PROTOCOLE RNIPH
RECHERCHE N'IMPLIQUANT PAS LA PERSONNE HUMAINE

TITRE	Etude sur données permettant d'améliorer les algorithmes du JIMINI en vue de meilleur performance analytique de rendu de résultat : une étude rétrospective	
TITRE PUBLIC	Amélioration des algorithmes du dispositif JIMINI pour la réalisation des analyses urinaires en laboratoire de biologie médicale.	
TITRE COURT / CODE BIOGROUP	JIMINI3-DATA / RC06.23H01	
ORGANISME CO-RESPONSABLE DU TRAITEMENT DE DONNEES	**USENSE** Dr Guillaume LEMETAIS (CEO), glemetais@usense.healthcare / 0617778641 Mr Jonathan FRANKLIN (CFAO, faisant fonction de DPO) jfranklin@usense.healthcare / 01 72 60 51 16 Adresse : 96 bis boulevard Raspail, 75006 Paris/ France Rôle : Traitement des données, dont archivage à l'issue de l'étude.	
	BIOGROUP – BIOESTEREL Dr Eric SAVOY, Président, eric.savoy@biogroup.fr Dr Laurent SCHLEGEL, Biologiste (DPO/DPD), laurent.schlegel@biogroup.fr Adresse : 405 avenue de cannes 06210 Mandelieu la Napoule /France Rôle : Recueil et traitement des données, dont archivage à l'issue de l'étude.	
RESPONSABLE SCIENTIFIQUE ET RESPONSABLE DE LA MISE EN ŒUVRE DE LA RECHERCHE	**USENSE** Dr Guillaume LEMETAIS (Dr en science), glemetais@usense.healthcare / 0617778641 Mr Amir OUERIEMMI (chef de projet, Head of Science) aoueriemmi@usense.healthcare / 01 72 60 51 16 Adresse : 96 bis boulevard Raspail, 75006/ FRANCE Rôle : développement du JIMINI et de ses algorithmes, coordination du projet, analyses statistiques, développement des modèles de machine learning et formation des partenaires.	
LISTE DES : **CENTRES ASSOCIES (CA)** **PARTENAIRES (P)** **DESTINATAIRE (D)** **SOUS-TRAITANT (ST)**	CA	**BIOGROUP-BIOESTEREL** Dr Adrien NEDELEC, Biologiste et responsable DSI adrien.nedelec@biogroup.fr , 06 81 11 60 82 Adresse : impasse des bruyères, zone industrielle de l'argile voie H 06370 Mouans-Sartoux /FRANCE Rôle : Sélectionner et informer les sujets, colliger et transmettre les données à USENSE

