |
| 4. Action 7.5 : un DMP enrichi pour tous les malades atteints de maladies rares | 17 |
| Axe n° 12 : Comité de Pilotage BNDMR..... | 17 |
| Annexe – budget et ressources humaines | 18 |
| 1. Budget..... | 18 |
| 2. Ressources humaines | 18 |

Glossaire

| | |
|-------------|--|
| AP-HP | Assistance Publique – Hôpitaux de Paris |
| CHU | Centre Hospitalier Universitaire |
| CNIL | Commission Nationale de l'Informatique et des Libertés |
| CO BNDMR | Cellule Opérationnelle BNDMR |
| CoSci | Comité Scientifique BNDMR |
| CoPil | Comité de Pilotage BNDMR |
| DGOS | Direction Générale de l'Offre de Soins |
| DGS | Direction Générale de la Santé |
| FSMR | Filière de Santé Maladies Rares |
| MR | Maladies rares |
| PNMR | Plan National Maladies Rares |
| RGPD | Règlement Général sur la Protection des Données |
| SDM | Set de Données Minimum |
| SI | Système d'information |
| SNDS | Système National de Données de Santé |

Introduction

1. Contexte

La création d'une banque nationale de données maladies rares est l'un des objectifs prioritaires des plans nationaux maladies rares. Lancé il y a 6 ans, le projet a avancé sur des sujets indispensables à la mise en place de cette banque de données pérenne et interopérable, au niveau national et international. En effet, en 2013 une équipe de développeurs a été recrutée, et le Set de Données Minimum maladies rares (SDM) a été créé, constituant les premières étapes en vue de la BNDMR. En 2015, la DGOS a confié la maîtrise d'œuvre du projet à l'AP-HP dont les modalités ont été fixées dans une convention cadre. En 2018, la publication du PNMR 3 réaffirmait l'importance du projet et prolongeait son financement et ses missions.

Le présent document, reflétant les orientations stratégiques de la BNDMR est destiné à l'ensemble des parties prenantes afin de leur permettre de suivre les travaux de mise en place de la BNDMR. Il indique les priorités de l'année 2020 et les délais de réalisation envisagés. Ceux-ci restent des orientations, et peuvent être amenés à évoluer en fonction des circonstances au cours de l'année.

Ce document est à lire conjointement au rapport annuel public 2019 et aux autres documents d'information disponibles sur le site bndmr.fr qui décrivent plus en détails le contexte et les briques existantes du projet.

2. Objectifs

Les objectifs suivants constitueront, en synthèse, le cœur de l'action de la Cellule Opérationnelle BNDMR en 2020 :

- **BaMaRa (« mode autonome »)** : après avoir dépassé les 70% de sites maladies rares déployés, les déploiements continueront selon le rythme affiché sur notre site internet bndmr.fr pour atteindre plus de 85% en fin d'année. Ce rythme est défini conjointement avec les établissements.
- **Interopérabilité (« mode connecté »)** : les établissements retenus dans l'appel à projet du ministère disposeront du mode « connecté » (saisie des données directement dans les DPI) d'ici la fin de l'année 2020.
- **BNDMR** : la campagne de signature des conventions inter-hospitalières se poursuivra et les données seront progressivement chargées dans l'entrepôt. Les développements informatiques permettront un fonctionnement optimal et facilitant sur le plan opérationnel pour toutes les finalités.
- **Appariement BNDMR – SNDS** : les démarches réglementaires entamées en 2019 seront poursuivies en vue d'une première étude sur les données appariées dès 2020. Un dossier en vue d'un entrepôt pérenne appariant BNDMR et SNDS sera soumis.
- **Patients sans diagnostic** : le projet pilote sur l'errance et l'impasse diagnostiques, mené avec la filière FILNEMUS, sera étendu à tous les centres de référence de la

filrière. L'extension nationale à toutes les filières sera également réalisée. Le service d'un recueil complémentaire sur la fœtopathologie sera ouvert.

Axe n°1 : BaMaRa

1. Evolutions de BaMaRa

1.1. Maintenance corrective et évolutive

Suite aux retours des utilisateurs recueillis au fil de l'eau (et à mesure que ces utilisateurs se font de plus en plus nombreux), des évolutions sont réalisées régulièrement afin d'assurer la maintenance de l'application dans le temps et l'amélioration de l'expérience utilisateur. Les résultats de l'enquête de satisfaction 2019 seront pris en compte pour prioriser les améliorations.

1.2. Modules d'informatique décisionnelle et d'administration

BaMaRa dispose actuellement de modules permettant aux utilisateurs d'exploiter et d'analyser les données de leurs patients de manière autonome. Ces modules seront complétés par des possibilités d'extraction de rapports sur la qualité des données et de rapports statistiques sur différents périmètres (le site, l'hôpital) et PIRAMIG.

1.3. Mise à jour des référentiels

Tous les référentiels de BaMaRa seront mis à jour une fois en 2020, de manière conforme aux préconisations du groupe de travail Européen RD-Action. Un document relatif à l'industrialisation de l'utilisation des référentiels sur le long terme (incluant la méthodologie de mise à jour des référentiels chaque année) a été produit et partagé notamment avec les établissements disposant de la fiche maladies rares.

Une première mise à jour a eu lieu fin 2019 concernant le référentiel « maladies rares », permettant ainsi aux utilisateurs d'utiliser dans BaMaRa la toute dernière classification ORPHANET, notamment pour coder les diagnostics.

