PROTOCOLE RNIPH
RECHERCHE N'IMPLIQUANT PAS LA PERSONNE HUMAINE

TITRE	Modéliser le suivi biologique des patients DNID pour déterminer des sous-populations de patients à partir de leur observance du suivi biologique.	
TITRE PUBLIC	Modéliser le suivi en laboratoire de biologie médicale du patient diabétique non insulino-dépendant	
TITRE COURT / CODE BIOGROUP	DNID-QH1 / RC59.23G06	
ORGANISME CO-RESPONSABLE DU TRAITEMENT (RT) DE DONNEES	QuantiHealth (SAS) Line GUFFOND LE GOANVIC, présidente, line@quantihealth.fr Mr Demian Ortega, DSI (faisant fonction, DPO/DPD), demian.ortega.m@gmail.com 06 46 45 06 94 Adresse : 28 rue du général de Gaulle 59242 Cappelle en Pévèle /France Rôle : RT conjoint, Propriété de la modélisation et exploitation des résultats. BIOGROUP – DIAGNOVIE Dr Marc Bethencourt, président, marc.bethencourt@biogroup.fr Mr Cyrille Pouchain, DSI (relais DPO), cyrille.pouchain@biogroup.fr 06 87 25 38 18 Adresse : 171 rue Jean Jaurès 59790 Ronchin/FRANCE Rôle : RT conjoint, propriété de la BDD DNID.	
SOUS TRAITANT DU RT RESPONSABLE SCIENTIFIQUE ET RESPONSABLE DE LA MISE EN ŒUVRE DE LA RECHERCHE	INRIA Lille Dr Sophie Dabo, Dr en sciences des mathématiques, sophie.dabo@univ-lille.fr Dr Sophie Perrin, juriste, sophie.perrin@inria.fr 06 25 01 31 28 Adresse : 6 rue Héloïse 59650 Villeneuve d'Ascq /FRANCE	
LISTE DES CENTRES ASSOCIES (CA) ET PARTENAIRES (P)	CA01	BIOGROUP – DIAGNOVIE Dr Magalie THOREZ, biologiste, relais RCI magalie.thorez@biogroup.fr 06 81 17 91 46 Adresse : 171 rue Jean Jaurès 59790 Ronchin/France Rôle : Sélection et information des sujets, recueil des données, propriété de la BDD DIAGNOVIE.
	CA02	BIOGROUP – BIOESTEREL Dr Adrien NEDELEC, biologiste, relais RCI adrien.nedelec@biogroup.fr 06 81 11 60 82 Adresse : Plateau technique du Muy - ZI Les Ferrières 83490 LE MUY/France Rôle : Sélection et information des sujets, recueil des données BIOESTEREL. Cette seconde base serait extraite si la BDD DIAGNOVIE était insuffisante pour modéliser.

CONTEXTE ET JUSTIFICATION DE LA RECHERCHE	**Contexte :** Le diabète est une maladie chronique, touchant 3,5 millions de personnes en France. Elle est très couteuse et génère de très graves complications qui pourraient être évitables. D'après plusieurs études réalisées, l'observance des patients est proche de 40% et le niveau de suivi biologique est également très faible. **Rationnel :** QuantiHealth est la solution intégrée d'amélioration de la qualité du parcours de soins du patient diabétique de type 2 (DNID) avec une action spécifique en prévention tertiaire (complications). Cette solution permettra de répondre au besoin des patients mais également aux recommandations des praticiens. Le suivi des patients se déroulent majoritairement en laboratoire de biologie médical de ville, l'extraction des données de suivi permettra de mettre en place un algorithme pertinent. **Hypothèse :** Nous pensons que la modélisation de l'impact du suivi biologique dont HbA1c sur 10 ans chez les patients atteints de DNID permettra de définir des sous-populations de patient ciblées qui pourraient bénéficier d'un accompagnement via QuantiHealth. Biblio : James, G., Witten, D., Hastie, T., & Tibshirani, R. (2013). An introduction to statistical learning (Vol. 112, p. 18). New York: springer. Witten, Daniela, and Gareth James. An introduction to statistical learning with applications in R. springer publication, 2013. Lebret, Rémi, et al. "Rmixmod: The R package of the model-based unsupervised, supervised, and semi-supervised classification Mixmod library." Journal of Statistical Software 67 (2015): 1-29. HAS : guide de parcours de soins DT2 Entred 1,2,3 : recours aux examens recommandés pour la surveillance du diabète Santé publique France : synthèse épidémiologique 2018 Diabète Etude observance thérapeutique du CRIP 2014 & seconde étude d'approfondissement en 2016.
DEROULEMENT	Le laboratoire Biogroup Diagnovie et QuantiHealh réutilisent les données issues du soin d'environ 10 000 patients. Une fois les patients inclus sont effectués le recueil de donnée dans la BDD de l'étude (Annexe 2).
OBJECTIF PRINCIPAL	La recherche porte sur le ou les objectif(s) suivant(s) : ☒Prévention et traitement ☒Compréhension des maladies OP : Modéliser le suivi biologique des patients DNID et déterminer des sous-populations de patients à partir de leur observance du suivi biologique.

