



BNDMR

Banque Nationale de Données
Maladies Rares

Rapport d'enquête nationale Bases de données maladies rares en France

Etat des lieux auprès des centres de référence maladies rares
Septembre 2015

Céline Angin, Amélie Ruel, Claude Messiaen, Rémy Choquet, Paul Landais
Banque Nationale de Données Maladies Rares (BNDMR)

Ce document a été produit par l'équipe de la Banque Nationale de Données Maladies Rares (BNDMR) dans le cadre d'un projet d'accompagnement des filières dans l'identification de leurs bases de données, suite à une proposition de la Direction de l'organisation des soins du Ministère de la Santé.

Plus d'informations sur la BNDMR et ses actions sont disponibles sur le site www.bndmr.fr.

Remerciements :

Nous souhaitons remercier chacun des participants pour le temps qu'ils ont pu nous accorder afin de répondre à cette enquête.

Nous remercions également les chefs de projet et animateurs des filières de santé maladies rares pour leur aide précieuse. Ils ont facilité la diffusion du questionnaire à leur réseau et ont su être disponibles pour les diverses questions qui en découlaient. Par ailleurs, nous souhaitons souligner l'engagement du groupe de travail qui nous a aidés à élaborer le questionnaire de l'enquête, et les remercier pour leurs contributions.

Enfin, ce travail a bénéficié des conseils du Pr Paul Landais sur les questions de méthodologie et d'épidémiologie. Nous lui adressons ici notre gratitude.

Avertissement :

Les résultats et conclusions de cette enquête découlent de données déclaratives. Les informations sur les bases de données rapportées dans ce rapport n'ont pas pu être vérifiées et sont sous la seule responsabilité des répondants.

Copyright :

Ce document et ses illustrations sont propriété de la BNDMR. Ils peuvent être utilisés et incorporé à d'autres matériels sous conditions :

- le contenu ne doit en aucune sorte être modifié
- aucun profit ne soit tiré de son utilisation (ne pas vendre le produit issu de l'usage de données extraites de ce rapport)
- la source des informations doit être clairement identifiée comme suit :

Céline Angin, Amélie Ruel, Claude Messiaen, Rémy Choquet, Paul Landais, **Rapport d'enquête nationale sur les bases de données maladies rares en France, Etat des lieux auprès des centres de référence maladies rare**, Septembre 2015. <http://bndmr.fr>

Table des matières

Contexte	4
Objectifs	4
Périmètre de l'enquête	5
Méthodologie	5
1. <i>Phase préparatoire</i>	5
2. <i>Modalités de collecte</i>	6
3. <i>Modalités d'exploitation</i>	6
Résultats	7
1. <i>Combien de bases de données maladies rares existe-t-il en France ?</i>	7
2. <i>Dans quels objectifs ont été créées les bases de données ?</i>	15
3. <i>Comment sont gérées les bases de données ?</i>	16
4. <i>Quelles sont les caractéristiques techniques des bases de données ?</i>	18
5. <i>Quelles données sont collectées dans les bases de données ?</i>	21
5.1. <i>Couverture et exhaustivité</i>	21
5.2. <i>Les données</i>	23
5.3. <i>Interopérabilité</i>	25
6. <i>Répondent-elles aux exigences éthico-juridiques ?</i>	28
7. <i>Quelle est la charge de travail qui en découle ?</i>	29
CEMARA	33
Discussion	37
Conclusion	39
Annexe 1 : liste des bases de données déclarées dans l'enquête	40
Annexe 2 : données brutes par filière hors CEMARA	44

Contexte

Les centres maladies rares doivent à la fois assurer une mission de soin, de recherche et de surveillance épidémiologique. Pour répondre à ces missions, de nombreuses bases de données, aux finalités différentes, ont été créées.

Un premier recensement de ces bases de données avait eu lieu auprès des centres de référence en 2012, lors de la structuration de la BNDMR. Cette enquête avait permis de souligner leur grand nombre, leur hétérogénéité et les difficultés rencontrées pour leur mise en œuvre (collecte, qualité, exploitation) et leur maintenance compte tenu de l'absence d'interopérabilité des systèmes. Cependant, cette étude globale a montré ses limites puisque les résultats étaient collectés par centre de référence et non par base de données.

Dans le cadre de la mise en place des filières et de la BNDMR, un nouveau travail de recensement des bases de données existantes dans le domaine des maladies rares a été identifié comme nécessaire. Proposé par la Direction générale de l'offre de soins, ce recensement par base de données a pu être l'occasion pour les filières de mieux définir leur plan d'action pluriannuel.

Enfin, d'un point de vue national, il contribuera à mieux comprendre l'organisation du recueil de données maladies rares en France et ainsi de situer la France dans le contexte international.

Objectifs

L'objectif était de mieux décrire les bases de données, leurs finalités, les catégories de données collectées et la charge de travail induite.

Pour cela une série de questions ont été posées afin de répondre à ces objectifs :

- 1- Combien de bases de données maladies rares existe-t-il en France ?
- 2- Dans quels objectifs ont-elles été créées ?
- 3- Comment sont-elles gérées (maintenance, gouvernance, financement) ?
- 4- Quelles sont leurs caractéristiques techniques (format, interopérabilité, hébergement) ?
- 5- Quelles données y sont collectées (couverture, quantité, exhaustivité) ?
- 6- Répondent-elles aux exigences éthico-juridiques ?
- 7- Quelle est la charge de travail qui en découle ?

Périmètre de l'enquête

Cette enquête concernait toute base de données dans laquelle les professionnels des centres de référence maladies rares en France collectent des données sur leurs patients atteints de maladies rares :

- Les applications spécifiques à des spécialités (génétique, biologie...)
- Toutes les autres bases de données (registres, cohortes, observatoires, PHRC, bases Excel, Access...) y compris CEMARA (questionnaire simplifié), qu'elles soient françaises, européennes ou internationales

Méthodologie

1. Phase préparatoire

Une évaluation de l'intérêt d'un recueil uniformisé au niveau national a été conduite par la BNDMR auprès de quelques filières. Le besoin d'accompagnement méthodologique et technique ainsi qu'une volonté d'harmonisation entre filières pour rendre cette enquête utile et exploitable ont été rapportés. Certaines filières avaient en revanche déjà fait un état des lieux.

La BNDMR a alors proposé d'apporter un soutien logistique et technique aux filières de santé pour mettre en place un recueil national standardisé.

Le questionnaire de l'enquête a été élaboré en collaboration avec un groupe de travail composé de chefs de projet de filières, afin de permettre de répondre aux questions identifiées au préalable (voir Objectifs).

Le questionnaire se compose d'une trentaine de questions, répartis en 6 axes :

- A. Identification
- B. Gouvernance
- C. Aspects éthico-juridiques
- D. Aspects techniques / informatiques
- E. Saisie et collecte
- F. Qualité et exploitation des données

Dans le cas de la base CEMARA, le questionnaire était simplifié et ne comportait que les points A, E et F.

2. Modalités de collecte

L'enquête a été mise à disposition par voie électronique sur la plateforme Zoho Survey, ceci afin de faciliter le recueil et d'homogénéiser l'exploitation. Le recueil a été ouvert du 20 mai au 9 juillet 2015.

Il était demandé de compléter l'enquête pour chacune des bases existantes dans les centres de référence – sites coordonnateurs et sites constitutifs – soit un questionnaire par base.

Les chefs de projet pouvaient soit remplir eux-mêmes l'enquête en s'appuyant sur les coordonnateurs de centres de référence, soit s'organiser avec les coordonnateurs pour assurer ce recueil.

Dans le cadre de cette enquête, les chefs de projet filière et leurs équipes ont eu un rôle essentiel de relais de l'information et d'accompagnement au recueil des résultats.

La BNDMR s'est chargée d'envoyer l'enquête aux centres de référence hors filières. Une relance a été effectuée début juin.

3. Modalités d'exploitation

Chaque chef de projet et animateur avait accès à l'ensemble des résultats recueillis pour sa filière, directement en ligne dans l'outil Zoho Survey. Ils avaient la possibilité de modérer les réponses afin de les rendre exploitables (fautes de frappe, réponse incomplète ou erronée, etc.). Ils pouvaient ainsi suivre à tout moment les résultats, agrégés mis en forme ou individuels, et les exporter afin d'établir des priorités en terme de bases de données (nouveau recueil, adaptation, interopérabilité pour alléger la charge de saisie, etc.) dans le cadre de leur plan d'action pluriannuel.

La BNDMR avait accès à tous les recueils de manière à suivre les résultats et accompagner les chefs de projet filière en cas de besoin.

Résultats

Rappel important :

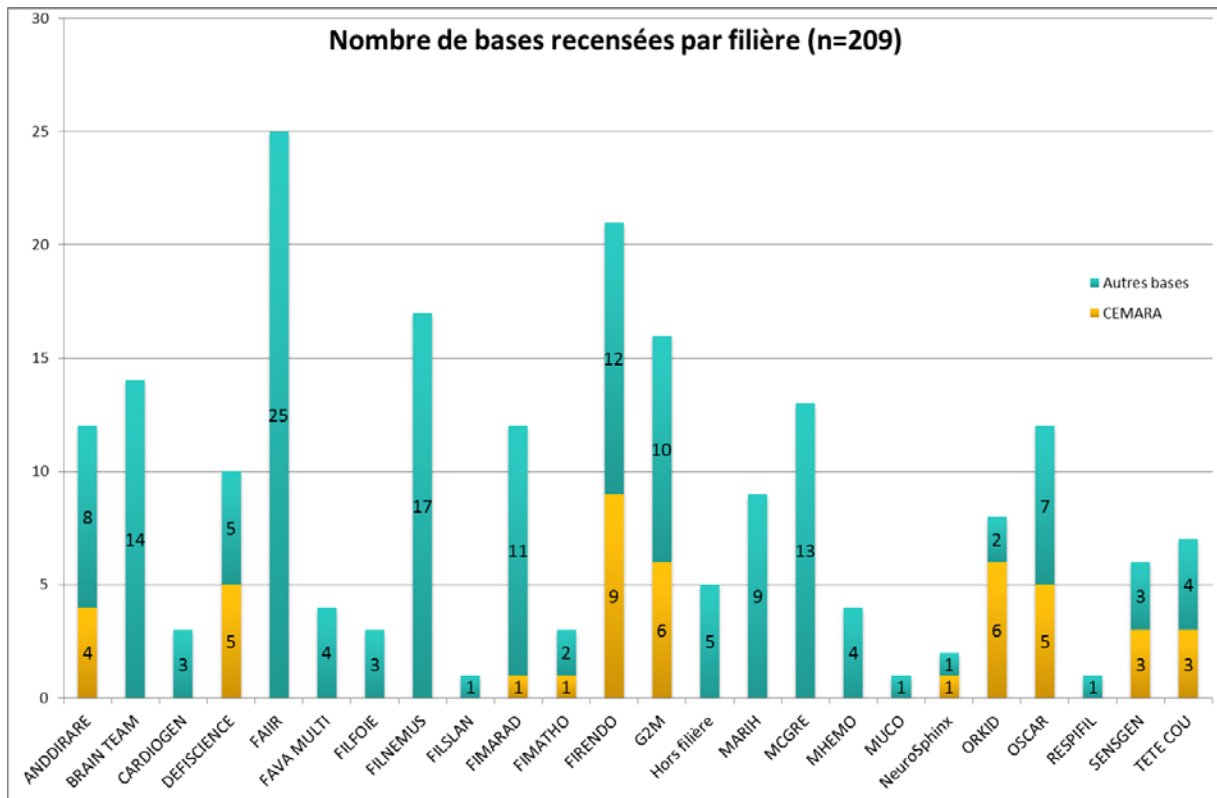
Ce recensement n'était pas obligatoire et chaque filière était libre d'utiliser l'outil mis à leur disposition de la manière qui lui semblait la plus adaptée à ses besoins. Nous soulignons ici que les résultats obtenus ne sont pas exhaustifs. Nous observons une très grande hétérogénéité entre les filières, due à des situations différentes sur le terrain mais aussi à des modes de recueil des informations différents.

1. Combien de bases de données maladies rares existe-t-il en France ?

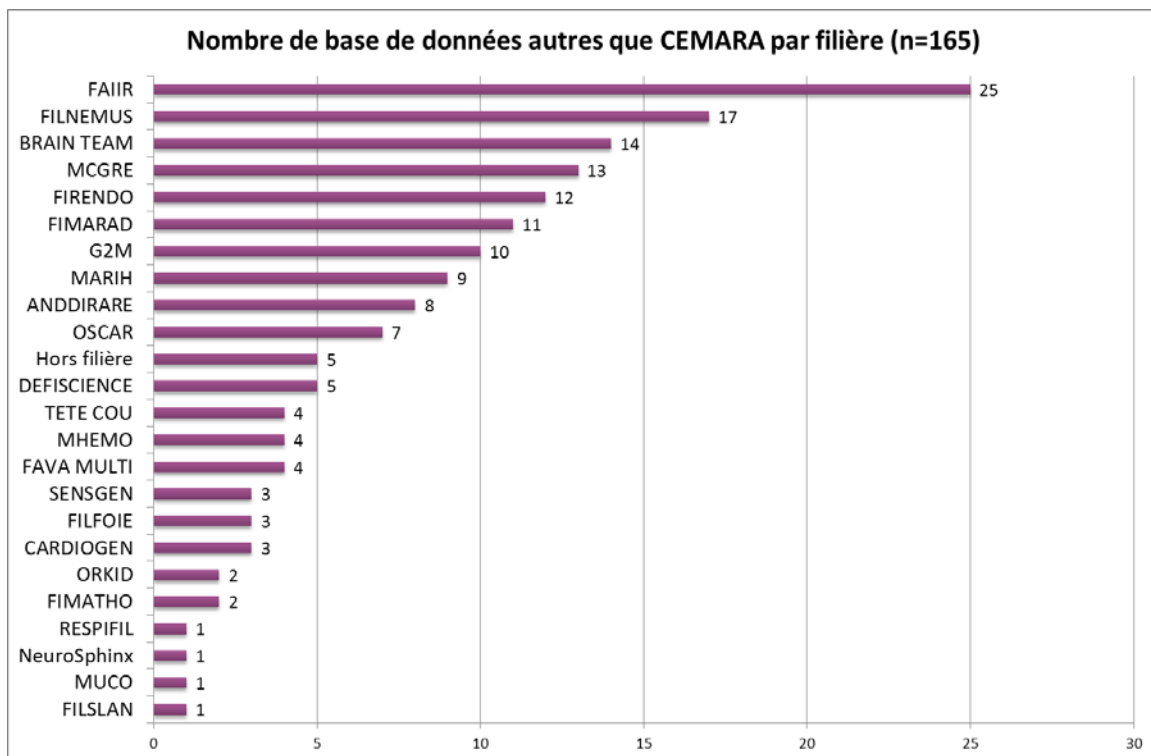
Nous avons collecté **227 réponses**, dont 168 étaient complètes. Parmi ces réponses, 18 ont été supprimées pour les analyses. Ces questionnaires étaient soit entrés en double (souvent suite à une première saisie incomplète, puis une nouvelle saisie complète), soit avaient été complétés pour signaler qu'aucune base de données n'existait pour le centre de référence concerné. Nous avons donc 209 réponses exploitables.

Seules 3 filières n'ont pas directement participé (FIMARAD, FILNEMUS et MARIH) mais nous ont transmis leur propre état des lieux. Ces bases de données ont été reportées dans le questionnaire par l'équipe de la BNDMR, bien que toutes les questions n'aient pu être complétées.

La base CEMARA a été citée 44 fois, dans 10 filières différentes. CEMARA se compose d'un tronc commun et peut être complétée de recueil de données complémentaires appelés pétales. CEMARA a été déclarée soit une seule fois pour la filière lorsque la complétion de l'enquête se faisait au niveau filière, soit pour chaque centre de référence lorsque la complétion était déléguée au coordonnateurs de CRMR. Les pétales et le tronc commun ont été rapportés sans distinction. C'est pourquoi pour une filière, la même base a pu être déclarée plusieurs fois.



Au total, **165 bases de données autres que CEMARA** ont été recensées.



La liste des 165 bases de données déclarées (données brutes) se trouve en annexe 1 (cf. page 39). Parmi elles, **29 bases sont européennes ou internationales**.

- **Redressement des résultats**

Afin d'évaluer le nombre de bases qui n'ont pas été recensées par ce questionnaire, nous avons tenté de rapprocher les résultats de notre enquête à d'autres recensements : ceux effectués par Orphanet¹ et par le portail des registres d'Aviesan².

Au 20 juillet 2015, la liste d'Orphanet des registres de patients maladies rares comportait 134 entrées, dont 28 concernaient des cancers. Nous les avons donc retirées et obtenons une liste de 106 entrées.

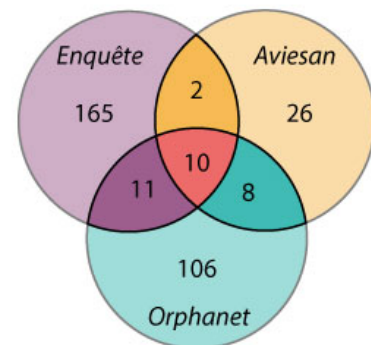
A la même date, la liste de registres identifiés « maladies rares » du portail Epidémiologie – France comportait 41 entrées, dont 11 n'étant pas des registres de maladies (par exemple le PMSI), 1 concernant un cancer et 3 ne concernant pas les maladies rares. Nous les avons donc retirées et obtenons une liste de 26 entrées.

21 bases de données ont pu être réalignées avec les 106 présentes dans Orphanet, et 12 avec les 26 identifiées dans Aviesan (dont 10 bases de données partagées par Orphanet et Aviesan).

	Enquête	Orphanet	Aviesan
Enquête	165	21	12
Orphanet	21	106	18
Aviesan	12	18	26

Au total, 91 potentielles bases supplémentaires ont ainsi pu être identifiées dans Orphanet et Aviesan.

Le faible taux de recouvrement peut être expliqué par le fait que nous avons comparé notre liste à la liste des registres français uniquement. Or dans la liste des bases de données recueillies dans notre enquête, 29 sont européennes ou internationales. De plus, nous n'avons comparé que la liste des registres de patients recensés par Orphanet (et non les registres de mutations). Au moins 3 bases de données de l'enquête sont des bases de mutations (UMD).



D'autre part, il nous a parfois été donné le nom du logiciel plutôt que le nom de la base, ce qui a pu rendre les alignements difficiles. L'une des cohortes RaDiCo a été déclarée dans notre enquête bien que celle-ci ne soit pas encore active. Nous pouvons ajouter les 15 autres cohortes RaDiCo³ au calcul final du nombre de base de données afin d'homogénéiser le calcul.