La mise à jour des autres référentiels aura lieu en Janvier 2020.

Les versions des référentiels utilisées dans l'application BaMaRa seront alors rendues publiques. Ce qui, couplé au document d'usage des référentiels ci-dessus, permettra aux éditeurs de DPI de proposer une mise à jour des référentiels utilisées dans les fiches maladies rares du mode connecté de leur logiciel.

1.4. Evolutions techniques

Les briques techniques de BaMaRa ont d'ores et déjà été progressivement revues depuis le lancement de l'application. En effet, si à l'origine la base de données était en MongoDB et l'application en Ruby on Rails, une approche par micro-services a permis d'ajouter certaines fonctionnalités en Java. Cette stratégie n'a pas encore porté sur les interfaces utilisateurs mais seulement sur des composants dits de « back end » qui constituent le moteur de l'application. Elle vise à garantir que les technologies utilisées sont les plus adéquates sur les plans de la performance, de la sécurité, ainsi que sur le plan opérationnel. En 2020, l'effort sera poursuivi

et pourra même s'étendre à l'affichage, certaines interfaces pouvant être basculée sur de nouveaux langages.

2. Certification « hébergeur de données de santé » (HDS) de l'AP-HP

Les exigences en matière de sécurité sont importantes dans le domaine des systèmes d'information de santé, en particulier pour l'application BaMaRa qui permet de recueillir des données de patients non pris en charge à l'AP-HP, et dont les données y sont stockées. C'est ainsi que BaMaRa s'inscrit dans la démarche de certification HDS de l'AP-HP.

Cette certification se matérialisera par :

- La finalisation de tout le corpus documentaire requis (analyse d'impact, analyse de vulnérabilités, test d'intrusion, homologation de sécurité, procédures diverses sur tous les aspects sécurité des normes ISO, notamment) : en effet, si les différents points techniques identifiés en 2019 dans le cadre de la préparation de cette certification sont d'ores et déjà respectés de manière satisfaisante, la documentation de tous les sujets sera mise à jour et précisée afin d'atteindre les attendus
- Un audit sur pièces et sur place réalisé par une société agréée

La cellule opérationnelle se montrera particulièrement diligente sur ces différents temps (préparation de l'audit, durée de l'audit, et suivi post-audit).

Axe n°2 : Interopérabilité

1. Suivi de l'Appel à Projets interopérabilité

La DGOS a lancé un appel à projets (AAP) incitatif à destination des CH(U) pour l'implémentation de la fiche maladies rares dans leur Dossiers Patient Informatisés (DPI) ainsi que le transfert de ces données dans BaMaRa dans le respect du cadre d'interopérabilité national des systèmes d'information en santé (CI-SIS) publié par l'ASIP-Santé.

18 établissements ont été retenus par le ministère, et ont engagé les travaux d'interopérabilité requis avec des cibles à atteindre en 2020.

La CO BNDMR se tient à la disposition des établissements concernés afin d'accompagner ces projets, conjointement avec les éditeurs de logiciels.

Au cours de l'année 2019, la BNDMR a pu collaborer avec les éditeurs et constater un bon avancement dans la mise en œuvre de fiches maladies rares au sein des DPI (soit la partie 1 « intégration » de l'appel à projets), allant parfois jusqu'à la mise à disposition de la fiche maladies rares aux utilisateurs finaux, bien que l'échange des données ainsi recueillies ne soit pas encore opérationnel. Cet effort se poursuivra tout au long de l'année à venir.

Dans le cadre de la partie 2 « transmission » relative à l'échange des données recueillies, une spécification complémentaire au volet est en cours de rédaction et indiquera les modalités précises avec lesquelles les documents devront être envoyés et seront contrôlés par la

BNDMR, ainsi que les retours que nous mettrons à disposition des émetteurs afin qu'ils puissent constater la bonne intégration des documents ainsi envoyés.

Aussi, des premiers échanges de fichiers de test ont eu lieu et des retours manuels ont pu être réalisés. Dans le courant de l'année 2020, la BNDMR automatisera le processus permettant la réception, le contrôle et l'intégration des CDA reçus par MSSanté et émis par les DPI.

L'avancement de ces deux composantes, à savoir l'intégration de la fiche maladie rare et la collaboration éditeur/BNDMR afin de permettre l'échange des documents au format CDA via MSSanté, est donc en progrès avec pour objectif de mise en œuvre la nouvelle date cible de septembre 2020, conformément à la mise à jour de l'appel à projets.

2. Connecteur – Cadre d'interopérabilité

Le volet maladies rares du CI-SIS définit les règles et les exigences nationales auxquelles la fiche « maladies rares » doit répondre. Il fixe ainsi le contenu (champs de la fiche maladies rares, terminologies associées permettant un langage commun à l'échelle nationale...) et les modalités techniques pour le partage et l'échange des données (transport des flux, sécurité et confidentialité). La mise à jour de ce « volet SDM-MR » a ainsi été publiée par l'ASIP-Santé en février 2019.

Les établissements retenus dans l'AAP produiront des fichiers conformes à ce volet, contenant pour un patient donné l'intégralité du set de données minimum, et les adresseront à la CO BNDMR. Cette dernière développe le module qui lira les fichiers reçus des établissements et chargera les données dans leur instance BaMaRa.