CONTEXTE ET JUSTIFICATION DE LA RECHERCHE	**Contexte et Justification de l'étude :** Dans de nombreux domaines médicaux, les professionnels de santé s'interrogent souvent sur la meilleure façon d'appréhender l'état de santé global d'un individu en intégrant la mesure de nombreux paramètres de dosage biologiques afin de mieux orienter leurs patients. A titre d'exemple, il n'existe pas, à ce jour, de données publiées sur les valeurs usuelles attendues des paramètres urinaires ou salivaires, individuels ou combinés. D'autant que de nombreuses études de cohortes prospectives ont montré un lien entre l'excrétion de certains paramètres urinaires et le risque de développement de pathologies (1, 2, 3) dans l'urine ou dans la salive (4, 5). De même le lien entre marqueurs biologiques sanguins, urinaires et salivaires n'a encore été que très peu exploré (6). Or, ces informations, contextualisées, pourraient permettre de mieux appréhender les caractéristiques physiologiques/métaboliques individuelles de chaque patient afin d'initier une démarche de médecine préventive et prédictive. Une étude des distributions de chaque biomarqueur pour différentes catégories de la population générale (âge, genre, BMI, pathologie…) ainsi que l'association entre biomarqueurs sanguin, urinaires et salivaires permettrait donc de disposer de données solides, afin d'améliorer encore la prise en charge des patients. **Hypothèse :** Il est possible d'identifier des valeurs usuelles spécifiques dans plusieurs sous-catégories de la population générale et de créer des indicateurs de risques combinant marqueurs sanguins, urinaires et salivaires afin d'identifier des risques de développement de pathologies. **Bibliographie :** 1. Mente, Andrew, et al. "Associations of urinary sodium excretion with cardiovascular events in individuals with and without hypertension: a pooled analysis of data from four studies." The Lancet 388.10043 (2016): 465-475. 2. O'Donnell, Martin, et al. "Urinary sodium and potassium excretion, mortality, and cardiovascular events." New England Journal of Medicine 371.7 (2014): 612-623. 3. Graudal, Niels, et al. "Compared with usual sodium intake, low-and excessive-sodium diets are associated with increased mortality: a meta-analysis." American journal of hypertension 27.9 (2014): 1129-1137. 4. Lee, Angie, et al. "Bacterial and salivary biomarkers predict the gingival inflammatory profile." Journal of periodontology 83.1 (2012): 79-89. 5. Zhang, Chen-Zi, et al. "Saliva in the diagnosis of diseases." International journal of oral science 8.3 (2016): 133-137. 6. Do, Kieu Trinh, et al. "Network-based approach for analyzing intra-and interfluid metabolite associations in human blood, urine, and saliva." Journal of Proteome Research 14.2 (2015): 1183-1194.
DEROULEMENT	Le centre de l'étude réutilise les données issues du soin d'environ 350 000 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) (choix multiple) ☒Diagnostics OP : Améliorer les algorithmes du JIMINI en vue d'une meilleure performance analytique et Calcul des valeurs usuelles des marqueurs biologiques urinaires, sanguins et salivaires prescrit lors d'analyse de routine et mesurés en laboratoires dans la population générale, stratifiés par catégorie d'âge et par genre

CRITERE DE JUGEMENT PRINCIPAL	CJP : Distribution de données pour chaque marqueur biologique et proportion de patient par catégorie d'âge et par genre	
OBJECTIFS SECONDAIRES (si applicable et si données disponible)	1. Distribution de chaque marqueur physiologique par type de pathologie et évolution au cours du temps 2. Comparaison données automates vs données bandelettes urinaires 3. Analyses complémentaires prescrites pour chaque biomarqueur 4. Temps moyen pour l'analyse selon type d'examen 5. Pourcentage de patients venant pour un examen en suivi et biomarqueurs associés 6. Volumétrie des examens urinaires par site et pour chaque marqueur biologique 7. Pathologies pour lesquelles le plus d'examens urinaires sont prescrits 8. Liens entre biomarqueurs 9. Identification index /facteur de risques de pathologies 10. Développement de ratios d'intérêt médical basés sur le croisement de données biologiques 11. Corrélation entre données de consommation et marqueurs biologiques	
CRITERE(S) DE JUGEMENT SECONDAIRE(S)	Chaque objectif est évalué statistiquement par : -Coefficient de corrélation, biais et erreur de la bandelette vs automate dans la population générale et dans la population stratifiée par catégorie d'âge et par genre -Distribution de la population sur les ratios développés par âge, genre et pathologie	
METHODOLOGIE	Recherche n'impliquant pas la personne humaine Rétrospective, Descriptive, Monocentrique	
NATURE DES DONNEES	Types de données recueillies : ☐ Commune de résidence de la personne étudiée ou Commune de décès ☒ Année et mois de naissance, Démographie ☒ Date de soins (JJMMAAAA), ☐ Date de décès (JJMMAAAA) ☐ Anthropométrie ☒ Données médicales	
Liste des variables	Issues du dossier médical souhaitable	Démographie : Sexe, âge (mois, année) Date de soin Résultats d'examen de biologie urinaire effectué dans le cadre du soin : -Examen microbiologique des urines (ECBU) ionogramme Urinaire (K, NA, Cl, Mg…), créatinine urinaire, albumine urinaire, ratio albumine/créatinine urinaire, protéine Urinaire - Volume urinaire 24h si urine des sur 24h. - Résultats d'examen de biologie sanguine effectué dans le cadre du soin : Créatinine, DFG