Par ailleurs, le travail ultérieur de réattribution des bases de données listées dans Orphanet et Aviesan à chaque filière nous a permis un contrôle qualité et un meilleur alignement. Pour cela, une liste de bases pouvant potentiellement être rattachées à chaque filière a été adressée aux chefs de projet filière. Les filières concernées ont ainsi pu nous confirmer que

¹ <http://www.orpha.net/orphacom/cahiers/docs/GB/Registries.pdf>

² <https://epidemiologie-france.aviesan.fr/epidemiologie-france/catalogue>

³ <http://radico.fr/fr/connaitre-radico/nos-cohortes-et-autres-programmes-associes>

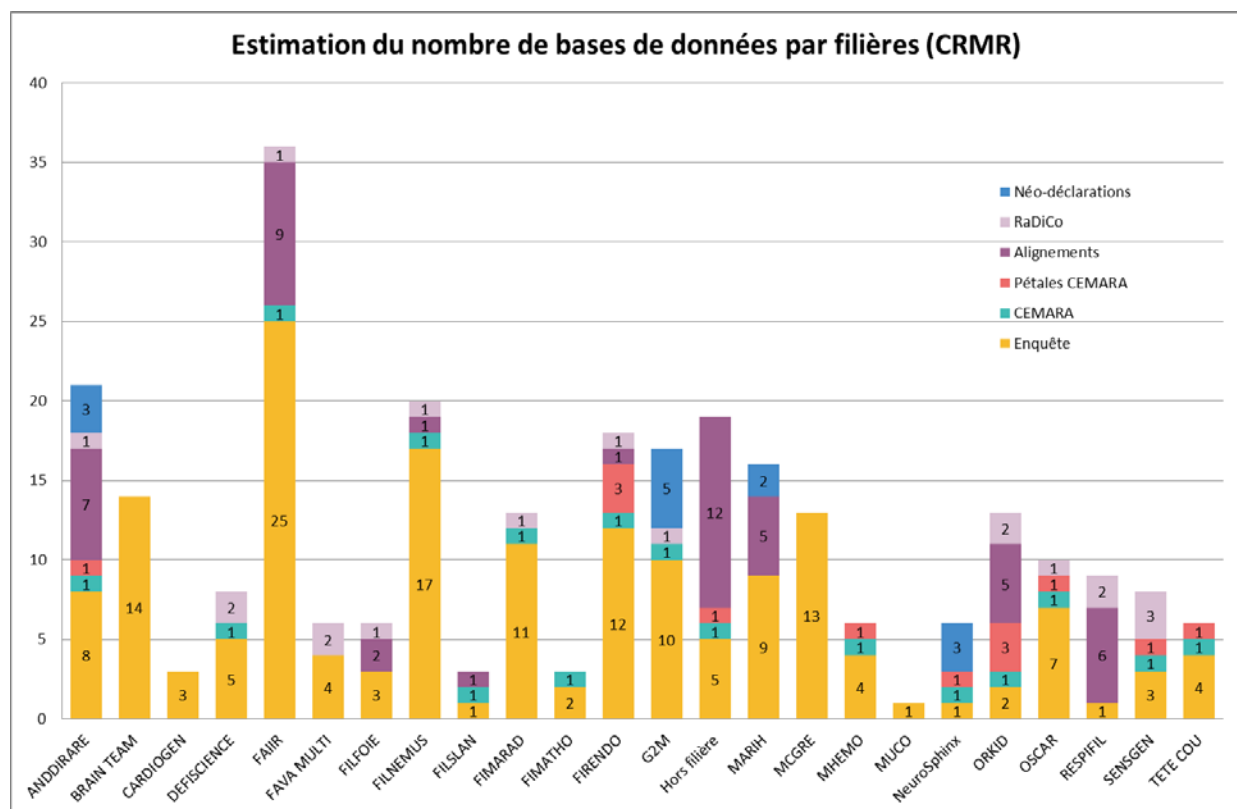
ces bases étaient bien complétées au sein de leurs centres. Certaines filières ont saisi l'occasion pour nous signaler l'existence d'autres bases de données qui n'avaient pas été recensées dans l'enquête (« néo-déclarations »).

Au final, 49 bases présentes soit dans Orphanet soit dans Aviesan ont pu être attribuées à une filière (ou à un centre de référence hors filière), et 42 n'ont pas pu être attribuées.

Une partie de ces bases de données non attribuées sont issus de PHRC. Ces bases de données ne sont remplies que le temps du protocole et ne sont plus actives par la suite. Après concertation avec les chefs de projet des filières de santé, il a été décidé de ne conserver dans nos résultats que les bases encore actives. Plusieurs bases de données dans Orphanet sont des bases d'industriels et n'ont pas été alignées ici. De plus, 2 bases n'ont jamais vu le jour.

Enfin, certaines bases citées dans Orphanet ou Aviesan n'ont pas pu être attribuées à une filière car les pathologies concernées n'entraient pas directement dans le champ de la filière. En revanche, ces pathologies sont généralement bien prises en charge par les CRMR. C'est le cas des diabètes rares qui ne sont pour le moment pas rattachés à la filière FIRENDO. Les bases de données ont alors été rattachées à la catégorie « hors filière ».

Toutes sources confondues (y compris CEMARA), la répartition des bases de données des CRMR dans les filières reste hétérogène comme le montre le graphique suivant.



Il nous a été indiqué qu'il existe en fait une multitude de bases de données (plus ou moins importantes en taille) relatives à la Mucoviscidose. La filière a fait le choix de ne déclarer

pour cette enquête que la principale (Registre National de la Mucoviscidose). La filière G2M a, quant à elle, inclus des réponses de CCMR dans les résultats.

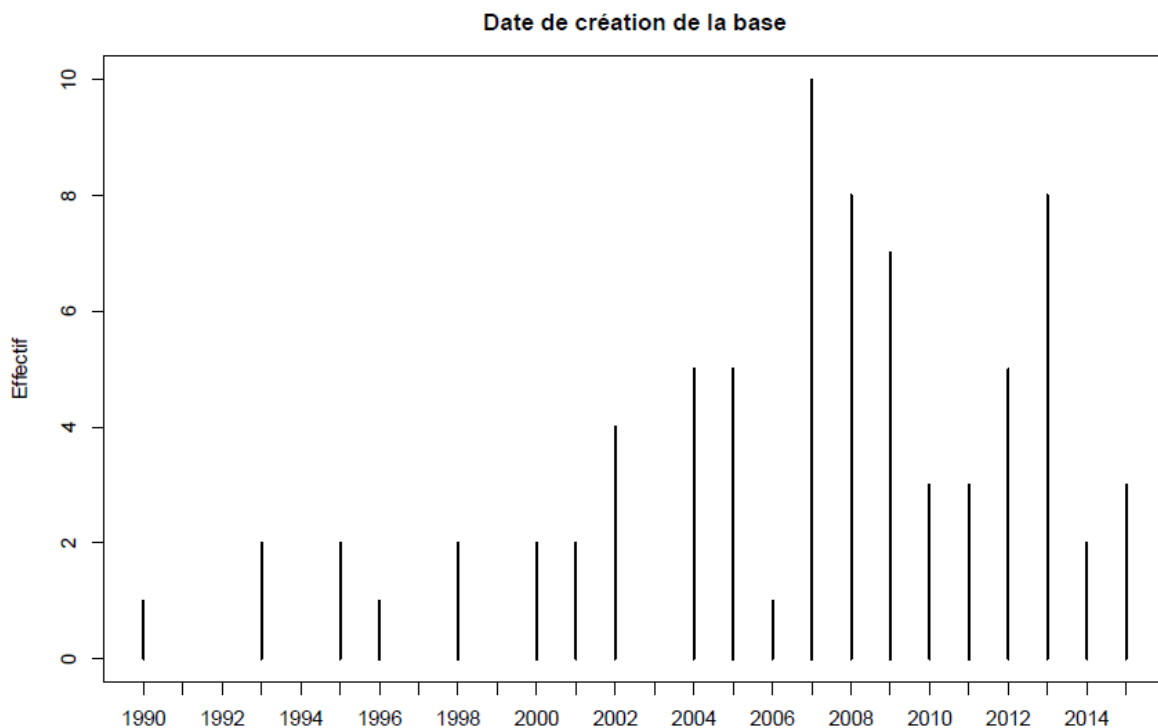
Une enquête complémentaire (et facultative) est en cours et révèle que des dizaines de nouvelles bases de données, souvent au format Excel mais aussi plus structurées, sont complétées par les CCMR. Au moins 2 bases d'Orphanet sont des bases issues de CCMR mais les filières n'ont pas souhaité les compter ici.

Une seule base de données est partagée entre 2 filières. Il s'agit de la base Spina Bifida, qui a été déclarée par les filières Anddirare et NeuroSphinx. Ces 2 filières partagent également un pétale CEMARA.

Nombre de bases de données maladies rares (hors cancers rares) identifiées en France :
 165 (enquête) + 1 (CEMARA) + 12 (pétales CEMARA) + 49 (alignements) + 15 (RaDiCo) +
 13 (néo-déclarations) – 1 (partagée entre 2 filières) = **254 bases de données**

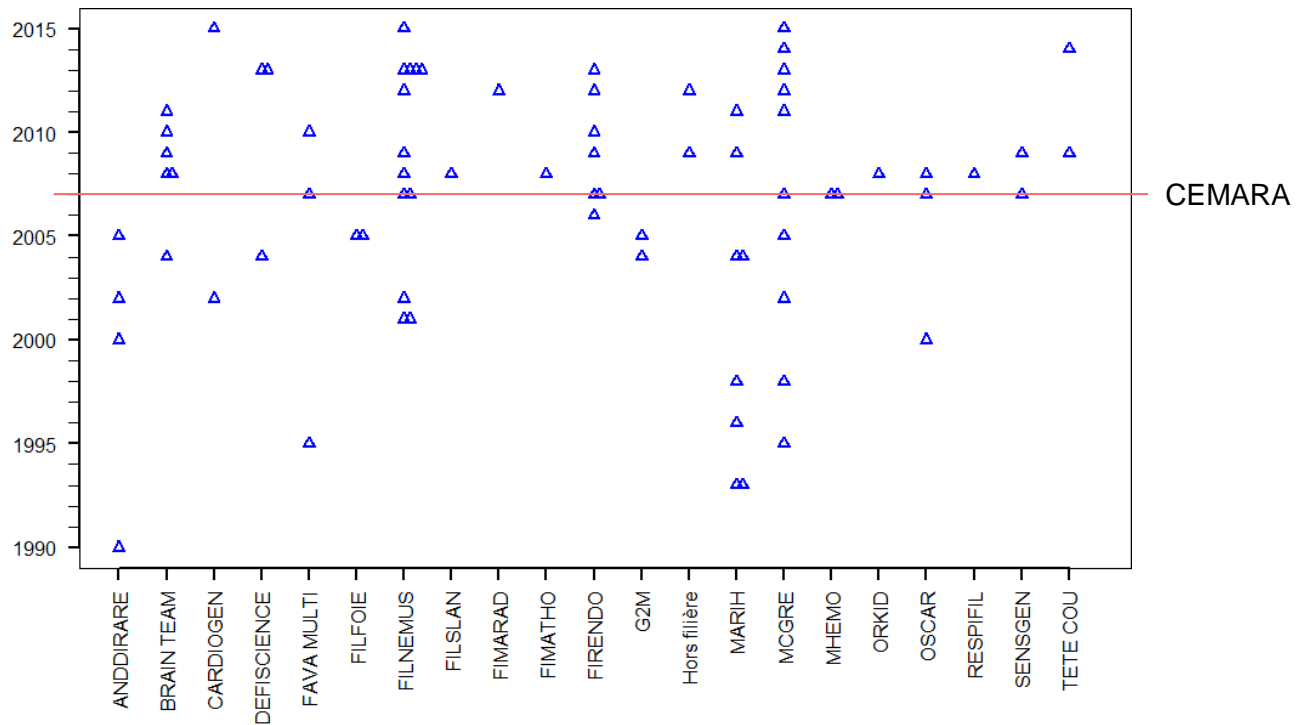
- **Dates de création des bases de données**

Cette question n'était posée que lorsque la personne complétant l'enquête indiquait être le concepteur de la base de données. Nous avons collecté **76 réponses**.



C'est en 2007 que le plus grand nombre de bases de données ont été créées (10). La base CEMARA a été créée en 2006 et ouverte le 23 mai 2007.

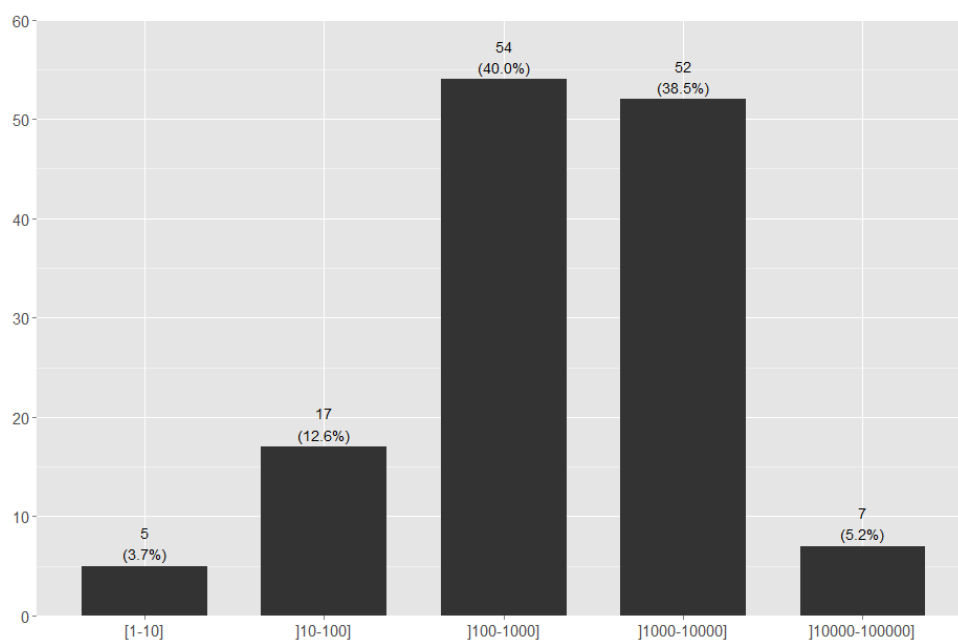
Nombre de création de bases de données par année et par filière (n=76)



- Nombre de patients**

Le nombre de patients par base de données varie de 5 patients à 100 000 patients (Banque nationale Alzheimer – BNA - qui contient en très grande majorité des cas non rares), hors base CEMARA. La médiane est de 793 patients et la moyenne de 3419 patients. Sans la BNA, la médiane est de 784 patients et la moyenne de 2098 patients par base.

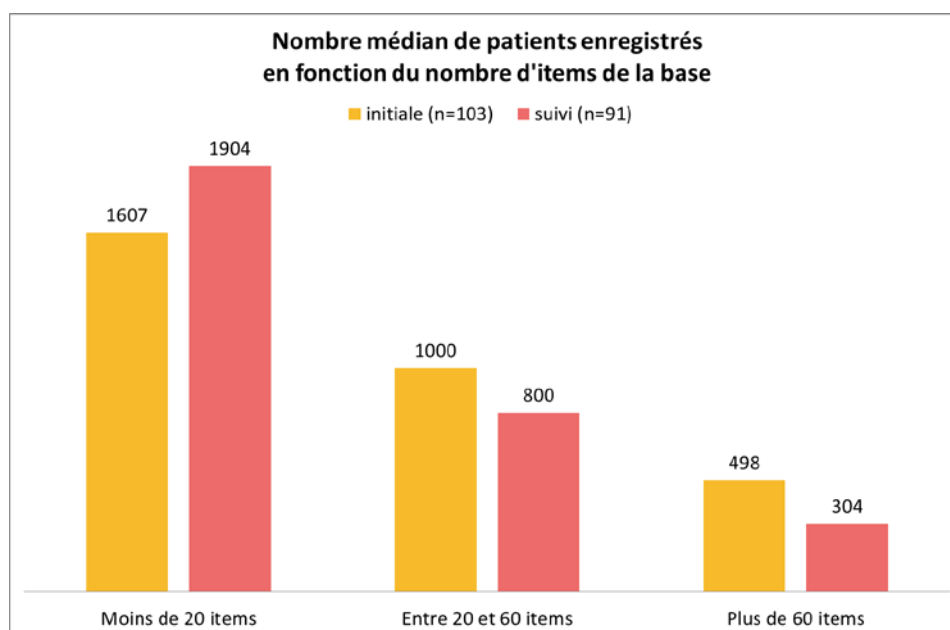
Nombre de bases de données en fonction du nombre de patients (n=135)



Une majorité des bases de données comptent entre 100 et 500 patients (41).

Nous pouvons estimer le nombre de patients dans les bases de données déclarées à **environ 360 000**. Pour ce calcul, nous n'avons pas pris en compte la BNA (soit 100 000 patients). Comme la BNA, d'autres bases de données non spécifiques aux maladies rares ont été déclarées à juste titre puisqu'incluant également des patients maladies rares. Nous ne pouvons pas à ce jour déterminer la proportion de patients maladies rares dans ces bases, cela nécessiterait une étude complémentaire. De même, pour les bases européennes et internationales, qui représentent 53700 patients (soit 15%), il ne nous est pas possible de déterminer la proposition de patients français.

Notre estimation est donc probablement surévaluée. Cependant, aucun effectif n'a été renseigné pour 30 bases de données déclarées. Cela représenterait **25 000** patients supplémentaires (calcul sur la médiane avec une distribution du nombre de patients en fonction du nombre d'items de la base de données en saisie initiale).



Plus une base de données contient d'items plus le nombre médian de patients est restreint.

Enfin, les bases supplémentaires que nous avons pu attribuer aux filières par croisement avec Orphanet et Aviesan (49), ainsi que les nouvelles bases déclarées a posteriori (13), représenteraient **48 000** patients supplémentaires (calcul sur la médiane).

Le nombre de patients (malades et porteurs sains) saisis dans CEMARA est de **310 000**.

Nous estimons ainsi qu'il y existe environ **743 000 patients** dans les bases de données citées plus haut (360 000 + 25 000 + 48 000+ 310 000). Enfin, plus de **7 000** patients sont attendus dans les cohortes RaDiCo, ce qui porterait ce chiffre à 750 000 patients.