Les échanges avec l'ASIP santé se poursuivent afin de faire évoluer le volet SDM-MR du cadre d'interopérabilité des systèmes d'information de santé suite aux remarques émises par les éditeurs dans le cadre du projet. Cela afin de permettre un transfert correct des informations SMD-MR recueillies dans les DPI.

3. Interopérabilité avec d'autres bases de données

Le modèle de reprise de données d'autres registres n'a pas encore été validé, la preuve de concept restant à apporter. En effet, le projet pilote avec la base de données eDBAI (portant sur les Dermatites Bulleuses Auto Immunes) a permis de mettre au jour différents degrés de complexité : au-delà des aspects légaux, de protection des données, techniques, et d'alignement des sets de données, l'interaction entre le mode autonome (la saisie directe dans BaMaRa) et le mode connecté (la reprise de données) doit être précisée. En effet, des données contradictoires peuvent se trouver dans les deux bases, et il convient donc de définir des règles de gestion de ces contradictions.

En parallèle, les travaux menés avec différents autres acteurs ont révélé la nécessité d'un investissement important de ces acteurs, ce qui n'est pas toujours atteignable.

En 2020, la cellule opérationnelle continuera donc de travailler avec eDBAI en vue de finaliser ce projet pilote et de valider la preuve de concept, ainsi qu'avec d'autres projets sous réserve de leur disponibilité (notamment au vu des ressources significatives nécessaires).

Axe n° 3 : Déploiement de BaMaRa

1. Prérequis : Migration des données de CEMARA

1.1. Correction des données de CEMARA et migration vers BaMaRa

Un contrôle qualité des données de CEMARA est effectué régulièrement par la cellule opérationnelle BNDMR. La data manager de l'équipe identifie les dossiers qui nécessitent d'être corrigés pour être intégrés au sein de l'application BaMaRa. Sans action, ces dossiers ne peuvent pas être migrés dans BaMaRa et les données risquent d'être perdues.

Depuis 2018, une procédure impliquant les chargés de mission des filières en charge des bases de données a été engagée. La CO BNDMR joue un rôle de coordination. Elle est responsable de la mise à disposition des fichiers de corrections aux chargés de mission des filières qui s'occupent de faire le lien avec les centres maladies rares et les utilisateurs pour qu'ils fassent les corrections correspondantes (champs obligatoires manquants, contrôle de cohérence ...). Les filières ont ainsi la responsabilité d'encourager et d'accompagner ce processus. Cela a permis de réduire le nombre d'interlocuteurs pour la cellule opérationnelle et d'assurer un traitement des corrections plus efficace.

Par ailleurs, l'équipe de la BNDMR assure une mission d'identitovigilance. Ainsi les équipes de soins et utilisateurs de CEMARA sont contactés individuellement par notre data manager (e-mail et téléphone) et notifiés des dossiers potentiellement en doublons au sein d'un hôpital ou d'un service, mais présentant des différences (noms égaux mais sexes différents, noms proches...).

En 2020, dans le cadre du déploiement généralisé de BaMaRa, les demandes aux utilisateurs vont s'intensifier afin de permettre la migration des données et assurer la correction des dossiers avant chaque déploiement d'hôpital. Pour chaque établissement, une date butoir de correction des données de CEMARA sera fournie. Les délais proposés prendront en compte la charge de travail correspondante, et viendront préciser le calendrier de déploiement. Au-delà de cette date, les dossiers non corrigés ne seront pas migrés automatiquement dans BaMaRa et devront donc être ressaisis manuellement par les équipes du centre.

1.2. Formation des référents filières

L'équipe opérationnelle de la BNDMR continuera à assurer des formations directes des référents bases de données des filières à BaMaRa (et par extension au principe de recueil dans les DPI des fiches maladies rares) afin de permettre une meilleure appropriation de l'outil. Une session de formation sera organisée en février 2020 pour les nouveaux chargés de missions ou chefs de projets. Par ailleurs, la CO BNDMR a été sollicité pour faire une session de rappel pour les personnes déjà formées. Afin de répondre à cette demande, pour ne pas mobiliser les équipes des filières en présentiel, un premier échange par téléconférence est proposé fin janvier sous le format de questions/réponses. Si cette formule s'avère efficace, elle pourra être reproduite régulièrement.

Les référents bases de données seront en charge de la formation des utilisateurs mais aussi des autres chargé(e)s de mission de la filière le cas échéant.

2. Méthodologie de déploiement et calendrier

2.1. Méthodologie de déploiement

Le déploiement est effectué par établissement. L'ouverture des accès à BaMaRa se fait ainsi de manière simultanée sur tous les sites de toutes les filières maladies rares au sein de l'établissement.

2.2. Calendrier

Un calendrier de déploiement est élaboré en prenant en compte le retour et la signature des conventions, la charge de travail nécessaire pour la correction et la migration des données CEMARA.

Ce calendrier est construit en concertation avec les établissements partenaires après leur signature de la convention BaMaRa. Il est disponible en ligne sur le site internet bndmr.fr.

L'objectif de l'année 2020 est que les 18 établissements retenus dans l'appel à projet sur le mode connecté puissent l'avoir mis en place et que BaMaRa (autonome ou connecté) soit ainsi disponible dans tous les CHU.

