La Présidente

Monsieur Yazdan YAZDANPANAH
DIRECTEUR
INSTITUT NATIONAL DE LA SANTÉ ET DE LA RECHERCHE MÉDICALE - ANRS-MALADIES INFECTIEUSES ÉMERGENTES
101 RUE DE TOLBIAC
75013 - PARIS

Paris, le 19 février 2021

N/Réf. : MLD/AVL/VCS/NDT211003

Objet : NOTIFICATION D'AUTORISATION

Décision DT-2021-006 autorisant l'INSTITUT NATIONAL DE LA SANTÉ ET DE LA RECHERCHE MÉDICALE - ANRS – MALADIES INFECTIEUSES ÉMERGENTES à mettre en œuvre un traitement de données ayant pour finalité la constitution d'un entrepôt de données hospitalier français de l'infection à VIH. (Demande d'autorisation n° 918266)

La Commission a été saisie, d'une demande d'autorisation relative à un traitement de données à caractère personnel. Ce traitement, a pour base légale l'exercice d'une mission d'intérêt public, au sens de l'article 6-1-e du Règlement européen sur la protection des données. Il relève de la procédure prévue aux articles 44 3° et 66 III de la loi du 6 janvier 1978 modifiée.

Responsable de traitement	L'Agence Nationale de Recherche sur le Sida et les hépatites virales (ANRS), agence autonome de l'Institut national de la Santé et de la Recherche Médicale (INSERM).
Finalité	L'entrepôt de données hospitalier français de l'infection à VIH a vocation, par la réalisation de recherches épidémiologiques, à permettre : - l'amélioration des connaissances sur les caractéristiques des personnes vivant avec le VIH et sur leurs conditions de prise en charge ; - l'évolution et l'adaptation des pratiques de soins ; - l'amélioration des stratégies thérapeutiques ;

	- l'évaluation de l'efficacité des nouveaux traitements. Les traitements de données de santé à caractère personnel qui seront mis en œuvre à des fins de recherche, d'étude ou d'évaluation dans le domaine de la santé à partir des données contenues dans cet entrepôt sont des traitements distincts qui devront faire l'objet de formalités propres au titre des articles 66 et 72 et suivants de la loi « informatique et libertés ».
Gouvernance	La base de données est coordonnée et gérée au sein de l'Institut d'épidémiologie et de santé publique Pierre Louis (IPLESP) par une unité mixte de recherche INSERM/Sorbonne Université. Un conseil scientifique composé de représentants des centres qui aliment la base de données, de représentants de l'équipe IPLESP, d'un représentant des associations des personnes vivant avec le VIH (TRT5 : Groupe inter associatif traitements & recherche thérapeutique) et d'un représentant de l'INSERM/ANRS se réunit tous les 6 mois, afin de mener une réflexion sur le développement et l'évolution de la base. Ce conseil est également chargé d'évaluer les projets de recherche et d'étude nécessitant l'utilisation des données de l'entrepôt selon leur faisabilité, leur pertinence scientifique et leur caractère d'intérêt public.
Données traitées	Les données à caractère personnel traitées sont issues des dossiers médico-épidémiologique du VIH (via un logiciel intitulé « DOMEVIH » ou via un dossier médical informatisé local) des patients pris en charge dans les hôpitaux ou services appartenant aux instances COREVIH (Coordination régionale de la lutte contre le VIH) : - numéro de pseudonymisation (constitué à partir du nom, prénom, jour et mois de naissance) ; - numéro de centre ; - date de naissance complète ; - sexe ; - département de résidence ; - pays de naissance ; - année d'arrivée en France ; - nationalité de naissance et nationalité « actuelle » ; - séjours hors de France ; - pathologies ; - affections ; - traitements antirétroviraux ; - données de suivi des infections VHB ; - comportements à risque (consommation de psychotropes) ; - habitudes de vie ; - situation familiale ; - type de logement ; - niveau d'étude et/ou diplôme ;

	- situation professionnelle ; - groupes de transmission du VIH, VHC, VHB ; - bénéficiaire de CMU/ALD ; - données de sérologie VIH, VHB, VHC ; - variables biologiques (ADN proviral, HbA1c, créatinine) ; - statut vital et cause de décès (dossier médical).
Destinataires	Sont destinataires des données les membres du personnels habilités, soumis au secret professionnel, dans les strictes limites de leur besoin d'en connaître pour l'exercice de leurs missions s'inscrivant dans les finalités de la base de données hospitalière française sur l'infection à VIH : - des centres hospitaliers participant à l'alimentation de la base de données ; - des membres des équipes de recherche de l'INSERM-ANRS ; - des organismes extérieurs issus du secteur public ou privé, dans le cadre de la mise en œuvre de projets de recherche ou d'étude nécessitant le traitement de données de l'entrepôt .
Information et droits des personnes concernées	*S'agissant des patients dont les données ont été collectées préalablement à la présente autorisation :* Compte tenu du nombre de patients concernés (180 000), de l'ancienneté d'une partie des données (patients pris en charge depuis 1990), du coût économique engendré par l'information individuelle, l'INSERM-ANRS estime qu'informer individuellement les patients constituerait un effort disproportionné. En application de l'article 14-5-b du RGPD et de l'article 69 de la loi informatique et libertés modifiée, l'obligation d'information individuelle de la personne concernée peut faire l'objet d'exceptions dans l'hypothèse où la fourniture d'une telle information se révèlerait impossible, exigerait des efforts disproportionnés ou compromettrait gravement la réalisation des objectifs du traitement. En pareils cas, conformément au RGPD, le responsable de traitement prend des mesures appropriées pour protéger les droits et libertés, ainsi que les intérêts légitimes de la personne concernée, y compris en rendant les informations publiquement disponibles. Le responsable de traitement diffusera sur son site web une information relative à la base de données ainsi que des informations sur les projets de recherche menés à partir des données qu'elle contient. Ces documents d'informations devront comporter l'ensemble des mentions prévues par le RGPD.