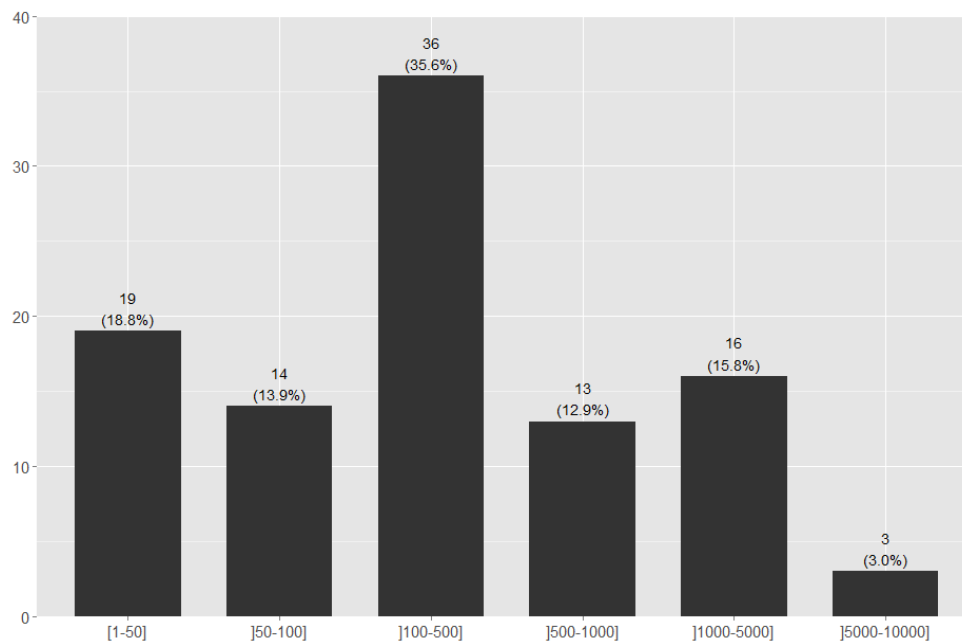
Ce chiffre ne prend pas en compte les doublons potentiels entre les bases.

Il ne prend pas non plus en compte les dossiers saisis dans la fiche maladie rare d'ORBIS (dossier patient informatisé), déployé en phase pilote à l'hôpital Bicêtre (soit environ 800 patients).

- **File active 2014**

La file active par base de données, soit le nombre de patients différents vus au cours de l'année 2014, varie de 2 patients à 6570 patients, hors base CEMARA. La médiane est de 264 patients et la moyenne de 706 patients.

Nombre de bases de données en fonction du nombre de patients pour la file active (n=101)



Au total, dans les bases de données citées dans l'enquête, la file active est d'environ **71 000** patients. Cependant, aucun effectif n'a été renseigné pour 30 bases de données déclarées. Cela représenterait environ **16 000** patients supplémentaires (calcul sur la médiane avec une distribution du nombre de patients en fonction du nombre d'items de la base de données en saisie initiale).

Enfin, les bases supplémentaires que nous avons pu attribuer aux filières par croisement avec Orphanet et Aviesan (49), ainsi que les nouvelles bases déclarées à posteriori (13), représenteraient **16 000** patients supplémentaires (calcul sur la médiane).

Dans CEMARA, la file active en 2014 était de **87 000** patients (y compris non malades).

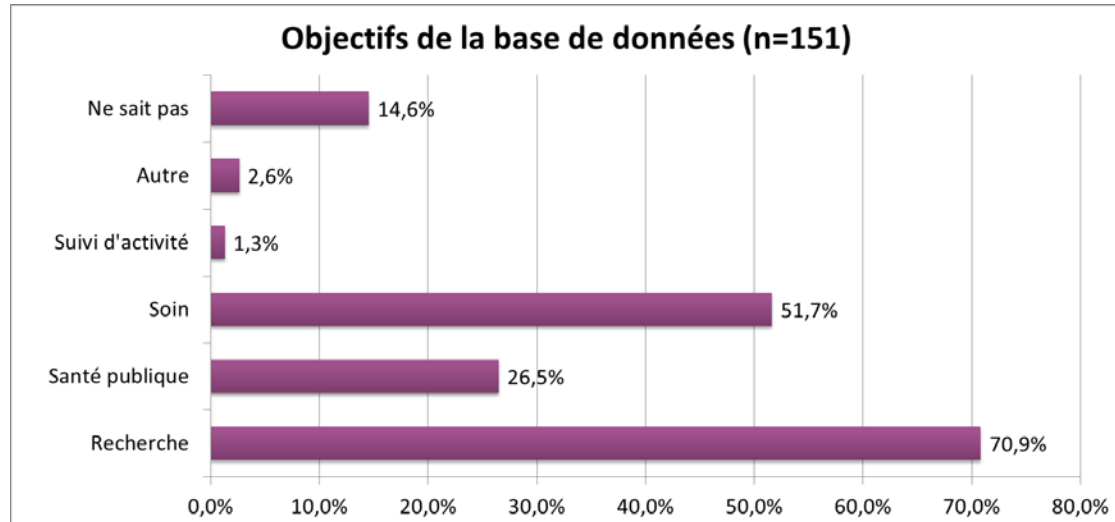
Nous estimons ainsi qu'il y existe environ **190 000 patients** dans les bases de données citées plus haut (71 000 + 16 000 + 31 000 + 87 000). Ce chiffre ne prend pas en compte les doublons potentiels entre les bases.

La file active représenterait donc 25% de la cohorte totale.

Les analyses suivantes se feront sur les 165 bases de données autres que CEMARA. Le cas de CEMARA est traité à part (voir page 39).

2. Dans quels objectifs ont été créées les bases de données ?

Cette question était à choix multiple, les objectifs pouvaient donc être combinés. Plus de 70% des bases de données déclarées ont été créées à des fins de recherche, et plus de 50% à des fins de soin.



Les objectifs déclarés dans la catégorie « Autre » sont : épidémiologie (2), administratif, et listing de patient avec les coordonnées (nom/prénom/date de naissance) et le diagnostic.

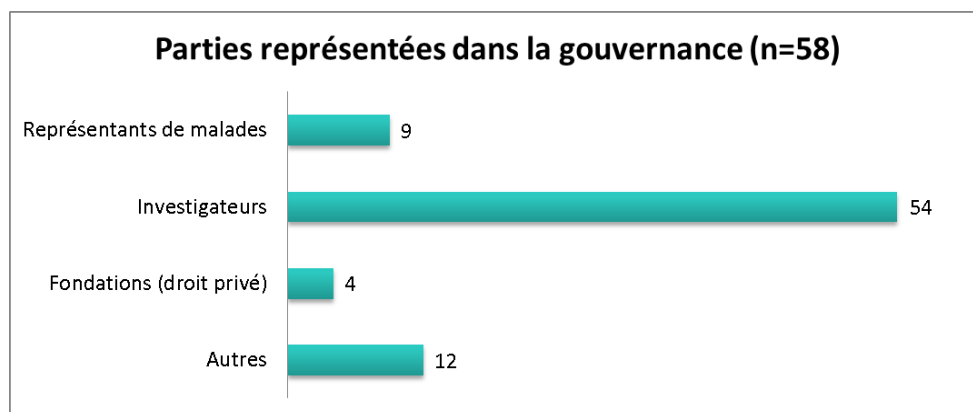
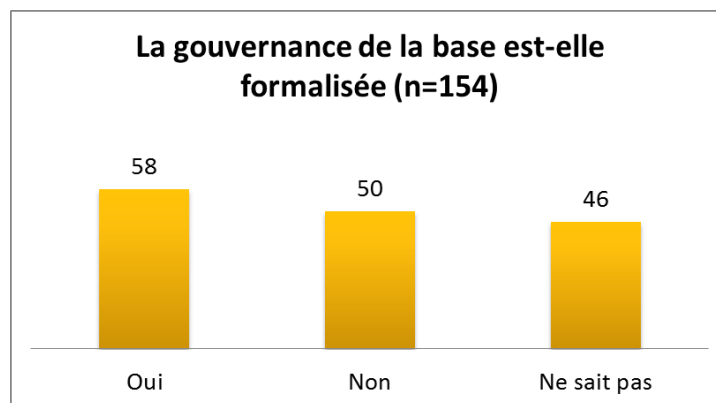
Objectifs de la base de données (nombre de réponse par filière)						
	Recherche	Santé publique	Soins	Suivi d'activité	Autre	
ANDDIRARE	6	1	5	0	0	
BRAIN TEAM	14	5	5	0	0	
CARDIOGEN	3	0	2	0	0	
DEFISCIENCE	4	4	3	0	0	
FAIIR	25	5	20	0	0	
FAVA MULTI	3	2	4	0	0	
FILFOIE	1	0	1	0	1	
FILNEMUS	5	2	3	0	0	
FILSLAN	1	1	0	1	0	
FIMARAD	1	0	2	0	0	
FIMATHO	2	1	0	0	0	
FIRENDO	11	1	3	0	2	
G2M	0	4	5	1	1	
Hors filière	4	1	3	0	0	
MARIH	0	0	0	0	0	
MCGRE	11	6	9	0	0	
MHEMO	2	2	2	0	0	
MUCO	1	1	1	0	0	
NeuroSphinx	1	1	0	0	0	
ORKID	2	0	0	0	0	
OSCAR	2	0	5	0	0	
RESPIFIL	1	1	1	0	0	
SENSGEN	3	0	1	0	0	
TETE COU	4	2	3	0	0	
total	107	40	78	2	4	

Plus de 70% des bases de données déclarées ont été créées à des fins de recherche, et plus de 50% à des fins de soin. Les objectifs pouvaient être combinés.

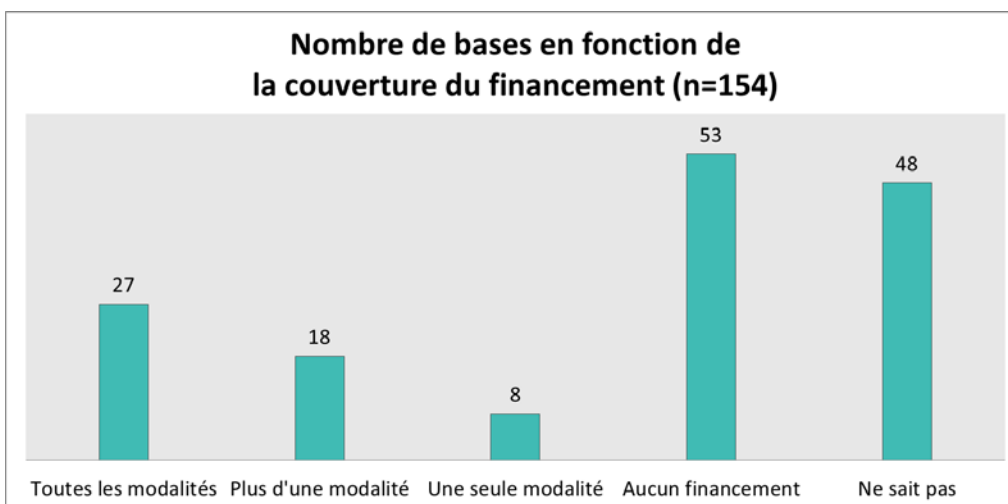
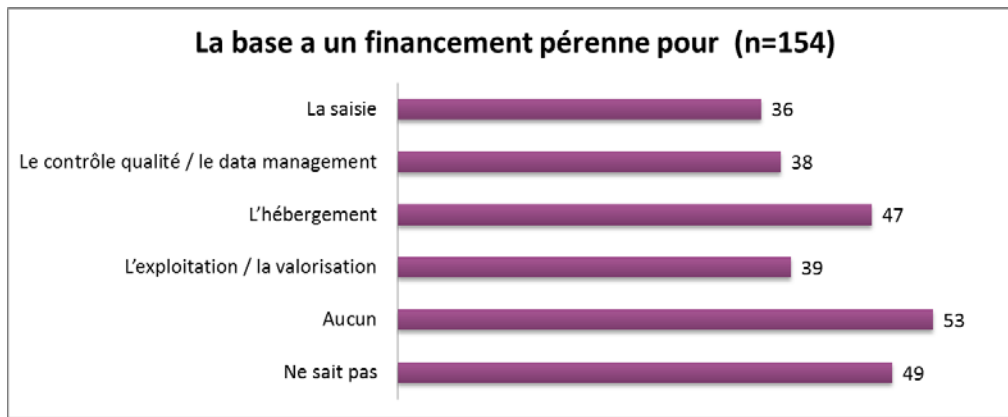
3. Comment sont gérées les bases de données ?

Plus d'un tiers des bases de données déclarées ont une gouvernance formalisée. Lorsque celle-ci existe (cette question était à choix multiple), les investigateurs sont représentés dans les comités mis en place dans plus de 90% des cas. Les représentants de malades le sont dans 15% des cas.

On notera qu'un tiers des répondants ne savaient pas si une gouvernance était en place pour les bases de données qu'ils utilisaient.



Lorsque la base de données possède un financement pérenne, celui-ci est en majorité pour l'hébergement (cette question était à choix multiple). La saisie est le plus souvent non financée. Près de 35% des bases n'ont pas de financement pérenne. Certaines bases de données possèdent néanmoins des moyens pour financer leur base. Si seules 5% des bases ont un financement pour une seule des modalités proposées dans l'enquête, plus de 29% ont un financement pour au moins deux de ces modalités. Vingt-sept bases, soit 17% des réponses, ont un financement concernant toutes les modalités, permettant d'assurer leur pérennité. Enfin, 32% ne savent pas si un financement pérenne existe pour leur base.



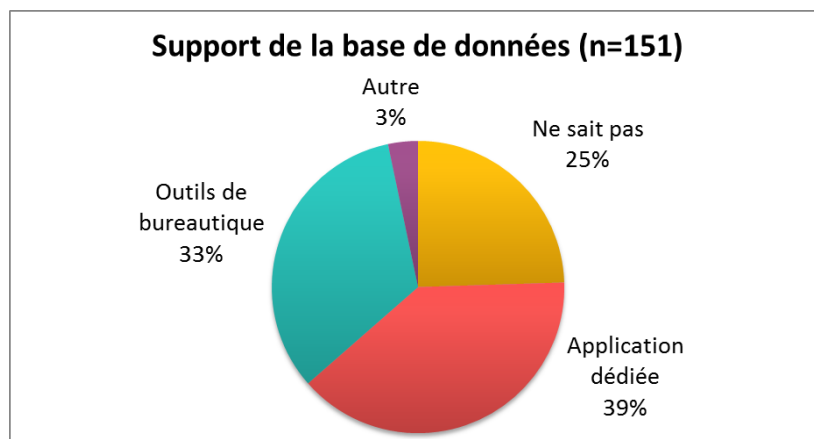
Financement pérenne pour la base de données (nombre de réponse par filière)							
	Saisie	Hébergement	Contrôle qualité / data management	Exploitation / valorisation	Aucun	Ne sait pas	
ANDDIRARE	1	3	2	2	4	1	
BRAIN TEAM	3	6	3	6	4	3	
CARDIOGEN	1	1	0	0	2	0	
DEFISCIENCE	0	0	0	0	4	0	
FAIIR	19	19	19	19	2	4	
FAVA MULTI	0	0	1	0	2	0	
FILFOIE	2	2	1	1	1	0	
FILNEMUS	0	0	0	0	1	16	
FILSLAN	0	0	1	1	0	0	
FIMARAD	0	0	0	0	0	4	
FIMATHO	0	0	0	0	2	0	
FIRENDO	2	3	3	3	8	2	
G2M	0	0	0	0	1	9	
Hors filière	3	1	2	2	2	0	
MARIH	0	0	0	0	0	9	
MCGRE	2	3	2	1	9	1	
MHEMO	2	3	2	3	0	0	
MUCO	1	1	1	1	0	0	
NeuroSphinx	0	1	0	0	0	0	
ORKID	0	0	0	0	2	0	
OSCAR	0	1	0	0	4	0	
RESPIFIL	0	1	1	0	0	0	

Financement pérenne pour la base de données (nombre de réponse par filière)						
	Saisie	Hébergement	Contrôle qualité / data management	Exploitation / valorisation	Aucun	Ne sait pas
SENSGEN	0	0	0	0	3	0
TETE COU	0	2	0	0	2	0
Total	36	47	38	39	53	49

La gestion des bases de données est assez mal connue. Une gouvernance est organisée dans un peu plus d'un tiers des cas. Les investigateurs y sont représentés dans 9 cas sur 10. Des financements pérennes sont disponibles pour différentes finalités mais 35% déclarent n'avoir aucun financement.

4. Quelles sont les caractéristiques techniques des bases de données ?

Les bases de données déclarées dans l'enquête sont en majorité des bases développées sur des applications dédiées (39%), mais un grand nombre reste développé sur des outils de bureautique (33%). Un quart des répondants ne connaissent pas le support de la base de données.

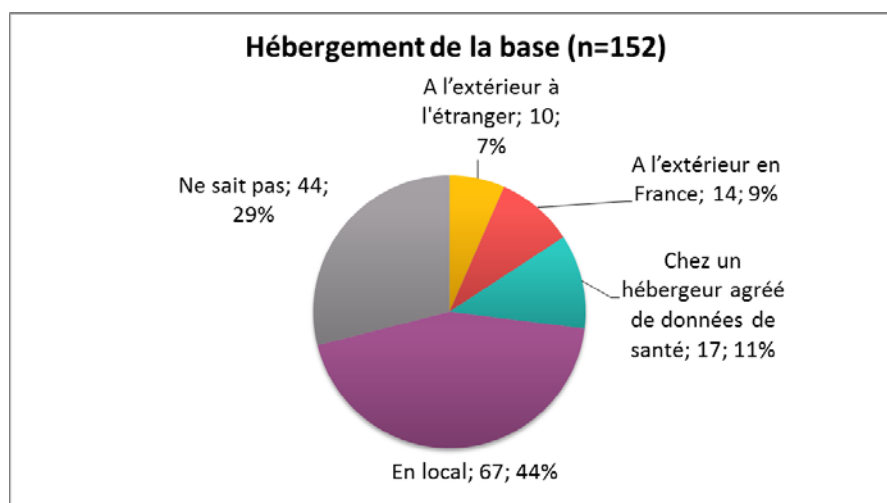


Les « autres » supports sont Lotus Approach, CLININFO, logiciel OCQM (système en ligne), et SPSS. De plus, l'une des bases est en cours de construction dans le cadre d'un projet RaDiCo.