Axe n° 4 : La Banque Nationale de Données Maladies Rares

1. Mise en œuvre de l'autorisation

La demande d'autorisation de constitution de l'entrepôt national maladies rares à caractère d'intérêt public a été validée par la CNIL en septembre 2019. Le détail de cette autorisation est consultable en ligne¹ et c'est dans le respect de ce cadre que les traitements relatifs à la BNDMR seront menés en 2020.

Suite à la réception de l'autorisation de la CNIL, et afin que les données de BaMaRa soient intégrées dans la BNDMR, il convient de cadrer les responsabilités, droits et modalités de traitement des données pour chaque partie impliquée. La convention cadrant ces modalités a été transmise en 2019 à 43 établissements (tous les établissements ayant accès à BaMaRa fin 2019 ainsi que tous les établissements de l'AAP mode connecté) et 15 établissements l'avaient signée fin 2019. L'objectif est que ces 43 établissements aient signé la convention BNDMR avant l'été 2020. Par ailleurs, les autres établissements seront contactés au premier trimestre 2020 (convention BaMaRa et BNDMR seront envoyées conjointement).

En matière d'information et de modalités d'exercice des droits, les dispositions prises en 2019² seront maintenues. S'agissant des mesures de sécurité, elles seront là aussi maintenues, ce qui sera garanti par l'audit prévu (voir la section relative à la certification « hébergeur de données de santé »).

¹ <https://www.legifrance.gouv.fr/affichCnil.do?id=CNILTEXT000039292715>

² Voir le Bilan 2019, disponible sur <http://www.bndmr.fr/publications/presentations-communications/>

2. Développements informatiques

En 2020, la BNDMR sera mise en production sur une plateforme bénéficiant des technologies innovantes Kubernetes et Docker, qui facilitent les évolutions et ont une forte capacité de mise à l'échelle (« scalability »), permettant de faire face à des périodes de calculs nombreux et complexes.

Le cadrage de l'intégration des données publiques relatives à l'environnement, la pollution de l'air et des sols, etc. viendra compléter les données de santé en vue d'études de causalité entre ces facteurs externes et les pathologies et signes répertoriés.

Enfin, le processus d'identito-vigilance nationale sera mis en production en 2020. Il s'agit d'une détection automatique, sans intervention humaine, de doublons potentiels (par exemple deux patients ayant des traits d'identité identiques à un caractère près), qui entraînera une alerte auprès des établissements concernés.

3. Comité Scientifique BNDMR

Les sollicitations du Comité Scientifique BNDMR sont en forte augmentation. Ainsi, une refonte significative du fonctionnement du comité a été organisée fin 2019, et validée par le Comité de Pilotage BNDMR. Cette refonte repose sur les principes suivants :

- Une réunion tous les premiers vendredis de chaque mois, garantissant la plus grande réactivité, annulable s'il n'y a pas de demande reçue
- Le fait de confier le remplissage d'une grille d'évaluation à un rapporteur, choisi à tour de rôle parmi les membres
- Réunions par téléphone, permettant à un maximum de membres de participer (le projet étant national, le Comité Scientifique est composé de personnalités basées sur tout le territoire)

En outre, afin de rendre possible sur le plan opérationnel cette nouvelle organisation, un rétro-planning strict a été mis sur pieds, par lequel la cellule opérationnelle s'engage à :

- Revenir sous une semaine vers tout demandeur pour préciser le processus et présenter le(s) document(s) à remplir (www.bndmr.fr/exploiter/donnees-bndmr/procedure-de-demande/)
- Arrêter la liste des demandes à deux semaines avant la date du prochain comité, le formulaire complet ayant dû être reçu (faute de quoi, la demande est présentée au comité suivant), afin de garantir un temps d'analyse suffisant par les membres du comité

A noter que certaines demandes de traitement simples et récurrentes ne nécessiteront pas l'avis du Comité Scientifique :

- Les indicateurs à destination des services de l'Etat ainsi que les rapports à destination des FSMR
- Les études de faisabilité (i.e. études visant à compter les patients répondant à certains critères fixés par l'investigateur), qu'elles soient demandées par des investigateurs membres du réseau d'expertise sur les maladies rares en lien ou non avec un

destinataire final issu du secteur privé, ou demandées par des investigateurs issus du secteur public.

- Le nombre de patients (agrégé au niveau national, le critère étant le code Orphanet et/ou un ensemble de signes cliniques) quand l'investigateur est une association de patients membres d'une filière.

Le Comité Scientifique sera toutefois informé de toutes les études conduites, en tant qu'il reverra systématiquement la liste exhaustive des études réalisées (cf. tableau disponible sur notre site internet).

4. Valorisation des données

La valorisation des données de la BNDMR est une priorité pour la CO BNDMR.

En 2020, la CO BNDMR va poursuivre sa collaboration avec plusieurs filières dans le cadre d'études longitudinales.

Le PNMR3 conduira à plusieurs analyses de données cette année, comme l'année passée. En particulier, la BNDMR continuera d'être une source de données pour le calcul d'indicateurs de performance du plan.

L'axe sur l'errance et l'impasse diagnostiques, en particulier, conduira à plusieurs analyses qui sont déjà en cours, notamment s'agissant de la valorisation de toutes les données collectées dans le pilote Filnemus.

Les associations de patients jouent un rôle déterminant en sollicitant les médecins experts afin que ceux-ci portent d'ambitieux projets d'analyse. Elles pourront également continuer d'interroger la BNDMR afin de connaître le nombre de patients atteints des pathologies qui les concernent.