	Issues du dossier médical ou acquises *de novo* (optionnelles)	Anthropométrie : Taille, Poids (IMC) Données clinique : Pathologie, Traitement Données biologiques si disponible : Bandelette urinaire, Heure de réception de l'échantillon si recueil mictionnelle, HbA1C, glycémie à jeun, protéinémie.
	Acquise *de novo* pour l'étude (obligatoire)	NA
SOURCES DE DONNEES UTILISEES (CHOIX MULTIPLES)	☒ Dossiers médicaux	
CALENDRIER DE RECUEIL DES DONNEES	- Historique des données (dates d'enregistrement des données de santé dans le dossier médical ou autres supports) : JANV/2023 à AOUT/2023 - Période d'inclusion et/ou de recueil des données pour le projet de recherche (dates prévisionnelles de début et de fin) : SEPT/2023 à DEC/2024 - Durée prévue du contrôle des données et de l'analyse statistique : 6 mois - Durée totale de la recherche (=période de recueil + durée d'analyse) : 10 mois	
COLLECTE DES DONNEES	La collecte des données cliniques reposera sur une base de données clinique conforme aux données présentes dans les documents sources. **<u>Recueil :</u>** • Les données seront obtenues directement auprès des sujets par questionnaire ou depuis leur dossier médical informatisé : logiciel KALISIL en version 3.0 ou ultérieure • Le recueil sera effectué par le responsable de la mise en œuvre pour le centre. • Le fichier de recueil est en annexe 2. **<u>Codage :</u>** Les données seront pseudonymisées par un codage au format suivant ID patient haché. **<u>Transfert :</u>** Les modalités de transfert de données seront sécurisées via Lockself / Locktransfert (ou autres solution ANSSI conforme). **<u>Conservation :</u>** Lieu de conservation du document établissant la correspondance entre le code du sujet et son identité : Serveur de chaque centre dans le dossier RCI>RNIPH>JIMINI3-DATA	

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 support 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 350000 sujets **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 essentiellement descriptive. Les variables qualitatives seront décrites par leur effectif et les variables quantitatives seront décrite par leur médiane, et écart type. Le critère de jugement principal sera analysé si possible en utilisant un test de CHI^2 / Mann Whitney non paramétrique. Les analyses statistiques seront effectuées par Mr Remy Wahnoun (Head of Data Science, Usense) à l'aide du langage Python en version 3.10 ou supérieure.
PERSONNES INCLUSES DANS LA RECHERCHE	**NOMBRE DE SUJETS A INCLURE :** environ 350000 sujets **<u>Critères d'inclusion :</u>** -Patient de la SELAS BIOESTEREL ayant un bilan urinaire dans le cadre du soin en 2023 **<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. Information individuelle des patients : Les sujets 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) et seront informé 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	Ce traitement relève de l'Intérêt légitime du RT car il est : -Légitime : USENSE et les LBM Biogroup ne peuvent pas améliorer le JIMINI et ses algorithme sans cette étude. -Licite au regard du droit, déterminé de façon suffisamment claire et précise. -Nécessaire pour les objectifs d'amélioration de la prise en charge des patients dans le diagnostic des pathologies néphrologiques et urologiques en France. De plus, le traitement ne heurte aucun droit ni intérêt des personnes dont les données sont traitées. Ces personnes ont bénéficié d'une information sur leurs droits.
RETOMBEES ATTENDUES	Amélioration de l'algorithme du JIMINI pour améliorer la performance analytique et le rendu résultat.

FLOW CHART :

Descriptif des données collectées	Sélection Avant J-30	Inclusion T 0
Information des sujets et absence d'opposition	✓	
Recueil de données		✓

ANNEXES

- Annexe 1 : Modalités d'information des sujets : G-PRA-M104 et le site internet Biogroup

LIBELLE	BIOLOGISTE REFERENT	PERIODE	SEL ET SITES CONCERNES	FICHE DESCRIPTIVE
Amélioration des algorithmes du dispositif JIMINI pour la réalisation des analyses urinaires en laboratoire de biologie médicale.	Dr Adrien NEDELEC (BIOESTEREL) Et son partenaire : USENSE	Sept à Déc 2023	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 EBM effectué : Non applicable