Support de la base de données (nombre de réponse par filière)				
	Application dédiée	Outils de bureautique	Autre	Ne sait pas
ANDDIRARE	6	2	0	0
BRAIN TEAM	9	4	1	0
CARDIOGEN	2	1	0	0
DEFISCIENCE	2	2	0	0
FAIIR	9	0	1	15
FAVA MULTI	1	1	1	0
FILFOIE	2	1	0	0
FILNEMUS	0	0	0	17
FILSLAN	1	0	0	0

Support de la base de données (nombre de réponse par filière)				
	Application dédiée	Outils de bureautique	Autre	Ne sait pas
FIMARAD	2	1	0	0
FIMATHO	2	0	0	0
FIRENDO	5	6	1	0
G2M	0	5	0	4
Hors filière	2	3	0	0
MARIH	1	8	0	0
MCGRE	4	8	1	0
MHEMO	3	1	0	0
MUCO	1	0	0	0
NeuroSphinx	1	0	0	0
ORKID	0	2	0	0
OSCAR	0	3	0	1
RESPIFIL	1	0	0	0
SENSGEN	2	1	0	0
TETE COU	3	1	0	0
Total	59	50	5	37

L'hébergement des bases de données se fait majoritairement au sein de l'hôpital (en local). Les bases sont pour 16% hébergées à l'extérieur de leur propre hôpital, soit en France (9%) soit à l'étranger (7%) ; 11% sont hébergées chez un hébergeur agréé de données de santé.



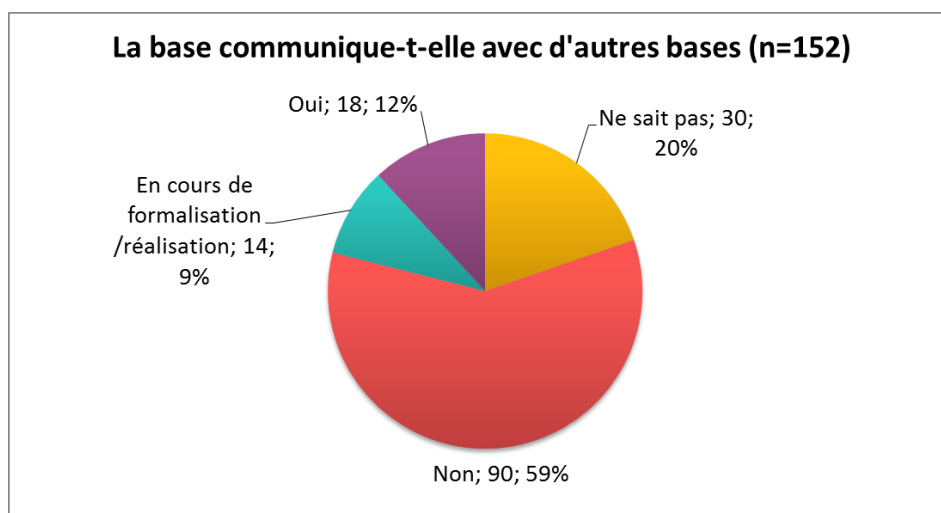
Hébergement de la base de données (nombre de réponse par filière)					
	Chez un hébergeur agréé de données de santé	En local	A l'extérieur en France	A l'extérieur à l'étranger	Ne sait pas
ANDDIRARE	0	6	1	1	0
BRAIN TEAM	2	8	1	2	1
CARDIOGEN	1	1	1	0	0
DEFISCIENCE	0	3	0	0	1
FAIR	8	0	0	1	16
FAVA MULTI	0	0	3	0	0
FILFOIE	2	1	0	0	0
FILNEMUS	0	1	0	0	16

Hébergement de la base de données (nombre de réponse par filière)					
	Chez un hébergeur agréé de données de santé	En local	A l'extérieur en France	A l'extérieur à l'étranger	Ne sait pas
FILSLAN	1	0	0	0	0
FIMARAD	0	0	1	0	2
FIMATHO	0	2	0	0	0
FIRENDO	1	6	2	3	0
G2M	0	7	0	0	3
Hors filière	0	3	1	1	0
MARIH	0	5	0	0	4
MCGRE	0	9	2	1	1
MHEMO	1	1	1	1	0
MUCO	1	0	0	0	0
NeuroSphinx	0	1	0	0	0
ORKID	0	2	0	0	0
OSCAR	0	4	0	0	0
RESPIFIL	0	0	1	0	0
SENSGEN	0	3	0	0	0
TETE COU	0	4	0	0	0
Total	17	67	14	10	44

Les bases de données hébergées à l'étranger le sont dans les pays suivants :

Pays	Nombre de bases
Allemagne	4
Belgique	2
Espagne	1
Suisse	1
Grande-Bretagne	1
USA	1

La majorité des bases de données ne communiquent pas avec d'autres bases (59%). Cependant, 9% indiquent qu'une communication avec une autre base est en cours de formalisation ou de réalisation, et 12% communiquent déjà avec d'autres bases.



La base communique avec d'autres bases (nombre de réponse par filière)		
	Oui	Non, mais c'est en cours de formalisation/réalisation
ANDDIRARE	2	1
BRAIN TEAM	4	1
CARDIOGEN	0	0
DEFISCIENCE	0	0
FAIR	2	0
FAVA MULTI	0	1
FILFOIE	1	0
FILNEMUS	5	2
FILSLAN	0	1
FIMARAD	0	0
FIMATHO	0	0
FIRENDO	1	0
G2M	0	0
Hors filière	0	0
MARIH	0	0
MCGRE	2	1
MHEMO	0	2
MUCO	1	0
NeuroSphinx	0	0
ORKID	0	0
OSCAR	0	1
RESPIFIL	0	1
SENSGEN	0	1
TETE COU	0	2
Total	18	14

Les bases déclarées sont majoritairement des bases développées sur des outils spécifiques, cependant de nombreuses bases de données sont développées sur des outils de bureautique. Elles sont majoritairement hébergées en local (44%) ; Seules 11% sont chez un hébergeur agréé de données de santé. Les bases de données sont 12% à communiquer avec d'autres bases et ce chiffre tend à augmenter puisque 9% sont en cours de formalisation d'une connexion.

5. Quelles données sont collectées dans les bases de données ?

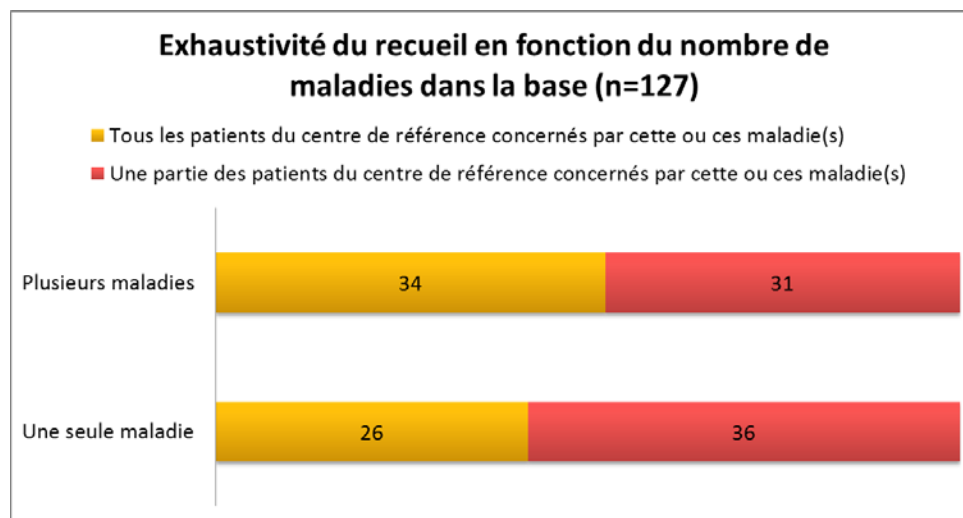
5.1. Couverture et exhaustivité

La base concerne :	Nombre de réponses (n=127)
Plusieurs maladies	65
Une seule maladie rare	62

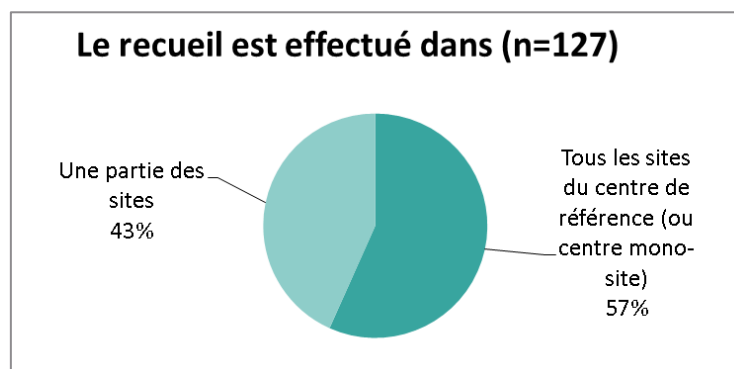
Les bases de données déclarées concernent autant une seule maladie que plusieurs maladies. Le recueil est aussi bien effectué pour une partie des patients que pour tous les patients des centres de référence concernés par la maladie. Une majorité recueille des

données dans tous les sites du centre de référence ou se déclarent centre mono-site. Il s'agit peut-être d'un biais de compréhension de cette notion de mono-site/multi-site.

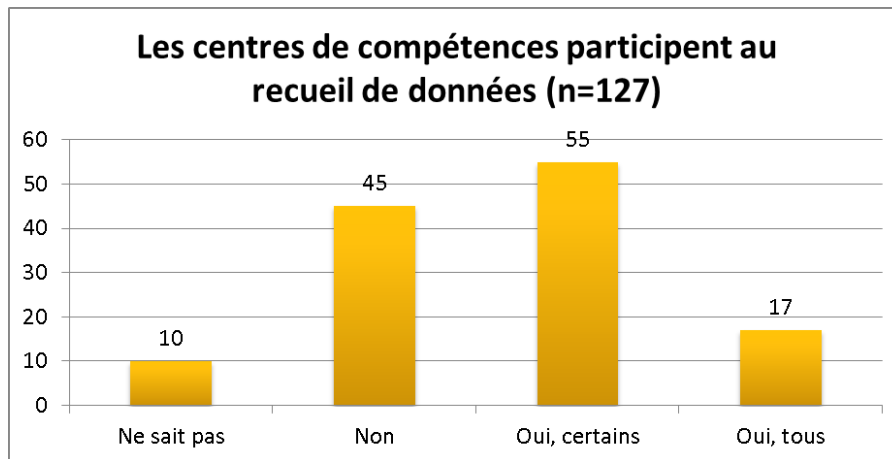
La base concerne :	Nombre de réponses (n=127)
Tous les patients du centre de référence concernés par cette ou ces maladie(s)	60
Une partie des patients du centre de référence concernés par cette ou ces maladie(s)	67



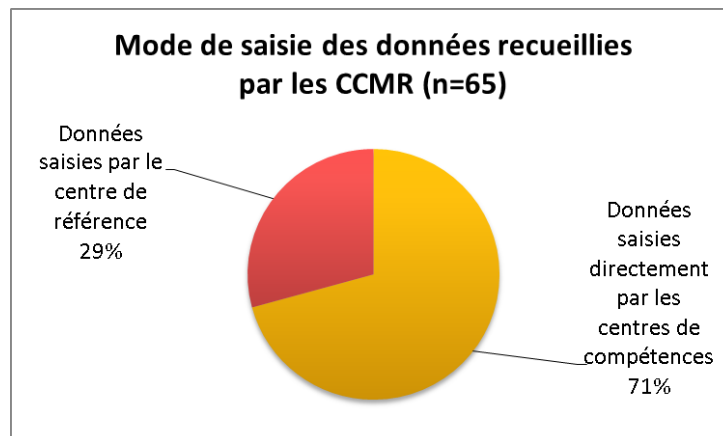
Le fait de recueillir des informations sur une seule maladie ne semble pas influencer positivement sur l'exhaustivité du recueil auprès de tous les patients concernés dans le centre de référence.



Les centres de compétences participeraient au recueil de données dans 57% des cas. Cependant, ce n'est que dans 13% des cas que tous les centres de compétences associés au centre de référence y participent.

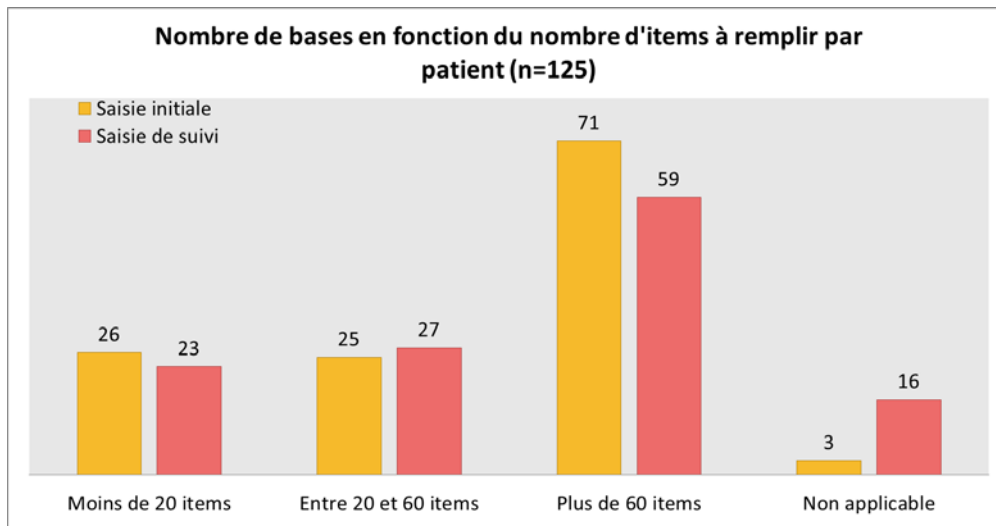


Lorsque les centres de compétences participent au recueil des données, ils saisissent également leurs données dans 71% des cas.

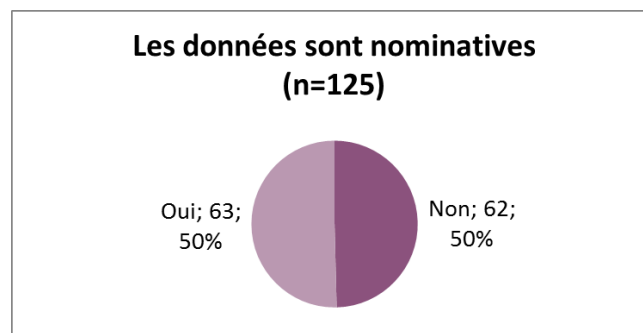


5.2. Les données

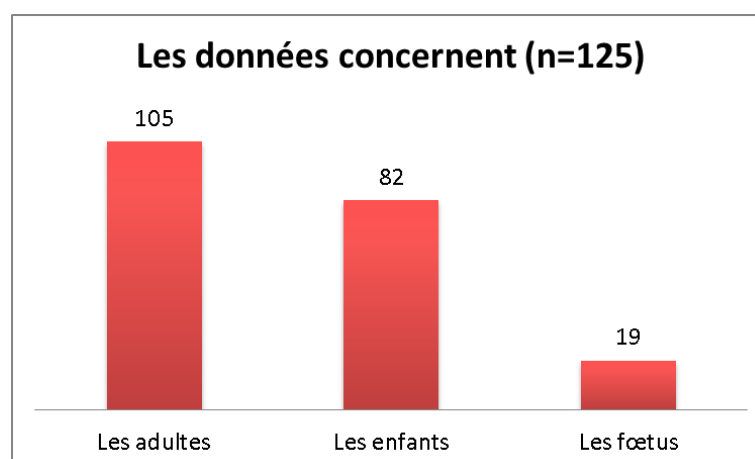
Les bases de données sont en majorité complexes (plus de 60 items). Les personnes ayant répondu « non applicable » pour la saisie initiale ont probablement utilisé cette modalité pour signifier qu'elles n'avaient pas la réponse à cette question. En revanche, la notion de non applicabilité est pertinente lorsqu'il n'existe pas de suivi dans la base de données. Les patients y sont alors entrés une seule fois et les professionnels n'assurent pas de suivi sur le dossier.



Les données collectées sont autant nominatives qu'anonymes.

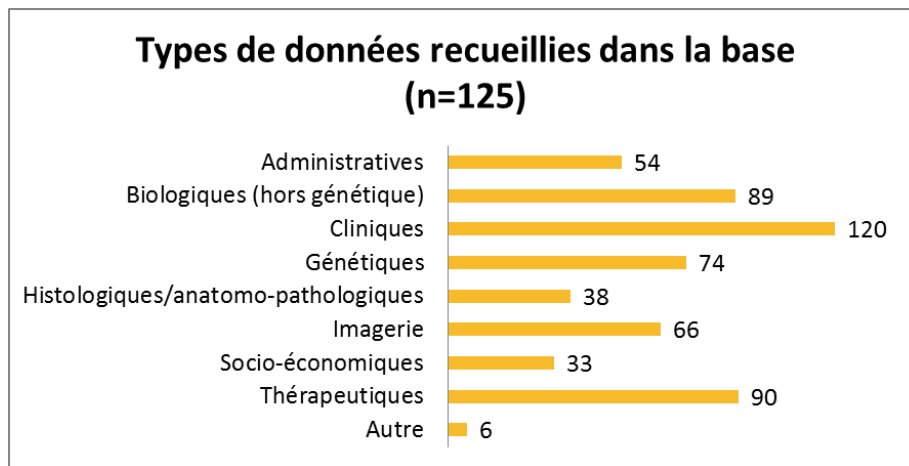


Les bases de données concernent en majorité les adultes (84%) (cette question était à choix multiple). Les données sur les fœtus sont collectées dans 15% d'entre elles.



Les bases de données recueillent des types de données hétérogènes mais concernent généralement la clinique (95%), les traitements (72%) et la biologie (71%). Il est plus rare de

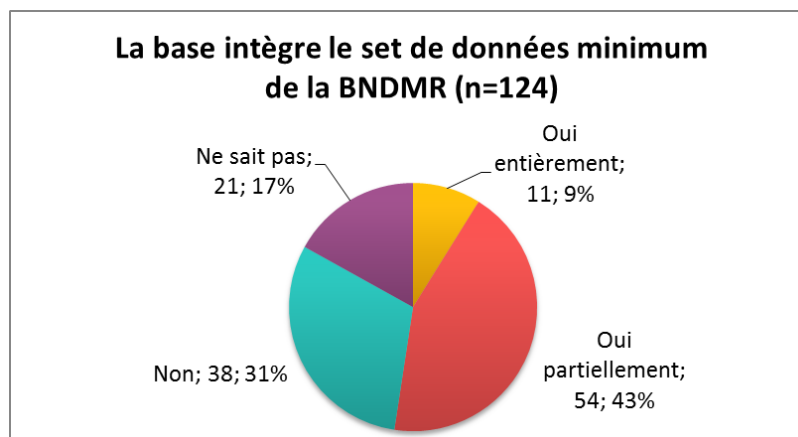
collecter des données socio-économiques (26%) ou histologiques (30%) (cette question était à choix multiple).



Les types de données décrits dans la catégorie « Autre » incluent des données démographiques, de diagnostic, d'examen clinique (ECG/test pharmacologiques), familiales, de Localisation/Origine ethnique, et de suivi.