Les universitaires, notamment dans le secteur de la data science, ont également à plusieurs reprises sollicité la BNDMR en 2019 et cela a vocation à continuer en 2020. Le prototype d'un outil d'alerte sur des niveaux anormaux d'incidence de certaines pathologies a ainsi développé et sa généralisation sera envisagée.

Enfin, plusieurs entreprises du secteur pharmaceutiques ont récemment transmis à la CO BNDMR des projets de collaboration ou des demandes d'analyse que la BNDMR rendra possible.

Axe n° 5 : L'appariement BNDMR - SNDS

Comme démontré par la littérature scientifique, l'exploitation du SNDS à des fins de santé publique sur les maladies rares est rendue difficile, voire parfois impossible, du fait de l'absence de codage spécifique des diagnostics maladies rares au sein du SNDS.

Ainsi, l'appariement entre la BNDMR et le SNDS est un objectif qui a présidé à celui de la création de la BNDMR elle-même : pouvoir conduire des études de santé publique de grande

envergure sur la cohorte entière des patients atteints de maladies rares (vus dans le réseau expert labellisé).

Dans ce cadre, en 2020, la BNDMR travaillera conjointement avec la CNAV et la CNAM, afin de mettre sur pied une solution technique d'appariement. Un dossier sera par ailleurs soumis à l'INDS, au CEREES et la CNIL afin de solliciter l'autorisation pour un tel appariement. Enfin, et sous réserve d'une telle autorisation, de premiers indicateurs de pilotage de l'offre de soins découlant de cet appariement seront produits au dernier trimestre 2020 sur la base des données présentes dans la BNDMR à cette date.

Axe n° 6 : L'Intelligence Artificielle au service du projet BNDMR

1. Algorithme de structuration des comptes rendus médicaux

A ce jour, les données relatives à la BNDMR sont saisies de manière dite « structurée », c'est-à-dire dans des formulaires dédiés. L'intelligence artificielle (IA), c'est-à-dire un ensemble de techniques statistiques, pourrait potentiellement permettre de pré-remplir voire de remplir directement ces formulaires à partir des comptes rendus disponibles dans les DPI. Dès 2019, des discussions ont eu lieu avec différents experts en Traitement Automatique du Langage, et avec différents entrepôts de données de santé contenant de tels comptes rendus.

Pour mettre sur pieds un tel algorithme, il convient dans un premier temps d'annoter les comptes rendus (afin de stocker, par exemple, l'information : « ici se trouve l'information selon laquelle l'âge aux premiers signes est de 2 ans et 3 mois »). Cette étape d'annotation, cruciale pour constituer ce que l'on appelle le « jeu de données d'entraînement » sur lequel l'algorithme pourra être construit, permettra aussi d'évaluer la faisabilité du projet. En effet, si l'IA permet de structurer (c'est-à-dire de récupérer) une donnée particulière (par exemple un item du SDM) parmi un ensemble de données non-structurées (par exemple un texte), elle ne peut en tout état de cause pas récupérer une donnée inexistante ni une donnée tout juste suggérée.

Ainsi, en 2020, la faisabilité d'utiliser l'IA pour pré-remplir, voire directement remplir le SDM, sera étudiée. Pour certains items, comme le lieu de naissance, il y a un risque que les données soient insuffisamment présentes. Pour d'autres, comme l'âge au diagnostic, il y a un risque qu'elles soient trop difficiles à récupérer dans les comptes rendus médicaux de par leur nature subjective et parfois ambiguë. Pour d'autres encore, comme le diagnostic selon les termes Orphanet ou les signes cliniques, il est très probable qu'un algorithme fonctionne bien (il existe de nombreux applications comparables dans d'autres domaines).

En fonction de cette faisabilité, le projet (c'est-à-dire le développement de l'algorithme, et sa diffusion aux établissements intéressés sous la forme d'une API) pourra être engagé.

2. Algorithme de dé-identification de textes libres

L'autorisation BNDMR ne prévoit pas que le champ en texte libre de BaMaRa remonte dans l'entrepôt. Un travail de dé-identification doit en effet être conduit au préalable. Si l'on estime parfois que de l'ordre de 80% des traits d'identité présents dans des comptes rendus médicaux correspondent à ceux du patient (et peuvent donc être retirés de manière déterministe, étant connus), il reste donc de l'ordre de 20% de traits d'identité correspondant à d'autres individus (apparentés, médecins, etc.) qu'il convient de retirer (au moins en partie). Un travail sera conduit en 2020 afin de réussir cette dé-identification.

3. Alerte sur incidences suspectes

Ainsi qu'évoqué dans la section « Valorisation des données », un travail a d'ores et déjà été conduit afin de tester des niveaux d'incidence suspects (par ex, nombre de cas suspects sur une période et/ou une région donnée). Ce travail, reposant sur l'application à un mélange de Gaussiennes d'une méthode de Monte Carlo par Chaîne de Markov (MCMC), sera poursuivi et pourra être mis en production si sa fiabilité est suffisante, s'accompagnant en outre d'un processus d'analyse en cas de remontée d'alertes.