CRITERE DE JUGEMENT PRINCIPAL	Sous population : nous utiliserons tous les paramètres, biologiques, individuels pour chercher des sous-groupes, Homogénéité des populations : les méthodes algorithmes et critère de clustering nous guiderons pour identifier les paramètres les plus représentatifs de la typologie optimale (en termes d'homogénéité des groupes). Classification : homogénéité en pourcentage
OBJECTIFS SECONDAIRES	OS1 : Variation des paramètres biologiques au cours du temps. Détermination de l'observance biologique des patients DNID OS2 : Variation spatiale en fonction des zones géographiques (code postal), et des données cliniques
CRITERE(S) DE JUGEMENT SECONDAIRE(S)	CJS1 : -Taux d'HbA1c en % -Taux de glycémie à jeun en mmol/L -Taux de RACu en g/g CJS2 : données biologiques complémentaires à celles listées dans CJS2, données anthropométriques, code postal, démographie, âge, sexe….

<table>
<tr><td>METHODOLOGIE</td><td colspan="2">Recherche n’impliquant pas la personne humaine :
Rétrospective, Comparative, Multicentrique</td></tr>
<tr><td>NATURE DES DONNEES</td><td colspan="2">Types de données recueillies :
☐ Commune de résidence de la personne étudiée ou Commune de décès
☒ Année et mois de naissance
☒ Date de soins (JJMMAAAA)
☐ Date de décès (JJMMAAAA)
☒ Démographie, Anthropométrie
☒ Données médicales
☐ Aucune</td></tr>
<tr><td rowspan="3">Liste des variables

*Annexe 4 : Lien code pour extraction Diagnovie</td><td>Issues du dossier médical obligatoires</td><td>-démographie : âge (mois et année de naissance), sexe.
-Date de venue en laboratoire
-4 premiers chiffres du code postale de résidence
-Données biologique : HbA1c, Glycémie à jeun.</td></tr>
<tr><td>Issues du dossier médical optionnelles*</td><td>-Anthropométrie : poids, taille, IMC.
-Données biologiques : créatinine (créatininémie, créatininurie), bilan lipidique (cholestérol total, HDL, LDL et triglycérides), microalbuminurie, protéinurie (Idéalement ratio albumine/créatinine urinaire RACu, sinon protéinurie de 24h, sinon protéinurie autre), bilan hépatique (ASAT, ALAT).
-données cliniques si disponible (antécédents médicaux et comorbidités, traitements, …).</td></tr>
<tr><td>Acquise de novo pour l’étude (obligatoire)</td><td>Aucune</td></tr>
<tr><td>SOURCES DE DONNEES UTILISEES (CHOIX MULTIPLES)</td><td colspan="2">☒ Dossiers médicaux</td></tr>
<tr><td>CALENDRIER DE RECUEIL DES DONNEES</td><td colspan="2">Recherche rétrospective :
- Historique des données (dates d’enregistrement des données de santé dans le dossier médical ou autres supports) : janvier 2014 à juin 2023
- Période de recueil des données pour le projet de recherche (dates prévisionnelles de début et de fin) : octobre 2023 à mars 2024
- Modélisation : novembre 2023 à mars 2025
- Durée prévue du contrôle des données et de l’analyse statistique : 6 mois (fin prévisionnelle en septembre 2025)
- Durée totale de la recherche (=période de recueil + durée d’analyse) : 24 mois</td></tr>
</table>

COLLECTE DES DONNEES	La collecte des données cliniques reposera sur une base de données clinique conforme aux données présentes dans les documents sources. **Recueil :** • Les données seront obtenues depuis leur dossier médical informatisé : logiciel KALISIL en version 3 ou ultérieure. • Le recueil sera effectué par le responsable de la mise en œuvre pour le centre principal et par le responsable de chaque centre associé. • Le fichier de recueil est en annexe 2. **Codage :** Les données seront pseudonymisées par un codage au format suivant Numéro de centre à 2 chiffres (01 pour Diagnovie et 02 pour Bioesterel) + 1 initiale du prénom + 1 initiale du nom + numéro unique de sujet (numéro d'ordre ou id patient haché). **Transfert :** Les modalités de transfert : Transfert sécurisée par serveur FTP / SSH / VPN ANSSI Conforme. **Conservation :** La conservation du document établissant la correspondance entre le code du sujet et son identité est faite sur le serveur de chaque CA dans le dossier RCI>RNIPH>DNID-QH1
DUREE DE CONSERVATION DES DONNEES	Les données à caractère personnel relatives aux personnes se prêtant à une recherche, et traitées à cette fin, peuvent être conservées dans les systèmes d'information du responsable de traitement, du centre participant ou du professionnel de santé intervenant dans la recherche au maximum deux ans après la dernière publication des résultats de la recherche ou, en cas d'absence de publication, jusqu'à la signature du rapport final de la recherche. Les données à caractère personnel des professionnels intervenant dans la recherche ne peuvent être conservées au-delà d'un délai de quinze ans après la fin de la dernière recherche à laquelle ils ont participé. Elles font ensuite l'objet d'un archivage sur support papier ou informatique pour une durée conforme à la réglementation en vigueur. Les données doivent ensuite être effacées.