5.3. Interopérabilité

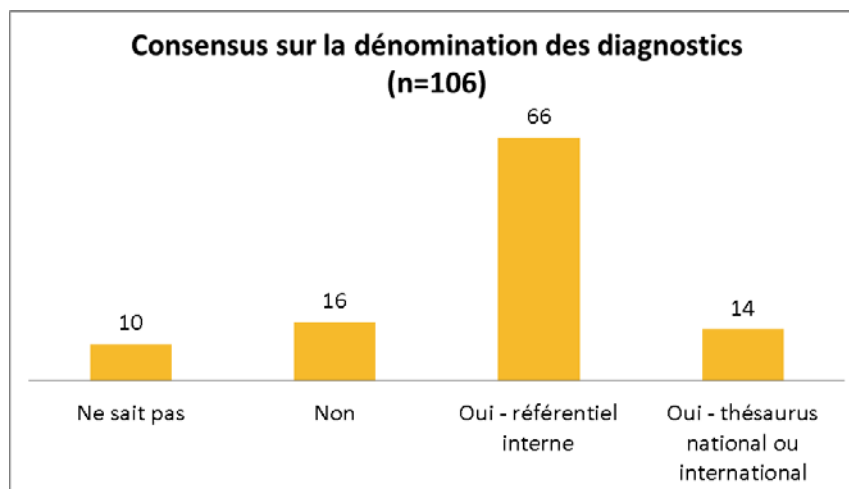
54% des répondants déclarent que la base intègre au moins une partie du set de données minimum de la BNDMR, en revanche seuls 9% l'auraient entièrement intégré. 31% pensent ne pas du tout retrouver les items du set de données minimum maladies rares dans la base de données.



La base intègre le SDM de la BNDMR (nombre de réponse par filière)				
	Oui entièrement	Oui partiellement	Non	
ANDDIRARE		0	2	5
BRAIN TEAM		1	9	1
CARDIOGEN		0	3	0
DEFISCIENCE		2	0	1
FAIR		3	0	15

La base intègre le SDM de la BNDMR (nombre de réponse par filière)			
	Oui entièrement	Oui partiellement	Non
FAVA MULTI	1	2	0
FILFOIE	0	2	1
FILNEMUS	0	0	0
FILSLAN	0	1	0
FIMARAD	0	1	0
FIMATHO	0	2	0
FIRENDO	2	6	1
G2M	1	2	4
Hors filière	0	2	3
MARIH	0	0	0
MCGRE	0	8	4
MHEMO	1	3	0
MUCO	0	0	1
NeuroSphinx	0	1	0
ORKID	0	0	2
OSCAR	0	2	0
RESPIFIL	0	1	0
SENSGEN	0	3	0
TETE COU	0	4	0
Total	11	54	38

Une majorité des bases de données utilisent un référentiel interne pour nommer les diagnostics maladies rares (62%).



Seuls 13% utilisent un thésaurus national ou international, répartis ainsi :

Modalité	Nombre de réponse
OMIM	2
Orphanet	3
Orphanet et/ou CIM10	4
thésaurus de spécialité internationale	3
CIM10	1
DSM-5	1

Consensus sur la dénomination des diagnostics (nombre de réponse par filière)				
	Oui – thésaurus national ou international	Oui - référentiel interne	Non	Ne sait pas
ANDDIRARE	1	2	0	0
BRAIN TEAM	3	6	1	3
CARDIOGEN	1	2	0	0
DEFISCIENCE	1	2	1	0
FAIIR	0	25	0	0
FAVA MULTI	0	1	1	1
FILFOIE	0	1	0	0
FILNEMUS	0	0	0	0
FILSLAN	1	0	0	0
FIMARAD	0	0	0	3
FIMATHO	1	0	0	1
FIRENDO	4	7	1	0
G2M	0	4	2	0
Hors filière	0	2	1	0
MARIH	0	0	0	0
MCGRE	0	6	4	1
MHEMO	1	2	0	1
MUCO	0	0	1	0
NeuroSphinx	0	0	1	0
ORKID	0	0	2	0
OSCAR	0	2	1	0
RESPIFIL	1	0	0	0
SENSGEN	0	2	0	0
TETE COU	0	2	0	0
Total	14	66	16	10

Les bases de données déclarées sont très hétérogènes. Elles concernent autant une seule maladie que plusieurs maladies. Le recueil n'est pas toujours exhaustif auprès des patients des centres de référence concernés par la maladie. Il est cependant majoritairement exhaustif en terme de sites du centre de référence collectant l'information. Les centres de compétences participent eux-aussi au recueil de données dans 57% des cas.

Les données collectées sont autant nominatives qu'anonymes et concernent en majorité les adultes (84%).

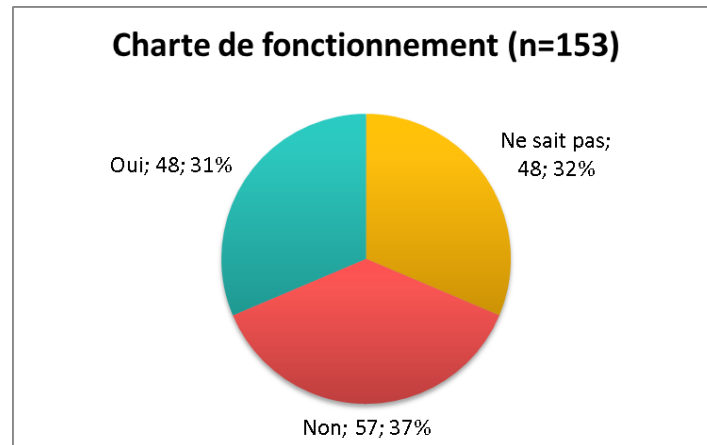
Les bases de données recueillent des types de données hétérogènes mais concernent généralement la clinique (95%), les traitements (72%) et la biologie (71%). Elles sont en majorité complexes (plus de 60 items).

63% des bases intègrent au moins une partie du set de données minimum de la BNDMR. Seules 13% utilisent un thésaurus national ou international pour la dénomination des diagnostics.

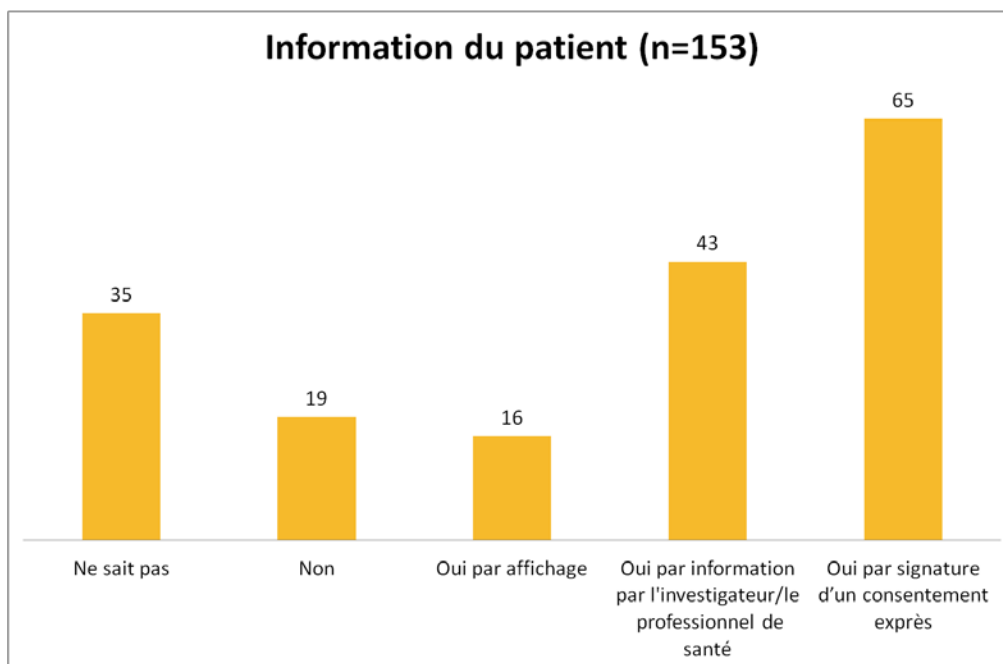
6. Répondent-elles aux exigences éthico-juridiques ?

Une démarche (déclaration ou autorisation) auprès de la CNIL a été renseignée dans 68% des cas.

Un tiers des bases de données déclarées possèdent une charte de fonctionnement. Cependant, un tiers des répondants n'ont pas pu nous informer sur ce point.



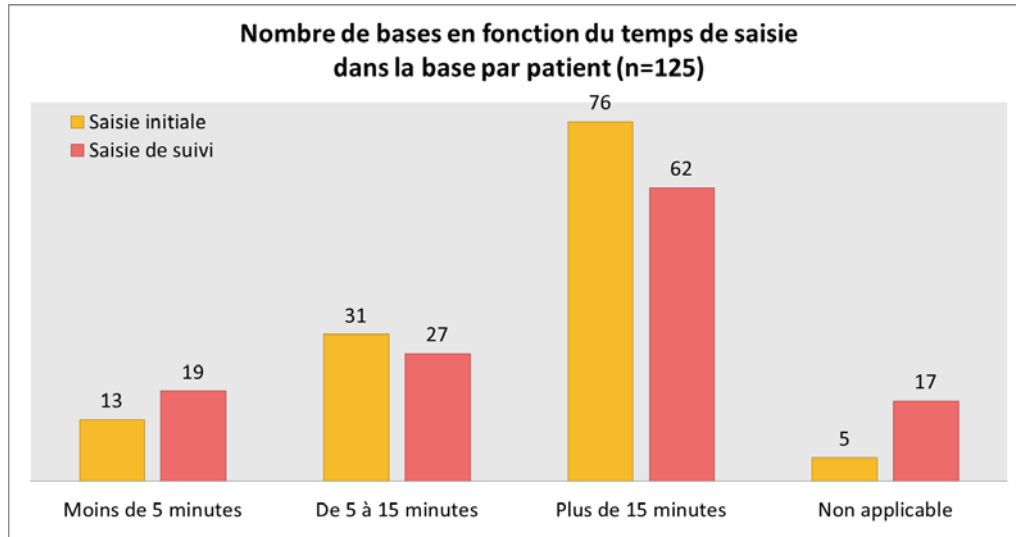
La majorité des bases requièrent un consentement signé du patient (36%) (cette question était à choix multiple). 20% des personnes ayant saisi l'information ne savaient pas si les patient étaient informés du recueil de leurs données.



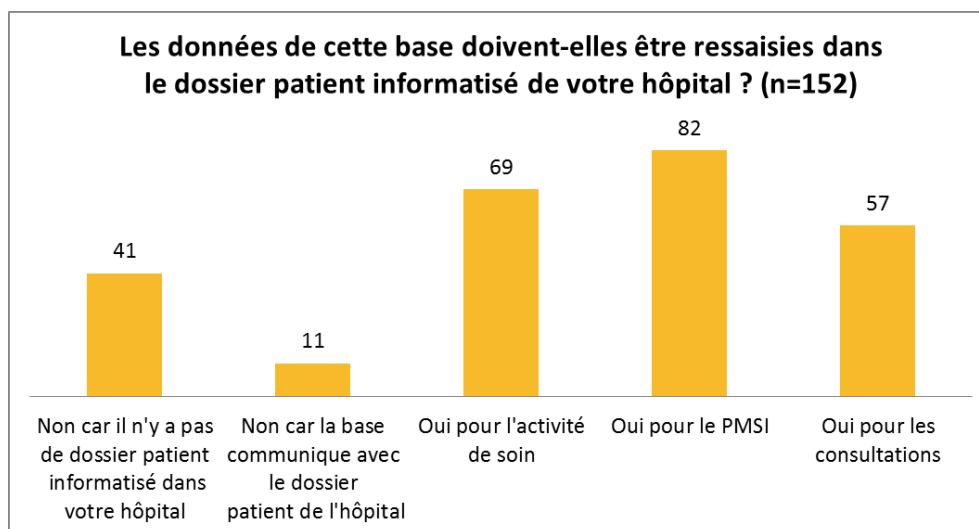
Les personnes ayant complété cette enquête ne semblaient pas avoir la maîtrise des aspects éthiques et réglementaires. Globalement, les démarches CNIL ont été entamées pour une majorité de bases de données. Il n'existe pas systématiquement de charte de fonctionnement pour les utilisateurs des bases de données. L'information des patients se fait majoritairement par signature de consentement.

7. Quelle est la charge de travail qui en découle ?

Par patient, le temps de saisie est majoritairement supérieur à 15 minutes que ce soit en saisie initiale (création de la fiche patient) ou de suivi (patient déjà en base).

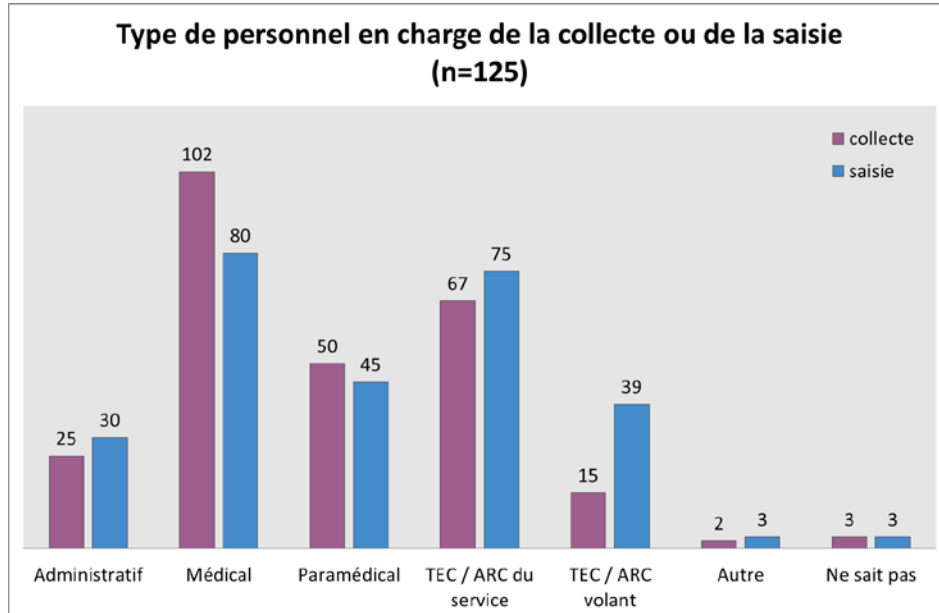


A nouveau, les personnes ayant répondu « non applicable » pour la saisie initiale ont probablement utilisé cette modalité pour signifier qu'elles n'avaient pas la réponse à cette question. En revanche, la notion de non applicabilité est pertinente lorsqu'il n'existe pas de suivi dans la base de données. Les patients y sont alors entrés une seule fois et les professionnels n'assurent pas de suivi sur le dossier.



Dans 27% des réponses, la ressaisie n'est pas nécessaire du fait de l'absence de dossier patient informatisé (DPI) dans l'hôpital (cette question était à choix multiple). Seuls 7% déclarent ne pas avoir à ressaisir car la base communique avec leur DPI. Dans tous les autres cas, bien qu'un DPI existe, les données doivent être ressaisies.

Le personnel en charge de la collecte et de la saisie d'information est hétérogène. Les médecins restent principalement en charge de la collecte mais aussi de la saisie des données (cette question était à choix multiple). Les TEC ou ARC des services sont également fortement sollicités pour la saisie ainsi que pour la collecte des informations.



La catégorie « Autre » regroupe pour la collecte un chercheur et un ingénieur de recherche, et pour la saisie un chercheur, un ingénieur de recherche et CLININFO.

- **Estimation de la charge de travail au niveau national**

La charge de travail qui découle de la saisie des données dans les bases de données peut être estimée en mettant en regard les valeurs suivantes :

- ↪ Le nombre de bases de données (254) et le nombre de patients dans ces bases (743 000)
- ↪ La file active (190 000 patients)
- ↪ Le temps de saisie estimé :

	CEMARA		Bases déclarées (hors CEMARA)	
	Temps de saisie par patient (minutes)	Patients par heure	Temps de saisie par patient (minutes)	Patients par heure
Saisie initiale	12	5,0	27	2,2
Saisie de suivi	5	11,3	25	2,4

Cette estimation est basée sur le déclaratif des personnes ayant répondu à l'enquête. Le temps de saisie, qu'il soit initial ou de suivi, était indiqué en fourchette de temps : moins de 5 minutes, entre 5 et 15 minutes, et plus de 15 minutes. Pour tous les calculs de temps, nous avons considéré le temps minimum et le temps maximum de la fourchette de temps (60 minutes pour « plus de 15 minutes »), puis nous en avons fait la moyenne. C'est ce chiffre qui est reporté dans le tableau ci-dessus.

Pour estimer de la charge de travail sur une année, nous proposons la simplification suivante :

La file active de CEMARA en 2014 était de 87000 patients : 46 000 nouveaux patients et 41 000 patients revus. Pour les nouveaux patients, nous appliquons le temps moyen de saisie initiale d'un dossier dans CEMARA ; pour les patients revus nous appliquons le temps de saisie de suivi dans CEMARA.

A cela s'ajoute la file active des bases de données déclarées dans l'enquête. Ces patients étant déjà entrés dans les bases de données, les prochaines saisies seront donc des suivis ; nous appliquons le temps de saisie de suivi dans les bases de données hors CEMARA.

Pour les bases de données identifiées par alignement avec les listes Orphanet et Aviesan, ainsi que pour les bases de données « néo-déclarée », il n'est pas possible de connaître la part de saisie initiale ou de suivi ; nous appliquons alors la moyenne du temps de saisie par dossier (soit 2.3 dossiers / heure).

Ces résultats sont alors rapportés en mois/homme. Nous considérons l'efficacité d'un professionnel sur une journée à 80%. Nous obtenons un total de 459 mois-homme pour une saisie unique (initiale ou de suivi) dans l'une des bases de données identifiées dans cette enquête.

D'autre part, dans CEMARA, nous avons 1,67 activités par dossier en moyenne en 2014. Bien que toutes les bases de données ne fassent pas de suivi de leurs patients, nous estimons au vu des résultats de cette enquête que nous pouvons néanmoins appliquer ce chiffre à l'ensemble des résultats. Au total, nous estimons donc la charge de travail annuelle pour la saisie dans ces différentes bases de données à **766 mois-homme**. Ce temps est **uniquement pour la saisie**. La recherche d'informations (retour au dossier etc.), le contrôle qualité des données et toute activité connexe n'y est pas compté. Nous comptons ici uniquement le travail au niveau des centres de référence, les bases des centres de compétence et leur travail n'étant pas l'objet de ce rapport.