4. Détection de doublons potentiels

La détection de doublons potentiels, évoquée dans la section « développements informatiques » ci-dessus, repose sur différentes stratégies de comparaison des traits d'identité au moyen de diverses « distances » (définies parmi les outils d'IA appliquée au langage). L'algorithme produisant les résultats (i.e. la probabilité pour deux groupes de traits d'identité d'être en fait un doublon) restera à finaliser en 2020. Un document détaillant l'approche retenue sera in fine publié.

Axe n° 7 : Dissémination

La CO BNDMR participera à une majorité des journées nationales de FSMR, ainsi qu'à certaines journées nationales de CRMR. La communication sera en outre continuellement assurée par le site internet bndmr.fr, qui promeut notamment les tutoriels et guides utilisateurs. En 2020, nous espérons pouvoir augmenter de 10% le trafic sur le site de la BNDMR.

La BNDMR organisera le 3 juillet 2020 (initialement prévu le 1^{er} avril) son second séminaire Utilisateurs, et attend plus de 150 personnes. Il sera l'occasion de présenter les avancées du projet et de montrer les possibilités d'analyses de données.

Le comité de pilotage de la BNDMR a souligné l'importance de la présence de la BNDMR lors de congrès sur les maladies rares, ou les congrès de spécialités médicales où elle pourrait soumettre, en lien avec les filières concernées, un poster sur les résultats d'une étude dans le domaine. La BNDMR sera également présente lors de congrès d'informatique médicale ou d'épidémiologie.

Au niveau européen, la CO BNDMR participera au congrès européen sur les maladies rares (ECRD) à Stockholm en mai 2020, où elle soumettra deux abstracts pour deux posters.

Par ailleurs, plusieurs projets de publications scientifiques sont en cours (cf. partie précédente sur la valorisation des données), ce qui permettra de faire connaître à la communauté scientifique et médicale le projet et son potentiel en matière de recherche.

Axe n° 8 : Travaux européens

La France, premier pays à s'être doté d'un Plan National Maladies Rares, a historiquement servi de modèle pour la définition des politiques publiques européennes en matière de maladies rares. La structuration française en filières maladies rares a fortement inspiré les ERNs (European Reference Network), et le set de données minimum maladies rares proposé récemment au niveau européen reprend en grande partie les travaux français. Cela constitue quelques exemples d'articulation entre la France et l'Europe.

Aujourd'hui, la France semble être un des pays les plus avancés dans la création d'une base nationale dédiée à l'étude de l'épidémiologie des maladies rares. La BNDMR doit donc être attentive à ce qui se met en place au niveau Européen pour ne pas travailler en silos.

1. RD Code

Dans la continuité de RD-ACTION, la CO BNDMR participe depuis début 2019 en tant qu'expert à l'action commune (joint action) de la commission européenne, RD CODE. Ce projet a pour objectif d'implémenter la nomenclature ORPHA dans 4 états membres (Espagne (6 régions ciblées), Malte, République Tchèque, Roumanie) suivant les recommandations émises dans le précédent projet RD-ACTION. La BNDMR assure un rôle de leader sur le programme de travail relatif au codage des patients sans diagnostic, en lien avec son expertise au niveau national.

Elle remettra en septembre 2020 une recommandation européenne pour le codage de ces patients dans les systèmes d'information.

2. EJP Rare Diseases

Les travaux de l'EJP RD, qui vont en s'accéléralant, consisteront notamment à recenser et rendre interopérable les bases de données existantes en Europe. Plusieurs groupes de travail ont ainsi été constitués. La CO BNDMR a officiellement été intégrée en tant que collaboratrice du projet en décembre 2019, dans le Pilier 2 et contribuera à ces groupes de travail en apportant son expertise sur les métadatas, les standards, le partage des données et leur analyse. La BNDMR sera un cas d'usage pour les travaux d'interopérabilité.

3. Registres ERN

Plusieurs ERN ont sollicité la BNDMR dans le cadre du *Call for projects* de la Commission Européenne en vue de bâtir des *ERN registries*³ pour la rédaction d'une lettre d'intention par

³ Call HP-PJ-2019

laquelle la cellule opérationnelle s'engage à travailler à l'interopérabilité avec ces futurs registres.

Si de nombreux points d'interrogation restent à lever, la cellule opérationnelle a systématiquement répondu favorablement : en 2020, la connexion entre la BNDMR et ces registres sera l'occasion de travaux préparatoires et prospectifs, visant à définir les modalités à envisager et à cadrer le projet dans la durée.

4. European University Hospital Alliance (EUHA)

La cellule opérationnelle contribuera au nouveau groupe de travail orienté sur les maladies rares de l'alliance des hôpitaux universitaires européens. La promotion du Set de Données Minimum Maladies Rares (SDM-MR) français – qui est nettement plus large que le Common Data Elements (CDE) européen, celle du volet SDM-MR du cadre d'interopérabilité (qui est conforme aux normes internationales), ou encore l'approche par collecte de ces données de manière structurée dans les DPI, constitueront les principaux axes de plaidoyer de la cellule opérationnelle.

Axe n° 9 : recueil complémentaire Fœtopathologie

Le PNMR 3 souhaite davantage « structurer les activités de fœtopathologie », et lutter contre l'impasse diagnostique. C'est à ces deux problématiques que la Société Française de Fœtopathologie en lien avec la CO BNDMR, souhaite s'adresser en développant un recueil complémentaire spécifique à la fœtopathologie, au sein de l'application BaMaRa. Comme pour les items du SDM, ce recueil complémentaire intégrera la BNDMR, afin de permettre des études de recherche à l'échelle nationale.