	S'agissant des patients dont les données seront collectées postérieurement à la présente autorisation : Une note d'information individuelle, qui devra comporter l'ensemble des mentions prévues par le RGPD sera remise aux patients. Les droits des patients s'exercent auprès du médecin les prenant en charge dans le centre hospitalier où les données sont recueillies, auprès du directeur de l'établissement ou par l'intermédiaire du délégué à la protection des données de l'INSERM-ANRS. La Commission rappelle que l'information relative à la constitution de la base ne peut se substituer à l'information individuelle préalable prévue par les dispositions du RGPD et de la loi « informatique et libertés », qui devra être réalisée pour chaque traitement de données réalisé à partir des données de la base.
Mesures de sécurité	En premier lieu, la Commission relève que le responsable de traitement a réalisé une analyse d'impact relative à la protection des données afin de démontrer la conformité de l'entrepôt de données de santé au RGPD et notamment à son article 32 ; elle prend acte de cette démarche ainsi que du plan d'actions qui l'accompagne. La Commission relève que l'entrepôt ne comporte pas de données directement identifiantes concernant les patients : seules des données pseudonymisées y sont stockées : plus spécifiquement, les noms, prénoms et identifiants patient sont remplacés par un numéro pseudonyme unique produit par une fonction de hachage cryptographique combinée avec une clé secrète. Lors d'une demande de réidentification d'un patient, ce numéro pseudonyme est remonté dans les centres qui réappliqueront cette fonction de hachage aux données d'identification afin de réidentifier le patient. Les données pseudonymisées transmises par les centres sont tout d'abord stockées temporairement pendant une année dans un serveur indépendant déporté dans un centre de données sécurisé. Ces données y sont chiffrées au repos. Des mesures de filtrage sont mises en place afin de restreindre l'émission et la réception des flux réseau à des machines identifiées et autorisées : seules des adresses IP préalablement définies peuvent se connecter à ce serveur. L'accès à ce serveur s'effectue au travers de tunnels sécurisés protégés par l'utilisation de protocoles d'authentification et de chiffrement asymétrique et symétrique à l'état de l'art. Ces accès administrateur sont tracés. Les données pseudonymisées sont ensuite transmises à des serveurs sécurisés de l'IPLESP. Une plateforme centralisée gère les accès administrateur à ces différents serveurs, qui sont effectués exclusivement depuis le réseau local.

	L'accès entre cette plateforme et les serveurs nécessite une authentification forte, et se fait au travers de tunnels sécurisés protégés par l'utilisation de protocoles d'authentification et de chiffrement asymétrique et symétrique à l'état de l'art. Ces accès administrateur sont tracés. La Commission recommande en outre que les données soient chiffrées au repos par un algorithme à l'état de l'art similairement au procédé utilisé pour le stockage temporaire des données. Elle note qu'une étude de faisabilité pour ce chiffrement au repos sera effectué dans les meilleurs délais. Concernant les accès des profils utilisateur, la Commission prend acte qu'une plateforme web permet la traçabilité des demandes d'accès, des actions effectuées et des données personnelles auxquelles les utilisateurs ont accès. Les données de recherche sont hébergées dans des espaces projets dédiés dont l'accès utilisateur est protégé par une authentification conforme à la délibération n° 2017-012 du 19 janvier 2017 portant adoption d'une recommandation relative aux mots de passe. La Commission recommande la mise en place pour ces profils d'une authentification forte similaire à celle mise en œuvre pour les profils administrateur. La Commission rappelle que les exports de données dans les espaces projets doivent être soumis à une procédure dédiée afin de vérifier notamment, conformément à l'état de l'art en la matière, le nombre minimal de patients acceptables dans un jeu de données, l'attribution d'un identifiant patient différent pour chaque export (sauf impossibilité due au suivi de cohorte) et le maintien du secret concernant le numéro pseudonyme et sa clé de génération, qui ne seront jamais exportés ni exposés. La Commission recommande ensuite que l'ensemble des traces techniques soient conservées pour une durée de six mois et qu'une surveillance manuelle ou automatique régulière de ces traces soit réalisée pour chaque serveur de l'entrepôt. La Commission recommande en outre que des sondes de détections d'intrusion soient mises en place afin de détecter les comportements anormaux et de lever des alertes le cas échéant. Sous réserve des précédentes observations, les mesures de sécurité décrites par le responsable de traitement sont conformes à l'exigence de sécurité prévue par les articles 5-1-f et 32 du RGPD. La Commission rappelle toutefois que cette obligation nécessite la mise à jour des mesures de sécurité et de l'analyse d'impact relative à la protection des données au regard de la réévaluation régulière des risques.
Transferts	La présente décision ne vaut pas autorisation de transfert de données en dehors de l'Union européenne vers un pays ne présentant pas un niveau de protection adéquat.

Durée de conservation des données	Les données sont conservées dans l'entrepôt pendant 25 ans à compter de la présente autorisation, puis seront supprimées.

AUTORISE l'INSTITUT NATIONAL DE LA SANTÉ ET DE LA RECHERCHE MÉDICALE - ANRS – MALADIES INFECTIEUSES ÉMERGENTES à mettre en œuvre le traitement, en application de l'article 13 de la loi précitée et de la délibération n° 2019-021 du 28 février 2019 portant délégation d'attributions de la Commission de l'informatique et des libertés à son président et à son vice-président délégué, j'autorise la mise en œuvre de ce traitement.

M-L. De

Marie-Laure DENIS