De plus, les calculs sont effectués sur le nombre de patients et non le nombre de bases de données. Le temps de saisie est par essence multiplié du fait de la multiplicité des bases et des personnes impliquées dans la saisie. La mise en commun de ressources, dans une optique de professionnalisation et d'industrialisation de la saisie, pourrait être un facteur d'efficacité.

Nous ne prenons pas non plus en compte les doublons au sein des bases de données. Or, une saisie initiale est généralement plus chronophage qu'un simple suivi.

Nous ne comptons pas ici le temps de saisie dans les cohortes RaDiCo.

Sur le même principe de calcul, nous évaluons que la saisie des dossiers déjà présents dans les bases de données a représenté environ 5561 mois / homme, pour en moyenne 2,53 activités par dossier (moyenne CEMARA).

La saisie des patients dans les bases de données est chronophage (majoritairement >15 minutes) et requiert par conséquent d'allouer du temps de personnel médical, paramédical, administratif ou de professionnels dédiés (TEC / ARC). L'absence d'interopérabilité des dossiers patients hospitaliers avec les bases de données ajoute au fardeau en multipliant les saisies.

CEMARA

La base de données CEMARA a été ouverte en mai 2007 afin de permettre aux professionnels des centres maladies rares de recenser leurs cas de patients maladies rares dans un objectif de Soins, de Recherche et de Santé publique.

A l'été 2015, CEMARA était utilisée par 61 des 131 centres de référence français, répartis dans 15 filières de santé.

CEMARA se compose d'un tronc commun et peut être complétée de recueils de données complémentaires appelés pétales. 12 pétales sont encore actifs à ce jour et concernent 9 filières différentes. Les pétales peuvent être des recueils de données communs à différentes filières.

Nous constatons que 3 filières n'ont pas déclaré saisir des données dans la base CEMARA bien que certains des professionnels de ces filières possèdent des comptes. La filière MHEMO a abandonné la saisie dans CEMARA suite au développement d'outils propres, davantage spécifiques aux pathologies de la filière. La filière FILNEMUS avait fait une enquête propre, qui portait exclusivement sur le mode de fonctionnement des bases spécifiques aux maladies neuromusculaires et non pas sur les bases remplies par chaque centre de référence, ce qui fait que CEMARA n'y apparaissait pas. La filière FILSLAN, probablement dans la même logique, n'a déclaré que la base de données CLEANWEB spécifique à la pathologie et complétée par tous les centres de la filière. Enfin, les centres hors filière n'ont pas non plus déclaré participer à CEMARA bien que plusieurs l'utilisent.

	Nombre total de sites CRMR dans la filière	Nombre de sites ayant CEMARA	% de sites ayant CEMARA	Pétales	Réponses CEMARA
ANDDIRARE	30	30	100%	1	4
BRAIN TEAM	17	0	0%	0	
CARDIOGEN	8	0	0%	0	
DEFISCIENCE	13	10	77%	0	5
FAIR	15	3	20%	0	
FAVA MULTI	4	0	0%	0	
FILFOIE	3	0	0%	0	
FILNEMUS	36	26	72%	0	
FILSLAN	3	2	67%	0	
FIMARAD	20	5	25%	0	1
FIMATHO	7	2	29%	0	1
FIRENDO	18	15	83%	3	9
G2M	17	11	65%	0	6
Hors filière	59	13	22%	1	
MARIH	27	0	0%	0	
MCGRE	14	0	0%	0	
MHEMO	18	11	61%	1*	*
MUCO	4	0	0%	0	

	Nombre total de sites CRMR dans la filière	Nombre de sites ayant CEMARA	% de sites ayant CEMARA	Pétales	Réponses CEMARA
NeuroSphinx	8	3	38%	1	1
ORKID	18	16	89%	3	6
OSCAR	13	5	38%	1	5
RESPIFIL	8	0	0%	0	
SENSGEN	13	8	62%	1	3
TETE COU	8	4	50%	1	3
Total	381	164		14	44

* Non rempli

- **Gestion**

Une gouvernance des données avait été mise en place à la création de CEMARA et ce comité scientifique se composait principalement des investigateurs. La base de données bénéficie d'un financement pérenne pour l'hébergement, le contrôle qualité et data management, l'exploitation et la valorisation des données. Seule la saisie reste entièrement à la charge des utilisateurs.

- **Caractéristiques techniques**

CEMARA est une application dédiée, accessible par Internet. La partie pétales a été développée par un prestataire externe. La base est hébergée à l'hôpital Necker-Enfants malades (Paris, France). CEMARA ne communique pas avec d'autres bases de données. C'est notamment pour cette raison que CEMARA va être prochainement remplacée par BaMaRa, base de nouvelle génération, qui permettra cette interopérabilité.

- **Ethico-juridique**

CEMARA a reçu une autorisation de la CNIL en 2007.

Une charte de fonctionnement est signée par chaque utilisateur afin que leur soit attribué un compte.

Les patients sont informés du recueil de leurs données par affichage ou par information par le professionnel de santé.

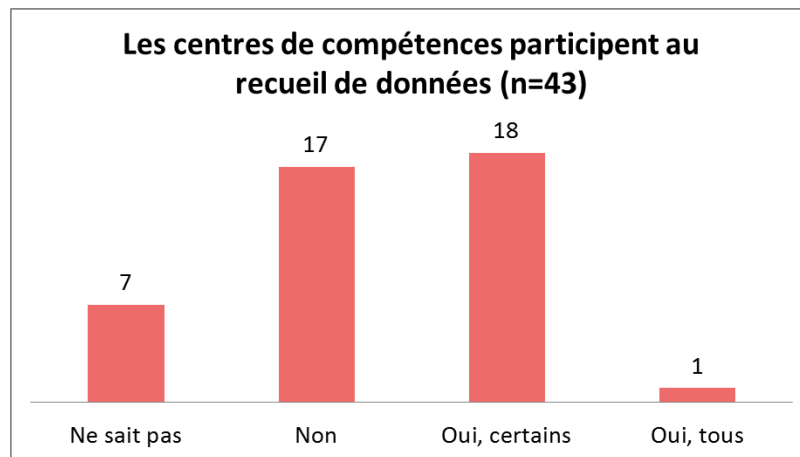
- **Les données**

Les résultats suivants se basent sur les réponses à l'enquête (n=43).

Par rapport aux autres bases de données recensées, peu d'utilisateurs de CEMARA recueillent des données pour une seule maladie (seuls 2 réponses soit moins de 5%). Le recueil est plus souvent effectué pour tous les patients du centre de référence concernés par la ou les maladies (79% pour CEMARA vs 43% pour les autres bases).

Le recueil est également plus souvent effectué dans tous les sites du centre de référence (72% pour CEMARA vs 57% pour les autres bases). Les centres de compétences participent

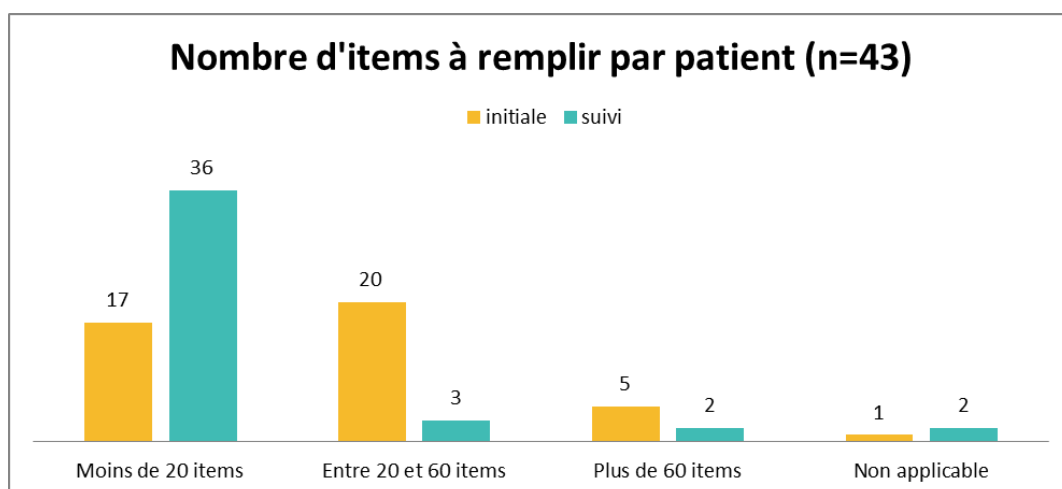
moins souvent au recueil (44%) et saisissent eux-mêmes les données dans 58% des cas contre 71% dans les autres bases.



Dans CEMARA, les données sont collectées de manière nominative. Elles concernent aussi bien les adultes que les enfants mais également plus de 17000 fœtus.

Les données recueillies dans le tronc commun sont principalement des données administratives, cliniques et génétiques. Le tronc commun intègre le Set de Données Minimum National Maladies Rares (SDM) de la Banque nationale de données maladies rares (BNDMR). Quelques répondants ne savaient pas que c'était le cas (8 réponses « ne sait pas » et 1 « non »). Par ailleurs, les diagnostics sont codés en Orpha (thésaurus issu d'Orphanet version 2009).

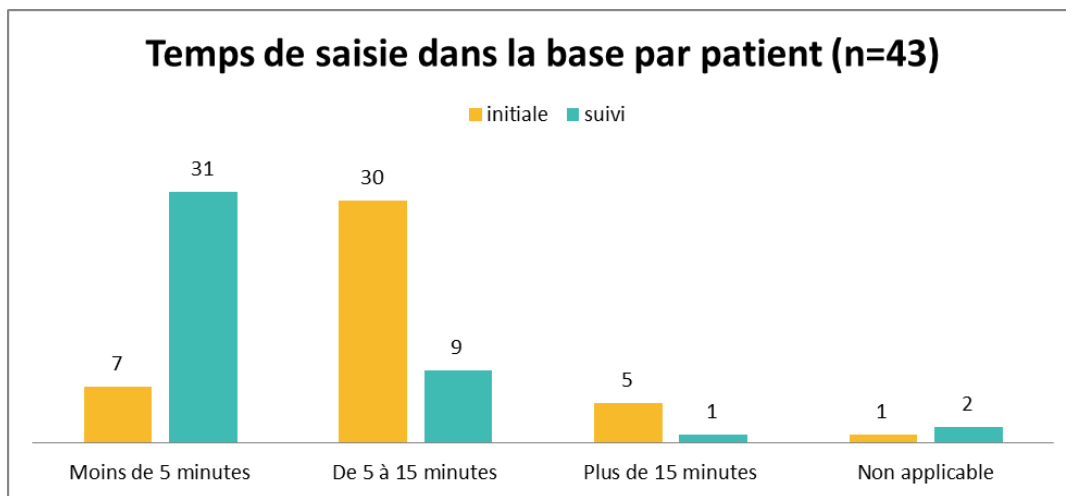
Le tronc commun de CEMARA compte une cinquantaine d'items, mais tous ne sont pas obligatoires.



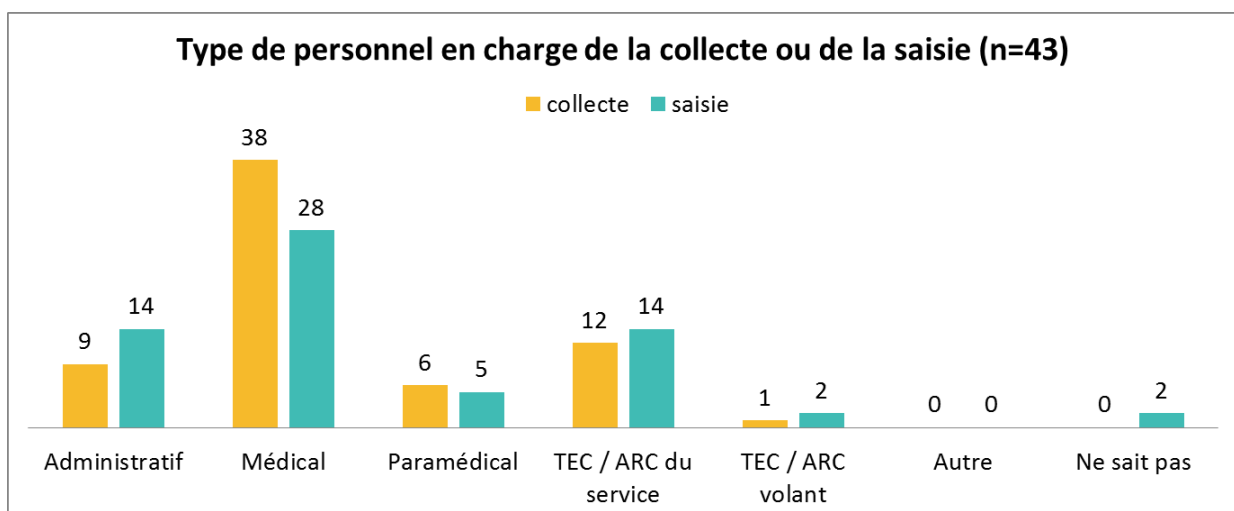
La personne ayant répondu « non applicable » pour la saisie initiale a probablement utilisé cette modalité pour signifier qu'elles n'avaient pas la réponse à cette question. En revanche, la notion de non applicabilité est pertinente lorsqu'il n'existe pas de suivi dans la base de

données. Bien que CEMARA ait été conçu pour suivre l'activité réalisée dans les sites de prise en charge, certains centres ne saisissent les patients que lors de la première activité et ne reviennent jamais sur le dossier pour indiquer les suivantes. C'est pourquoi le nombre d'activité médian par dossier dans CEMARA est de 1 et le nombre d'activité par dossier varie de 0 à 349. Le nombre d'activité moyen dans CEMARA est de 2,53.

Le temps de saisie moyen est de 12 minutes pour la saisie initiale et de 5 minutes pour le suivi. Les temps de saisie de plus de 15 minutes correspondent très probablement aux pétales spécifiques.



Enfin, le type de professionnels en charge de la collecte ou de la saisie des données dans CEMARA sont principalement des médecins. On note que la proportion de TEC / ARC et de paramédicaux est inférieure aux autres bases de données.



Discussion

Cette enquête présente des résultats hétérogènes entre les filières de santé, reflétant à la fois la complexité du paysage des bases de données en France mais aussi les biais liés à la méthode de recueil de données pour cette enquête.

Cette enquête a été proposée aux filières comme un outil pour les accompagner dans leur démarche d'état des lieux nécessaire à la structuration de leur plan d'action. Elle n'était pas obligatoire, c'est pourquoi 3 filières n'ont pas directement participé (FIMARAD, FILNEMUS et MARIH). Ces filières avaient réalisé leur propre état des lieux, souvent moins détaillé que l'enquête nationale. Elles nous ont transmis leurs résultats et ces bases de données ont été reportées dans le questionnaire par l'équipe de la BNDMR, avec les informations disponibles. Cela sous-entend que les résultats présentés ne reflètent que très peu les bases de données de ces 3 filières pour lesquelles de nombreux items n'ont pas pu être complétés.

Par ailleurs, deux modes de collecte distincts ont cohabité dans cette enquête. Dans 8 filières, le recueil et la saisie des données ont été effectués par le chef de projet ou le chargé de mission de la filière (ou l'ARC), suite à des entretiens avec les responsables de centres de référence. Dans les autres cas, le questionnaire a été géré par les responsables de centres de référence eux-mêmes. Un accompagnement en direct par téléphone leur était parfois proposé par certains chefs de projet, afin de s'assurer de l'harmonisation des réponses au sein de la filière.

Certaines filières ont eu une approche par centre de référence et d'autres par la filière. Ainsi, certaines filières ont déclaré plusieurs fois la même base de données puisqu'elle était utilisée par plusieurs centres (avec parfois des différences de réponses en fonction du centre concerné, par exemple en terme de personnel qui saisit, temps de saisie, nombre de patients...). Cette approche était informative pour la filière, mais un travail de dédoublonnage a été nécessaire pour l'analyse nationale. Dans la démarche d'approche globale au sein de la filière, les chefs de projet ont joué un rôle essentiel d'homogénéisation et de dédoublonnage le cas échéant. Dans les deux cas, les chefs de projet avaient la possibilité de suivre les réponses dans l'outil mis à leur disposition (Zoho Survey) afin de les compléter ou les corriger tout au long de l'enquête.

Pour la filière MUCO, le fait de ne déclarer qu'une seule base de données, le registre national, a été un choix délibéré. Il nous a été indiqué qu'il existait d'autres bases, non déclarées ici.

Dans le cas des 23 centres hors filière, les médecins des centres de référence ont été directement sollicités. Nous n'avons obtenu que 5 réponses après 2 relances. Certaines filières nous ont informé avoir également eu des difficultés à obtenir des retours de la part des CRMR. La filière avait parfois connaissance de l'existence d'autres bases de données mais ne parvenait pas à recueillir davantage d'information auprès des coordonnateurs des centres. Cela a été un facteur de sous-déclaration. Néanmoins, la dynamique des filières a favorisé la participation des centres de référence à l'enquête nationale.

Les centres de référence ont répondu à l'enquête pour leurs principales bases de données. Le questionnaire étant relativement long et complexe, le temps nécessaire pour le compléter a pu être un frein au recueil exhaustif de toutes les bases de données utilisées. Un autre biais repose sur le fait que les centres de référence répondaient au nom de tous leurs sites constitutifs. Or ils ne connaissent pas forcément toutes les bases développées et utilisées, notamment les petites bases locales. Ainsi, les bases (Excel) personnelles ont probablement été sous-déclarées. Enfin, certaines bases ont pu être omises du fait de leur absence de déclaration à la CNIL. Enfin, lorsqu'une base française et une base européenne coexistaient, il ne nous a été déclaré que l'une des deux.

Les filières ne recouvrent pas exactement le périmètre des CRM et cela a pu être un obstacle à la déclaration exhaustive de toutes les bases de données. L'enquête s'adressait aux centres de référence, mais elle a été portée par les filières. Il serait donc intéressant de savoir si les bases de données n'entrant pas dans le champ de la filière ont malgré tout été déclarées dans l'enquête.

Le travail d'alignement des résultats de l'enquête avec les listes d'Orphanet et d'Aviesan reste partiellement incomplet. En effet, il serait intéressant d'investiguer davantage sur les 42 bases qui n'ont pu être attribuées à aucune filière afin de déterminer si ces bases sont obsolètes ou si elles sont rattachées à des centres de référence (inclus ou non dans les filières). Le fait d'avoir travaillé avec les chefs de projet des filières et non avec les coordonnateurs de centres de référence pour la réattribution de ces bases a pu être un biais.

De plus, certains professionnels sont rattachés à plusieurs centres de référence qui font partie de différentes filières. Ces professionnels ont pu être contactés plusieurs fois, au titre des différentes filières, et cela a pu porter à confusion. Nous avons de ce fait réaffecté 2 réponses qui n'avaient pas été saisies pour la bonne filière.