Le recueil complémentaire, en phase de test fin 2019, sera présenté au premier semestre 2020 devant la communauté des fœtopathologistes.

En 2020, ce recueil complémentaire sera mis en production, et les fœtopathologistes seront formés à la saisie.

A noter que certains établissements ayant expertise en matière de fœtopathologie ne sont pas dans le réseau maladies rares. Les conventions BaMaRa et BNDMR, de par leur participation au réseau de fœtopathologie, leurs seront adressées au premier semestre 2020.

Un premier rapport annuel devrait pouvoir être établi à la fin de l'année 2020 pour présenter les chiffres d'activité et de premières analyses sur le réseau de fœtopathologie.

Axe n° 10 : Patients sans diagnostic

1. Pilote sur l'errance et l'impasse diagnostique avec la filière FILNEMUS

Les résultats positifs du pré-pilote avec la filière FILNEMUS, qui a consisté à la mise en place et l'utilisation de recueils complémentaires dans BaMaRa, seront valorisés dans une publication scientifique. Ces résultats seront, conjointement avec le retour d'expérience des médecins, exploités pour réviser les recueils et les rendre encore plus pertinents. Pour cela des sessions de travaux sont prévues dès le début de l'année avec les médecins des centres ayant participé.

Ce travail qui a débuté avec trois sites coordonnateurs de la filière des maladies neuromusculaires sera étendu à d'autres centres.

2. Déploiement national inter-filières

Au cours de l'année 2019, en parallèle du projet pré-pilote initié avec la filière FILNEMUS, la cellule opérationnelle avait travaillé sur un axe inter-filières visant à préparer le déploiement à l'échelle nationale et sur toutes les FSMR. Un rapport a été rendu en ce sens à la DGOS.

Ce rapport préconise le déploiement du projet de manière adaptée à chaque FSMR, en retenant trois scénarios :

- Pour certaines, des règles de codage plus précises du SDM et notamment de son volet relatif au diagnostic seraient suffisantes
- Pour d'autres, des recueils complémentaires pour certaines pathologies ou certains cas particuliers d'errance et d'impasse diagnostics seraient nécessaires
- Pour d'autres encore, une approche très similaire à celle suivie par FILNEMUS (à savoir plusieurs recueils complémentaires couvrant, in fine, la totalité des patients sans diagnostic) est privilégiée

La cellule opérationnelle contribuera en 2020 sur ce déploiement via le maintien d'un poste dédié à l'accompagnement du projet, et le développement des outils requis par chaque scénario.

L'observatoire du diagnostic, matérialisé en 2019 par la réunion au 1^{er} octobre de toutes les filières, se réunira aussi en 2020.

Axe n° 11 : Autres actions du PNMR 3

1. Action 3.2 : nouveaux entrepôts de données interopérables

L'action 3.2 prévoit le lancement d'un appel projets (AAP) visant à financer 10 nouveaux entrepôts de données interopérables, ayant notamment pour vocation d'encourager l'innovation. Lorsque ces AAP seront publiés, la cellule opérationnelle proposera un

partenariat aux candidats qui le souhaitent en vue de garantir l'interopérabilité entre BaMaRa/BNDMR et les futurs entrepôts.

2. Action 4.3 : générer des connaissances en vie réelle

Cette action est intimement liée au projet de créer à terme un entrepôt pérenne BNDMR-SNDS. En effet, un tel appariement permettra de connaître la consommation de tous types de soins, notamment des médicaments, par pathologie (et donc y compris les consommations de médicaments hors-AMM).

3. Action 6.1 : groupe de coordination de l'innovation sur les maladies rares

La cellule opérationnelle a d'ores et déjà été contactée en 2019 pour rejoindre ce groupe de coordination, et continuera d'y contribuer en 2020.

4. Action 7.5 : un DMP enrichi pour tous les malades atteints de maladies rares

Plusieurs réunions ont déjà permis en 2019 de dessiner les contours de ce DMP enrichi pour les maladies rares, en 2020 la cellule opérationnelle continuera de contribuer à ce projet, en proposant notamment que tout ou partie du SDM collecté dans les établissements soit transmis au DMP pour enregistrement comme fiche de synthèse.

Axe n° 12 : Comité de Pilotage BNDMR

Comme chaque année, un Comité Scientifique et un Comité de Pilotage seront réunis une fois par trimestre, et fixeront les grandes orientations du projet.

Le Comité de Pilotage BNDMR (CoPil BNDMR) validera les choix opérationnels effectués par la cellule opérationnelle afin de produire les résultats attendus selon les différents axes de travail prévus, et conservera un pouvoir décisionnaire par rapport aux avis rendus par le Comité Scientifique BNDMR (CoSci BNDMR), notamment en matière de contractualisation et de valorisation monétaire des traitements de données.