DESCRIPTIF DE L'ANALYSE STATISTIQUE DES DONNEES ET/OU REFERENCES DU BIOSTATISTICIEN RESPONSABLE	**NOMBRE DE SUJETS A INCLURE** Environ 10 000 sujets sur la base des patients issus de Diagnovie. Si nécessaire ajout de 10 000 sujets de la base de Bioesterel **ANALYSE DE DONNEES** La base de données sera stockée conformément à l'engagement investigateur signé. S'agissant d'une première étude dans la thématique, l'analyse statistique sera à la fois descriptive et supervisée. Les variables qualitatives seront décrites par leur effectif et les variables quantitatives seront décrites par leur médiane, et écart type. Le critère de jugement principal sera analysé si possible en utilisant une analyse factorielle, des tests statistiques (test de student, test de CHI^2, Mann Whitney non paramétrique…) et méthodes de clustering Les analyses statistiques seront effectuées par le Pr Sophie DABO (INRIA Lille) à l'aide du logiciel R en version 4.3.1
PERSONNES INCLUSES DANS LA RECHERCHE	**<u>Nombre de sujet inclus :</u>** Minimum 10 000 **<u>Critères d'inclusion :</u>** -Patients diabétique défini comme ayant eu une glycémie à jeun > 7mmol/L (à valider) ou à défaut une HbA1c > 7% -suivi dans un centre de l'étude **<u>Critères de non-inclusion :</u>** -Patient(es) s'opposant à l'utilisation de leurs données dans le cadre de la recherche
Modalités d'information et de traçabilité de la non-opposition/opposition	Les patients seront informés de façon complète et loyale, en des termes compréhensibles, des objectifs de l'étude et de la nature des informations recueillies, et de leur droit de s'opposer à tout moment à l'exploitation des données recueillies. <u>Information individuelle des patients selon G-DPO-M004 :</u> -ils ont reçu, intégré à leur compte rendu d'examen de biologie médicale, une lettre d'information de la collection d'échantillon et données associées de Biogroup (DC-2020-4014 ou son renouvellement ultérieur) ou ils ont été informé par voie d'affichage selon G-PRA-M104 en version 3 ou ultérieure. -ils seront informés de cette étude via le site internet : https://biogroup.fr/specialites/innovation-scientifique-recherche-medicale/ En cas d'opposition du sujet au traitement de ses données personnelles de santé à des fins de recherche, l'opposition sera consignée dans son dossier médical. Ce droit d'opposition s'exerce à tout moment par tout moyen auprès soit du responsable de la recherche soit de l'établissement détenteur des données qui s'engagent à donner suite à cette demande dans un délai maximal de 1 mois.

ETHIQUE	Une soumission au Comité d'Ethique sera effectué si nécessaire avant publication.
LICEITE	L'intérêt est : -légitime : les LBM Biogroup et QuantiHealth ne peuvent pas effectuer cette étude épidémiologique sans cette étude. -licite au regard du droit, déterminé de façon suffisamment claire et précise. -nécessaire pour les objectifs de compréhension de la maladie et de prévention sur les volets du diabète. De plus, le traitement ne heurte aucun droits et intérêts des personnes dont les données sont traitées. Ces personnes ont bénéficié d'une information sur leur droit.
RETOMBEES ATTENDUES	L'analyse des données de suivi des patients DNID issue des laboratoires de ville permettront de déterminer et de définir des sous-groupes de population. Ceux-ci pourraient bénéficier d'un accompagnement de suivi par QuantiHealth Amélioration de l'observance des patients DNID. Publication d'articles de recherche et communication lors d'évènements médicaux ou de recherche académique.

FLOW CHART :

Descriptif des données collectées	Sélection Avant J-30	Inclusion T 0
Information des sujets et absence d'opposition	✓	
Recueil de données		✓
Réutilisation des données pour les objectifs de l'étude		✓

ANNEXES

Annexe 1 : Modalités d'information des sujets :

- Cf. G-DPO-M004 – Annexe 2, 3 et 4 Lettre d'information type
- Cf. G-PRA-M104
- Site internet :

LIBELLE	BIOLOGISTE REFERENT /RT	PERIODE	SEL ET SITES CONCERNES	FICHE DESCRIPTIVE
Modéliser le suivi biologique des patients DNID pour déterminer des sous-populations de patients à partir de leur observance du suivi biologique.	Dr Magalie THOREZ Dr Line GUFFOND LE GOANVIC	Janvier 2014 à Juin 2023	DIAGNOVIE BIOESTEREL	Sera inséré ici le lien vers ce protocole

- Annexe 2 - Fichier de recueil de données (Fichier Excel/ Autre, document séparé)

- Annexe 3 : Circuit des échantillons issus du corps humain (EICH) et examen de biologie effectué :

Non applicable

- Annexe 4 : Lien code pour extraction des données Diagnovie (séparé).