Une seule base de données recensée dans l'enquête concerne deux filières. Il est possible que, dans le cas où une base de données concernait plusieurs filières, elle n'ait été déclarée qu'une seule fois. Le nombre de bases par filière serait donc à nouveau sous-estimé.

Enfin, les résultats de cette enquête et notamment l'estimation de la charge de travail qui découle des bases de données devrait aussi s'appuyer sur les résultats de l'enquête PIRAMIG.

Conclusion

Cette enquête permet de mieux décrire le paysage des bases de données élaborées par les professionnels des centres de référence maladies rares dans le cadre des missions qui leurs ont été confiées. De nombreuses bases de données, aux finalités différentes, ont été créées. Une grande disparité est observée entre les filières.

Il a été impossible d'obtenir un paysage exhaustif du fait du caractère facultatif de cette enquête. Nous avons tenté de redresser les résultats afin d'estimer de la manière la plus fine possible le nombre de bases de données dans les centres de référence en France. Les 254 bases de données identifiées dans les centres de référence pourront être complétées par les données recueillies auprès des centres de compétences dans l'enquête en cours, mais elle aussi facultative, qui devrait se terminer mi-décembre 2015.

Des stratégies nationales pourront ainsi être établies par les filières de santé maladies rares pour les prochaines années afin de mieux couvrir les besoins : création de nouvelles bases spécifiques, adaptation de bases existantes. Cependant, le grand nombre de bases identifiées souligne l'importance de simplifier le recueil des données et de limiter les saisies multiples en privilégiant l'interopérabilité. Des outils d'interopérabilité sont proposés par la BNDMR mais ne couvrent pas tout le champ des besoins recensés. Il serait intéressant de mutualiser les bases et les ressources entre centres, sites ou hôpitaux de manière à factoriser les coûts de tels recueils.

En effet, la multiplication des bases de données est un frein à leur mise en œuvre effective en terme de collecte de l'information, de contrôle qualité des données et d'exploitation. Leur maintenance dans le temps est également difficile et coûteuse.

Enfin, le grand nombre de maladies ainsi que la spécialisation des réseaux rendent difficile leur intégration dans les systèmes d'information hospitaliers. Cela allègerait pourtant la charge de travail des équipes en limitant les saisies multiples et améliorerait l'identification des patients (enregistrement centralisé à l'hôpital).

Annexe 1 : liste des bases de données déclarées dans l'enquête

Liste des bases de données déclarées (données brutes non corrigées)

1. 4D
2. Acces
3. Access
4. Access Epidermolyses bulleuses héréditaires
5. Access
6. Adherence
7. Adjuvite
8. AIR (Auto-Immunité et Rituximab)
9. Aplasie
10. ASSESS
11. Autre base privée pour une autre maladie
12. Banque de données Gilles de la Tourette
13. Base Access devenue base Excel
14. Base administrative
15. Base de données ACCESS - HDJ UMGR
16. Base de données Atrophie Multisystématisée
17. Base de données de l'Unité transversale de la drépanocytose
18. Base de données du CRMW
19. Base de données du PHRC ECOPHEN
20. Base de données Excel
21. Base de données Excel Perso
22. Base de Données Internationale pour les Dysferlinopathies
23. Base de données interne
24. Base de données MFM
25. Base de données pédiatriques GR-Ex "PEDIA DREPA"
26. Base Excel
27. Base Excel des données initiales des nouveaux patients suivi dans le CR (Bilan biologique de base de la 1ère consultation dans le CR)
28. Base maison
29. Base nationale Alzheimer (BNA)
30. Base privée
31. Base Sclérodermie
32. Base spina
33. B-Confident
34. Biomarqueurs (RHLF)
35. Blackfan diamond
36. BRAINDYS
37. CADASIL - Middlec@re
38. CARREG
39. CARTAGENIA
40. Chilrane
41. CIROCO
42. CleanWeb
43. Cohorte Moyamoya - Middle care

44. CRANIOFACIAL
45. Cyclolup
46. Cyto AI adultes
47. D[4]/Phenodent
48. Database CERA
49. DEFGEN
50. DEFI
51. DM-Scope
52. DSD-Life
53. DUO
54. Dystonies
55. eDbai
56. EFACTS
57. Encephalites du gris et autres
58. Epidémiologie clinique des myopathies inflammatoires idiopathiques : Myosites Auto-anticorps Spécifiques Classification (MASC)
59. ERCUSYN (European register on cushing's syndrom)
60. e-RespiRare
61. ETAP
62. Etaplus
63. EUHASS (European Haemophilia Safety Surveillance)
64. EUROMGDB
65. Evacarney
66. Evans enfant
67. Excel
68. Excel
69. Excel
70. Excel base donnée locale
71. Excel Dysplasie ectodermique
72. Excel Epidermolyses Bulleuses héréditaires
73. Excel Ichtyoses
74. Excel Incontinentia Pigmenti
75. Excel Syndrome de Netherton
76. Excel XP
77. Excell
78. Excell personnel
79. Ferti-preserved
80. Fichier excel - adultes drépanocytaires
81. Fichier Patients
82. FilemakerPro
83. FranceCoag
84. Genecoeur, Integralis
85. Gensource
86. GR2
87. HC-Forum (Technidata)
88. Histo
89. HNF1B
90. Holoprosencéphalie
91. Hyperplasie congénitale des surrénales
92. Hypophosphatasie - D[4]/Phenodent
93. Hypospades
94. JIR cohorte

95. KIDBIOSEP
96. LEUKOFRANCE / LEUKOTREAT Data Base
97. Listing patient DSD Paris, fichier excel
98. Listing patients DSD CR Lyon
99. Logiciel CleanWeb
100. LogiFer
101. Lotus Approach IBM
102. Lupsenic
103. MAINRITSAN2
104. Malformations des membres
105. MAOLYA 4D
106. MAT
107. MG
108. Microcéphalies primaires héréditaires
109. Microsoft Access
110. MIVB
111. Neuropathies Périphériques
112. Neutro LGL
113. Neutropénie
114. NEXT iMEDIDATA
115. NF-France
116. NF-Vancouver
117. NHEMO (Net-HEMOstase)
118. Nordinet
119. Observatoire des patients atteints de laminopathies and emerinoopathies (OPALE)
120. Observatoire français des patients atteints de glycogenose de type III
121. ONF DMFSH
122. OpenCrf
123. OSCAR - BASE DE DONNEES
124. Patients CDR
125. PCDREP (Access)
126. Pédopsybase
127. Pexivas
128. PGRx lupus
129. Pituigene
130. Predex
131. RASopathie
132. RECOVER
133. REDCap - Etude GenMod SS
134. REGISTRE CEREDIH
135. Registre de l'atrésie de l'oesophage
136. Registre des maladies mitochondriales
137. Registre européen de l'acromégalie LAS
138. REGISTRE FRANCAIS DE LA MUCOVISCIDOSE
139. Registre national de "Cushing"
140. Registre National de la Maladie de Pompe
141. Registre National des Hernies Diaphragmatiques
142. REGISTRE NATIONAL DES THALASSEMIES/BASE ACCESS
143. Registre National Français de la Maladie de Gaucher
144. Registry
145. RESOCANAUX
146. RHEOP

147. RTU de l'infliximab dans la maladie de Takayasu
148. SABLE
149. Sclerocyc
150. Scleromics
151. SDM-VOOZANOO / NEODREP
152. SED-VASC ; DFM
153. SICLOPEDIE - Dossier Patient partagé informatisé
154. SPATAX
155. Spirale
156. Statvas
157. Stomatocytoses héréditaires
158. Thromboses veineuses cérébrales
159. UMD-DMD
160. UMD-EMD
161. UMD-LMNA
162. UMD-SMN1
163. VALDIG
164. Vedoss
165. WIN LUPUS

Annexe 2 : données brutes par filière hors CEMARA

Question	modalité	ANDDIRARE	BRAIN TEAM	CARDIOGEN	DEFISCIENCE	FAIIR	FAVA MULTI	FILFOIE	FILNEMUS	FILSLAN	FIMARAD	FIMATHO	FIRENDO	G2M	Hors filière	MARIH	MCGRE	MHEMO	MUCO	NeuroSphinx	ORKID	OSCAR	RESPIFIL	SENSGEN	TETE COU
Nombre total de patients enregistrés dans la base dans son ensemble :	Min	500	70	2224	50	5	150	793	6	12000	800	913	99	408	1743	237	60	1470	9000	371	227	15	4859	2000	10
	Max	45000	1,00E+05	43526	450	1530	9585	1155	5000	12000	3000	1584	2800	408	5700	2051	3708	31529	9000	371	3000	1500	4859	7838	6000
	Median	3474	702	19561	328	90	2958	974	421	12000	2000	1248	235	408	3500	820	620	5818	9000	371	1613	184	4859	7030	3039
File active pour 2014 : (Nombre de patients différents vus au cours de l'année 2014)	Min	45	45	604	50	2	50	140		3500		46	28	240	60		11	264	6500	136	125	3	1443	700	10
	Max	1385	5000	6570	50	1400	1509	684		3500		155	1053	592	2500		2180	5006	6500	136	300	300	1443	2325	724
	Median	275	190	1099	50	109	804	412		3500		100	80	416	300		300	1782	6500	136	212	70	1443	1000	400
Date de création de la base	Min	1990	2004	2002	2004		1995	2005	2001	2008	2012	2008	2006	2004	2009	1993	1995	2007			2008	2000	2008	2007	2009
	Max	2005	2011	2015	2013		2010	2005	2015	2008	2012	2008	2013	2005	2012	2011	2015	2007			2008	2008	2008	2009	2014
	Median	2001	2008	2008	2013		2007	2005	2009	2008	2012	2008	2009	2004	2010	2001	2009	2007			2008	2007	2008	2008	2011
Êtes-vous le concepteur de cette base ?	Non	4	8	1	1	25	1	1	4	0	2	1	5	6	3	0	3	2	1	1	0	2	0	1	2
	Non Répondu	0	0	0	1	0	0	0	0	0	0	7	0	0	0	0	0	0	0	0	0	2	0	0	1
	Oui	4	6	2	3	0	3	2	13	1	2	1	7	4	2	9	10	2	0	0	2	3	1	2	2
Cette base est-elle une base européenne ou internationale ?	Non	6	10	3	3	17	3	2	15	1	4	1	9	10	5	8	12	3	1	1	2	4	1	3	2
	Non Répondu	0	0	0	1	0	0	0	0	0	7	0	0	0	0	0	0	0	0	0	0	2	0	0	1
	Oui	2	4	0	1	8	1	1	2	0	0	1	3	0	0	1	1	1	0	0	0	1	0	0	2
Quel est l'objectif principal de la base ?	administrative	0	0	0	administrative	0	0	0	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
	épidémiologie	0	0	0	0	0	0	0	0	0	0	0	2	0	0	0	0	0	0	0	0	0	0	0	0
	Non Répondu	0	0	0	1	0	0	0	12	0	8	0	4	0	4	0	9	0	0	0	0	2	0	0	1
	Recherche	6	14	3	4	25	3	1	5	1	1	2	11	0	4	0	11	2	1	1	2	2	1	3	4
	Santé publique	1	5	0	4	5	2	0	2	1	0	1	1	4	1	0	6	2	1	1	0	0	1	0	2
	Soin	5	5	2	3	20	4	1	3	0	2	0	3	5	3	0	9	2	1	0	0	5	1	1	3
	Suivi d'activité	0	0	0	0	0	0	0	0	0	1	0	0	0	1	0	0	0	0	0	0	0	0	0	0
Uniquement listing de patients avec coordonnées (nom/prénom/date de naissance) et diagnostic.	0	0	0	0	0	0	0	0	0	0	0	0	0	1	0	0	0	0	0	0	0	0	0	0	
La gouvernance de la base est-elle formalisée (comité scientifique/groupe de pilotage/ comité opérationnel...)?	Ne sait pas	1	2	0	0	4	0	0	16	0	4	0	1	8	0	9	1	0	0	0	0	0	0	0	0
	Non	6	10	1	3	0	1	2	0	0	0	0	5	2	2	0	7	2	0	1	1	3	1	2	1
	Non Répondu	0	0	0	1	0	1	0	0	0	7	0	0	0	0	0	0	0	0	0	0	2	0	0	1
Quelles sont les parties représentées dans la gouvernance ? (plusieurs réponses possibles)	Oui	1	2	2	1	21	2	1	1	1	0	2	6	0	3	0	5	2	1	0	1	2	0	1	3
	Autres	1	1	0	0	2	1	0	0	0	0	0	1	0	0	0	2	1	1	0	0	0	0	0	2
	Fondations (droit privé)	0	1	0	1	1	0	0	0	0	0	0	1	0	0	0	0	0	0	0	0	0	0	0	0
La base a-t-elle un financement pérenne pour : (plusieurs réponses possibles)	Investigateurs	0	1	2	1	20	2	1	1	1	0	2	6	0	3	0	4	2	1	0	1	2	0	1	3
	Non Répondu	7	12	1	4	4	2	2	16	0	11	0	6	10	2	9	8	2	0	1	1	5	1	2	2
	Représentants de malades	0	1	0	0	1	0	0	0	0	0	0	1	0	0	0	1	2	1	0	0	1	0	0	1
	Aucun	4	4	2	4	2	2	1	1	0	0	2	8	1	2	0	9	0	0	0	2	4	0	3	2
	L'exploitation / la valorisation	2	6	0	0	19	0	1	0	1	0	0	3	0	2	0	1	3	1	0	0	0	0	0	0
	L'hébergement	3	6	1	0	19	0	2	0	0	0	0	3	0	1	0	3	3	1	1	0	1	1	0	2
	La saisie	1	3	1	0	19	0	2	0	0	0	0	2	0	3	0	2	2	1	0	0	0	0	0	0
La base a-t-elle fait l'objet d'une déclaration ou autorisation CNIL ?	Le contrôle qualité / le data management	2	3	0	0	19	1	1	0	1	0	0	3	0	2	0	2	2	1	0	0	0	1	0	0
	Ne sait pas	1	3	0	0	4	0	0	16	0	4	0	2	9	0	9	1	0	0	0	0	0	0	0	0
	Non Répondu	0	0	0	1	0	1	0	0	0	7	0	0	0	0	0	0	0	0	0	2	0	0	0	1
Existe-t-il une charte de fonctionnement de la base ?	Ne sait pas	1	0	0	0	3	1	0	16	0	0	0	1	7	0	0	3	0	0	0	0	0	0	0	1
	Non	1	2	1	0	0	0	1	0	0	0	0	2	2	0	0	1	1	0	0	1	3	0	1	0
	Non Répondu	0	0	0	1	0	1	0	0	0	8	0	0	0	0	0	0	0	0	0	0	2	0	0	1
Les patients sont-ils informés du recueil de leurs données ? (plusieurs réponses possibles)	Oui	6	12	2	4	22	2	2	1	1	3	2	9	1	5	9	9	3	1	1	1	2	1	2	3
	Ne sait pas	1	3	0	0	20	1	0	1	0	3	0	0	7	0	9	2	0	0	0	0	1	0	0	0
	Non	4	8	1	3	0	1	2	6	0	0	0	7	2	2	0	7	2	0	1	2	3	1	3	2
	Non Répondu	0	0	0	1	0	1	0	0	0	8	0	0	0	0	0	0	0	0	0	0	2	0	0	1
	Oui	3	3	2	1	5	1	1	10	1	0	2	5	1	3	0	4	2	1	0	0	1	0	0	2
Les patients sont-ils informés du recueil de leurs données ? (plusieurs réponses possibles)	Ne sait pas	1	1	0	0	2	0	0	12	0	2	0	0	8	0	9	0	0	0	0	0	0	0	0	0
	Non	2	1	1	1	0	1	1	0	0	0	0	3	2	1	0	1	1	0	0	1	2	0	0	1
	Non Répondu	0	0	0	1	0	1	0	0	0	8	0	0	0	0	0	0	0	0	0	0	2	0	0	1
	Oui par affichage	2	0	0	1	2	1	0	0	1	0	0	0	0	1	0	4	2	0	0	0	0	1	0	1
Les patients sont-ils informés du recueil de leurs données ? (plusieurs réponses possibles)	Oui par information par l'investigateur/le professionnel de santé	3	6	2	1	0	0	0	0	1	0	2	6	0	3	0	9	2	1	1	0	2	1	1	2
	Oui par signature d'un consentement exprès	3	8	2	3	21	1	2	5	0	1	0	5	0	2	0	4	1	0	0	1	2	0	2	2