Annexe – budget et ressources humaines

1. Budget

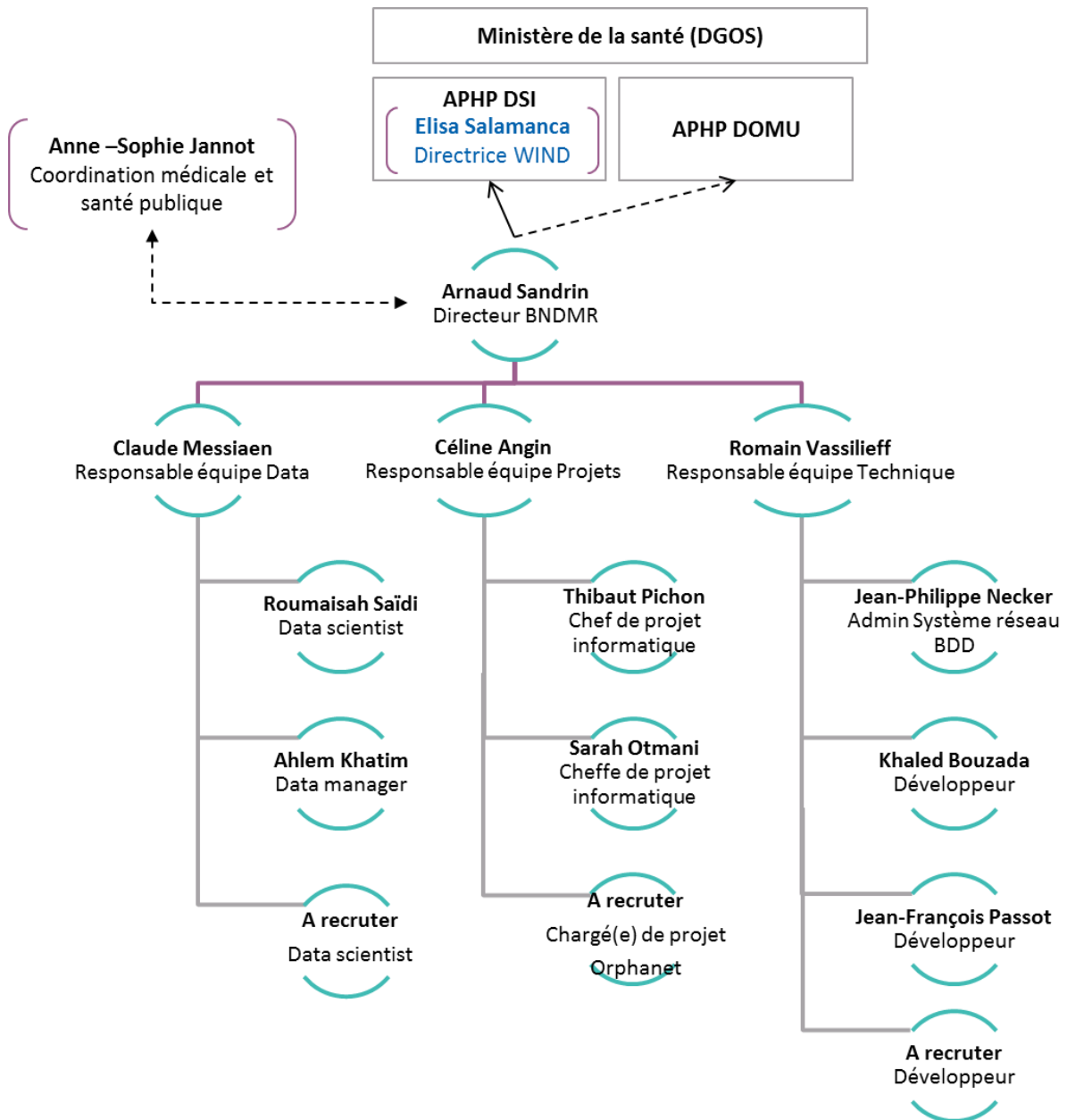
Le budget à fin 2019 est décrit dans le rapport annuel.

La projection ci-dessous pour l'année 2020, tenant compte des prévisions de personnel (augmentées de trois recrutements prévus) et de recours à des prestations externes (pour lesquelles une diminution est prévue), est la suivante.

| | 2019 | 2020 |
|---|--------------------|-----------------------|
| ETP moyen sur l'année | 9,9 | 13,6 |
| TOTAL TITRE 1 - PERSONNEL | 575 000 € | 806 112,39 € |
| TOTAL TITRES 2 et 3 - CHARGES D'EXPLOITATION | 210 000 € | 179 366,48 € |
| TOTAL TITRE 4 - AMTs, PROVs, FIN, EXC | 0,00 € | 0,00 € |
| TOTAL CHARGES avant frais de structure | 785 000 € | 985 478,87 € |
| Frais de structure (% des charges de personnel) | 7% | 7% |
| Frais de structure | 40 250 € | 56 428 € |
| Total Charges | 825 250 € | 1 041 906,74 € |
| MIG versée | 939 753 € | 605 690,00 € |
| PCA N-1 | 351 249 € | 465 751,57 € |
| Total Recettes | 1 291 002 € | 1 071 441,57 € |
| Résultat net | 465 752 € | 29 534,83 € |

2. Ressources humaines

Un renforcement via des recrutements sera recherché notamment au sein de l'équipe technique (fortement sollicitée par l'étoffement de BaMaRa et de la BNDMR), de l'équipe data (notamment les différents projets d'IA), ainsi que sur la mise à jour de la terminologie Orphanet.



La cellule opérationnelle fera en outre appel à des prestataires externes tant au titre de l'accompagnement sur le développement de la BNDMR et de l'interopérabilité de BaMaRa que sur les projets relatifs à la fœtopathologie et à l'errance et l'impasse diagnostiques.