Question	modalité	ANDDIRARE	BRAIN TEAM	CARDIOGEN	DEFISCENCE	FAIR	FAVA MULTI	FILFOIE	FILNEMUS	FILSLAN	FIMARAD	FIMATHO	FIRENDO	G2M	Hors filiére	MARIH	MCGRE	MHEMO	MUCO	NeuroSphinx	ORKID	OSCAR	RESPIFIL	SENSGEN	TETE COU
Quel est le support de la base de données ?	Application dédiée développée spécifiquement pour ce recueil (dossier patient informatisé, application hospitalière type dossier de spécialité, application web développée par des informaticiens internes ou externes...)	6	9	2	2	9	1	2	0	1	2	2	5	0	2	1	4	3	1	1	0	0	1	2	3
	Base de données Lotus Approach. Données extraites de la base et collées dans le dossier Word du patient (données basales et résultats du bilan annuel ou suivi post-greffe.	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1	0	0	0	0	0	0	0	0
	CLININFO	0	0	0	0	0	0	0	0	0	0	0	1	0	0	0	0	0	0	0	0	0	0	0	0
	En cours de construction via RADICO	0	0	0	0	0	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
	logiciel OCQM, système en ligne	0	0	0	0	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
	Ne sait pas	0	0	0	0	15	0	0	17	0	0	0	0	4	0	0	0	0	0	0	0	0	1	0	0
	Non Répondu	0	0	0	1	0	1	0	0	0	8	0	0	1	0	0	0	0	0	0	0	3	0	0	1
	Outils de bureautique (Excel, Access, File Maker Pro...)	2	4	1	2	0	1	1	0	0	1	0	6	5	3	8	8	1	0	0	2	3	0	1	1
SPSS	0	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	
Où est hébergée cette base ?	A l'extérieur (prestataire externe, autre site hospitalier, association...) à l'étranger	1	2	0	0	1	0	0	0	0	0	0	3	0	1	0	1	1	0	0	0	0	0	0	0
	A l'extérieur (prestataire externe, autre site hospitalier, association...) en France	1	1	1	0	0	3	0	0	0	1	0	2	0	1	0	2	1	0	0	0	0	1	0	0
	Chez un hébergeur agréé de données de santé	0	2	1	0	8	0	2	0	1	0	0	1	0	0	0	0	1	1	0	0	0	0	0	0
	En local (sur votre site hospitalier)	6	8	1	3	0	0	1	1	0	0	2	6	7	3	5	9	1	0	1	2	4	0	3	4
	Ne sait pas	0	1	0	1	16	0	0	16	0	2	0	0	3	0	4	1	0	0	0	0	0	0	0	0
	Non Répondu	0	0	0	1	0	1	0	0	0	8	0	0	0	0	0	0	0	0	0	0	3	0	0	1
Précisez le pays dans lequel est hébergée votre base :	Allemagne	0	2	0	0	0	0	0	0	0	0	0	1	0	1	0	0	0	0	0	0	0	0	0	0
	Belgique	1	0	0	0	0	0	0	0	0	0	0	1	0	0	0	0	0	0	0	0	0	0	0	0
	Espagne	0	0	0	0	0	0	0	0	0	0	0	1	0	0	0	0	0	0	0	0	0	0	0	0
	Non Répondu	7	12	3	5	24	4	3	17	1	11	2	9	10	4	9	12	3	1	1	2	7	1	3	5
	Suisse	0	0	0	0	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
	UNITED KINGDOM	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1	0	0	0	0	0	0
Êtes-vous en mesure de modifier votre base de données / application ?	USA	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1	0	0	0	0	0	0	0	0	
	Ne sait pas	0	0	1	0	0	16	0	2	0	8	0	1	0	0	1	0	0	0	0	0	0	0	0	0
	Non	0	3	0	1	23	0	0	1	0	0	0	3	0	0	0	0	1	0	0	0	0	0	1	0
	Non Répondu	0	0	0	1	0	1	0	0	0	8	0	0	0	0	0	0	0	0	0	0	3	0	0	1
	Oui par prestation externe	3	1	2	0	0	3	0	0	1	1	0	2	0	3	0	1	2	0	0	0	0	0	1	1
Les données de cette base doivent-elles être ressaisies dans le dossier patient informatisé de votre hôpital ?	Oui par vous-même ou en interne	5	10	1	3	1	0	3	0	0	0	2	7	2	2	8	12	1	1	1	2	4	1	1	3
	Non car il n'y a pas de dossier patient informatisé dans votre hôpital	4	4	0	2	0	1	0	0	0	0	0	3	9	3	0	9	0	0	0	1	0	1	2	2
	Non car la base communique avec le dossier patient de l'hôpital	0	4	0	0	0	0	1	0	0	0	1	0	0	1	0	2	0	1	0	0	1	0	0	0
	Non Répondu	0	0	0	1	0	1	0	0	0	8	0	0	0	0	0	0	0	0	0	0	3	0	0	1
	Oui pour l'activité de soin	4	5	3	3	25	1	2	0	0	2	0	9	1	1	0	2	3	0	1	1	2	1	0	3
	Oui pour le PMSI	1	4	3	2	25	0	2	17	0	3	1	5	1	0	9	2	2	0	0	1	1	1	1	1
La base communique-t-elle avec d'autres bases (inter-régionales, nationales, européennes, internationales) ?	Oui pour les consultations	1	3	2	3	25	1	2	0	1	2	0	5	1	1	0	2	3	0	1	1	1	1	1	0
	Ne sait pas	0	0	0	0	4	0	0	6	0	2	0	1	8	0	9	0	0	0	0	0	0	0	0	
	Non	5	9	3	4	19	2	2	4	0	1	2	10	2	5	0	10	2	0	1	2	3	0	2	2
	Non mais c'est en cours de formalisation ou de réalisation	1	1	0	0	0	1	0	2	1	0	0	0	0	0	0	1	2	0	0	0	1	1	1	2
Ya-t-il eu une déclaration CNIL spécifique pour ce transfert de données ?	Non Répondu	0	0	0	1	0	1	0	0	8	0	0	0	0	0	0	0	0	0	0	0	3	0	0	1
	Oui	2	4	0	0	2	0	1	5	0	0	0	1	0	0	0	2	0	1	0	0	0	0	0	0
	Ne sait pas	0	1	0	0	1	0	0	5	0	0	0	1	0	0	0	1	0	0	0	0	0	0	0	0
	Non	1	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
Non Répondu	6	10	3	5	23	4	2	12	1	11	2	11	10	5	9	11	4	0	1	2	7	1	3	5	
	Oui	1	2	0	0	1	0	1	0	0	0	0	0	0	0	0	1	0	1	0	0	0	0	0	

Question	modalité	ANDDIRARE	BRAIN TEAM	CARDIOGEN	DEFISCIENCE	FAHR	FAVA MULTI	FILFOIE	FILNEMUS	FILSLAN	FIMARAD	FIMATHO	FIRENDO	G2M	Hors filière	MARIH	MCGRÉ	MHEMO	MUCO	NeuroSphinX	ORKID	OSCAR	RESPIFIL	SENSGEN	TETE COU
La base concerne-t-elle :	Non Répondu	0	0	0	1	0	1	0	17	0	8	0	0	0	0	8	0	0	0	0	0	3	0	0	1
	Plusieurs maladies	7	5	3	3	9	1	3	0	1	1	0	4	7	4	1	4	3	0	0	0	3	1	2	3
	Une seule maladie rare	1	9	0	1	16	2	0	0	0	2	2	8	3	1	0	9	1	1	1	2	1	0	1	1
Le recueil est-il effectué pour :	Non Répondu	0	0	0	1	0	1	0	17	0	8	0	0	0	0	8	0	0	0	0	0	3	0	0	1
	Tous les patients du centre de référence concernés par cette ou ces maladie(s)	6	8	2	2	0	3	3	0	1	1	2	4	5	4	0	6	3	1	1	1	3	0	1	3
	Une partie des patients du centre de référence concernés par cette ou ces maladie(s)	2	6	1	2	25	0	0	0	0	2	0	8	5	1	1	7	1	0	0	1	1	1	2	1
Le recueil est-il effectué dans :	Non Répondu	0	0	0	1	0	1	0	17	0	8	0	0	0	0	8	0	0	0	0	0	3	0	0	1
	Tous les sites du centre de référence (ou centre mono-site)	3	9	2	0	25	2	3	0	1	0	1	2	5	2	0	4	3	1	1	2	0	1	3	2
	Une partie des sites	5	5	1	4	0	1	0	0	0	3	1	10	5	3	1	9	1	0	0	0	4	0	0	2
Les centres de compétences participent-ils au recueil de données dans cette base ?	Ne sait pas	1	2	0	0	0	1	0	0	0	2	0	0	2	0	1	1	0	0	0	0	0	0	0	0
	Non	6	5	1	2	0	0	2	0	0	1	0	6	1	3	0	9	0	0	1	0	3	0	3	2
	Non Répondu	0	0	0	1	0	1	0	17	0	8	0	0	0	0	8	0	0	0	0	0	3	0	0	1
	Oui, certains	1	4	1	2	25	0	1	0	0	0	0	6	6	1	0	0	3	0	0	1	1	1	0	2
Pour ces centres de compétences, quel est le mode de saisie des données recueillies ?	Oui, tous	0	3	1	0	0	2	0	0	1	0	2	0	1	1	0	3	1	1	0	1	0	0	0	0
	Données saisies directement par les centres de compétences	0	1	0	1	24	2	1	0	1	0	0	4	4	0	0	0	3	1	0	0	1	1	0	2
	Données saisies par le centre de référence	1	3	2	1	1	0	0	0	0	0	0	2	3	2	0	1	1	0	0	2	0	0	0	0
Les données collectées sont-elles nominatives ?	Non Répondu	7	10	1	3	0	2	2	17	0	11	2	6	3	3	9	12	0	0	1	0	6	0	3	3
	Non	1	7	1	1	25	1	1	0	0	2	1	7	6	1	0	5	2	0	0	1	0	0	0	0
	Non Répondu	1	0	0	1	0	1	0	17	0	8	0	0	0	0	9	0	0	0	0	0	3	0	0	1
Les données concernent-elles : (plusieurs réponses possibles)	Oui	6	7	2	3	0	2	2	0	1	1	1	5	4	4	0	8	2	1	1	1	4	1	3	4
	Les adultes	6	12	3	2	22	3	3	0	1	2	1	10	9	5	0	11	4	1	1	1	3	0	3	2
	Les enfants	7	7	3	4	6	2	2	0	0	2	2	7	8	4	0	9	4	1	0	2	4	1	3	4
Quels types de données sont recueillis dans la base ? (plusieurs réponses possibles)	Les fœtus	6	1	0	1	0	0	0	0	0	0	1	3	3	1	0	0	0	0	0	1	1	0	0	1
	Non Répondu	1	0	0	1	0	1	0	17	0	8	0	0	0	0	9	0	0	0	0	0	3	0	0	1
	Administratives	4	8	3	2	1	3	1	0	1	1	1	4	9	3	0	3	1	0	1	0	3	1	2	2
	Biologiques (hors génétique)	3	9	1	3	24	2	2	0	1	1	1	9	6	4	0	11	4	0	1	2	2	1	0	2
	Cliniques	7	14	3	4	25	3	2	0	1	3	2	12	9	5	0	12	3	1	1	2	3	1	3	4
	DEMOGRAPHIQUES	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1	0	0	0	0	0	0
	Diagnostic	0	0	0	0	0	0	0	0	0	0	0	0	1	0	0	0	0	0	0	0	0	0	0	0
	examens cliniques : ECG/test pharmacologiques	0	0	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
	familiales	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1	0	0
	Génétiques	5	8	3	4	8	3	1	0	1	1	2	5	7	2	0	12	3	1	0	2	2	1	1	2
	Histologiques/anatomo-pathologiques	4	3	0	0	12	0	2	0	0	1	1	6	0	2	0	1	0	0	0	1	2	1	0	2
	Imagerie	3	6	3	2	14	2	2	0	0	1	2	9	4	2	0	8	0	0	1	2	1	1	1	2
	Localisation/Origine ethnique	0	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
	Non Répondu	1	0	0	1	0	1	0	17	0	8	0	0	0	0	9	0	0	0	0	0	3	0	0	1
	Socio-économiques	2	6	0	3	1	2	1	0	0	0	0	3	3	0	0	5	0	1	1	1	0	1	1	2
	suiti	0	0	0	0	0	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
Thérapeutiques	1	10	2	3	23	3	2	0	1	1	2	9	7	3	0	11	3	1	1	0	1	1	1	4	

Question	modalite	ANDDIRARE	BRAIN TEAM	CARDIOGEN	DEFISCIENCE	FAIR	FAVA MULTI	FILFOIE	FILNEMUS	FILSLAN	FIMARAD	FIMATHO	FIRENDO	G2M	Hors filiére	MARIH	MCGRE	MHEMO	MUCO	NeuroSphinx	ORKID	OSCAR	RESPIFIH	SENSGEN	TETE COU
Nombre d'items à remplir par patient [Saisie initiale]	Entre 20 et 60 items	3	3	1	1	0	1	0	0	0	1	1	3	1	0	0	6	1	0	0	0	2	0	0	1
	Moins de 20 items	2	3	0	0	0	0	1	0	0	0	0	3	5	2	0	3	2	0	0	0	2	1	0	2
	Non applicable	0	0	0	0	0	0	0	0	0	1	0	1	0	0	0	0	0	0	0	1	0	0	0	0
	Non Répondu	1	0	0	1	0	1	0	17	0	8	0	0	0	0	9	0	0	0	0	0	3	0	0	1
	Plus de 60 items	2	8	2	3	25	2	2	0	1	1	1	5	4	3	0	4	1	1	1	1	0	0	3	1
Nombre d'items à remplir par patient [Saisie de suivi]	Entre 20 et 60 items	2	3	0	0	1	3	1	0	0	0	1	2	2	3	0	4	1	0	1	0	1	0	1	1
	Moins de 20 items	3	4	2	1	0	0	1	0	0	0	1	2	3	0	0	1	2	0	0	0	2	0	0	1
	Non applicable	2	2	0	0	0	0	0	0	0	2	0	3	1	1	0	2	1	0	0	1	1	0	0	0
	Non Répondu	1	0	0	1	0	1	0	17	0	8	0	0	0	0	9	0	0	0	0	0	3	0	0	1
	Plus de 60 items	0	5	1	3	24	0	1	0	1	1	0	5	4	1	0	6	0	1	0	1	0	1	2	2
Temps de saisie dans la base par patient [Saisie initiale]	De 5 à 15 minutes	3	6	1	2	0	1	1	0	0	0	1	1	1	1	0	5	2	1	0	1	2	0	0	2
	Moins de 5 minutes	0	0	0	0	0	0	0	0	0	0	0	3	4	1	0	1	0	0	0	0	1	1	0	2
	Non applicable	1	0	0	0	0	0	0	0	0	2	0	1	0	0	0	0	0	0	0	1	0	0	0	0
	Non Répondu	1	0	0	1	0	1	0	17	0	8	0	0	0	0	9	0	0	0	0	0	3	0	0	1
	Plus de 15 minutes	3	8	2	2	25	2	2	0	1	1	1	7	5	3	0	7	2	0	1	0	1	0	3	0
Temps de saisie dans la base par patient [Saisie de suivi]	De 5 à 15 minutes	2	4	0	0	0	2	1	0	1	0	1	1	1	3	0	3	2	0	1	1	3	0	1	0
	Moins de 5 minutes	2	4	2	2	0	0	0	0	0	0	1	3	0	0	0	1	1	1	0	0	0	0	0	2
	Non applicable	3	2	0	0	0	0	0	0	0	2	0	3	1	1	0	2	1	0	0	1	1	0	0	0
	Non Répondu	1	0	0	1	0	1	0	17	0	8	0	0	0	0	9	0	0	0	0	0	3	0	0	1
	Plus de 15 minutes	0	4	1	2	25	1	2	0	0	1	1	7	5	1	0	7	0	0	0	0	0	1	2	2
Quel type de personnel est en charge de la collecte des informations ? (plusieurs réponses possibles)	Administratif (secrétaire médicale/opérateur de saisie)	2	2	1	0	0	1	1	0	0	0	1	5	1	0	3	1	1	0	0	2	0	2	2	
	Chercheur	0	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	
	Ingénieur de recherche	0	0	0	0	0	0	0	0	0	0	0	1	0	0	0	0	0	0	0	0	0	0	0	
	Médical	7	14	2	4	25	2	1	0	1	0	2	8	6	5	0	10	4	1	1	0	2	1	2	4
	Ne sait pas	0	0	0	0	0	0	0	0	0	3	0	0	0	0	0	0	0	0	0	0	0	0	0	0
	Non Répondu	1	0	0	1	0	1	0	17	0	8	0	0	0	0	9	0	0	0	0	0	3	0	0	1
	Paramédical (infirmier/consillier en génétique/psychologue)	4	6	2	0	25	1	0	0	0	0	0	2	2	0	0	2	1	1	1	0	1	2	0	
	TEC / ARC du service	1	9	2	2	25	2	2	0	0	0	1	4	3	4	0	4	4	0	0	1	0	0	1	2
TEC / ARC volant	1	1	1	0	1	0	0	0	0	0	1	4	2	0	0	2	1	0	0	1	0	0	0	0	
Quel type de personnel est en charge de la saisie informatique ? (plusieurs réponses possibles)	Administratif (secrétaire médicale/opérateur de saisie)	3	2	2	0	0	2	1	0	1	0	0	3	5	1	0	1	2	1	0	0	1	1	2	2
	Chercheur	0	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	
	CLININFO	0	0	0	0	0	0	0	0	0	0	0	1	0	0	0	0	0	0	0	0	0	0	0	
	Ingénieur de recherche	0	0	0	0	0	0	0	0	0	0	0	1	0	0	0	0	0	0	0	0	0	0	0	
	Médical	2	9	0	1	25	2	1	0	1	0	2	7	5	2	0	8	3	1	1	0	2	1	3	4
	Ne sait pas	0	0	0	0	0	0	0	0	0	3	0	0	0	0	0	0	0	0	0	0	0	0	0	0
	Non Répondu	1	0	0	1	0	1	0	17	0	8	0	0	0	0	9	0	0	0	0	0	3	0	0	1
	Paramédical (infirmier/consillier en génétique/psychologue)	3	4	2	0	25	1	0	0	1	0	0	0	0	1	0	0	2	1	1	0	1	1	2	0
TEC / ARC du service	3	11	2	3	25	3	2	0	1	0	1	4	3	4	0	4	3	0	0	1	0	1	2	2	
TEC / ARC volant	1	1	0	1	25	0	0	0	0	0	1	3	2	0	0	2	1	0	0	1	0	1	0	0	
Ne sait pas	0	3	0	1	7	0	0	0	0	2	0	3	3	0	0	1	0	0	0	0	1	0	0	0	
La base intègre-t-elle le set de données minimum de la BNDMR ?	Non	5	1	0	1	15	0	1	0	0	0	1	4	3	0	4	0	1	0	2	0	0	0	0	
	Non Répondu	1	0	0	1	0	1	0	17	0	8	0	0	0	0	9	0	0	0	0	0	4	0	1	
	Oui entièrement	0	1	0	2	3	1	0	0	0	0	0	2	1	0	0	0	1	0	0	0	0	0	0	
Oui partiellement	2	9	3	0	0	2	2	0	1	1	2	6	2	2	0	8	3	0	1	0	2	1	3	4	
Y a-t-il un consensus sur la dénomination des diagnostics au sein de la base ?	CIM10	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1	0	0	
	DSM-5	0	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	
	INUTILE	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1	0	0	0	0	0	
	Ne sait pas	0	3	0	0	0	1	0	0	3	1	0	0	0	0	0	1	1	0	0	0	0	0	0	
	Non	0	1	0	1	0	1	0	0	0	0	1	2	1	0	4	0	0	1	2	1	0	0	0	
	Non Répondu	5	1	0	1	0	1	2	17	0	8	0	0	4	2	9	2	0	0	0	0	4	0	1	3
	OMIM	0	0	1	0	0	0	0	0	0	0	1	0	0	0	0	0	0	0	0	0	0	0	0	
	Orphanet	1	0	0	1	0	0	0	0	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	
	Orphanet, CIM10	0	2	0	0	0	0	0	0	0	0	0	2	0	0	0	0	0	0	0	0	0	0	0	
	Oui mais nous utilisons un référentiel interne (termes propres au site/au centre)	2	6	2	2	25	1	1	0	0	0	0	7	4	2	0	6	1	0	0	0	2	0	2	2
thésaurus	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1	0	0	0	0	0	0		
thésaurus de spécialité international	0	0	0	0	0	0	0	0	0	0	0	2	0	0	0	0	1	0	0	0	0	0	0		