

Délibération n° 2025-014 du 13 février 2025 autorisant l'Agence européenne des médicaments à mettre en œuvre des traitements automatisés de données à caractère personnel ayant pour finalité des études portant sur l'estimation d'incidence et de prévalence des pathologies dans la population générale en France dans le cadre du projet DARWIN EU.

()

La Commission nationale de l'informatique et des libertés,

Saisie le 24 octobre 2024 par l'Agence européenne des médicaments d'une demande d'autorisation concernant des traitements automatisés de données à caractère personnel ayant pour finalité des études portant sur l'estimation d'incidence et de prévalence des pathologies dans la population générale en France dans le cadre du projet DARWIN EU ;

Vu le règlement (UE) 2016/679 du Parlement européen et du Conseil du 27 avril 2016 relatif à la protection des personnes physiques à l'égard du traitement des données à caractère personnel et à la libre circulation de ces données, et abrogeant la directive 95/46/CE (règlement général sur la protection des données) ;

Vu le règlement (UE) 2018/1725 du Parlement européen et du Conseil du 23 octobre 2018 relatif à la protection des personnes physiques à l'égard du traitement des données à caractère personnel par les institutions, organes et organismes de l'Union et à la libre circulation de ces données, et abrogeant le règlement (CE) n° 45/2001 et la décision n° 1247/2002/CE ;

Vu la loi n° 78-17 du 6 janvier 1978 modifiée relative à l'informatique, aux fichiers et aux libertés, notamment ses articles 66, 72 et suivants ;

Vu l'avis favorable avec recommandations du Comité éthique et scientifique pour les recherches, les études et les évaluations dans le domaine de la santé du 3 octobre 2024 ;

Vu le dossier et ses compléments ;

Sur la proposition de Mme Marie Zins, commissaire, et après avoir entendu les observations de M. Damien Milic, commissaire du Gouvernement,

Formule les observations suivantes :

1. La demande d'autorisation s'inscrit dans le cadre des activités du réseau DARWIN EU (« Data Analysis and Real-World Interrogation Network »), coordonné par l'Agence européenne des médicaments (EMA).
2. Préfigurateur déjà opérationnel de l'espace européen des données de santé, ce réseau vise à favoriser l'utilisation de données de santé en vie réelle afin d'étudier l'utilisation, la sécurité ainsi que l'efficacité des médicaments et des vaccins à usage humain. Les études commanditées par les comités de l'EMA tels que le Comité des médicaments à usage humain (CHMP) et le Comité des médicaments orphelins (COMP) complètent celles présentées par les demandeurs et titulaires d'autorisations de mise sur le marché de médicaments. Elles leur permettent d'éclairer leurs décisions tout au long du cycle de vie des médicaments.
3. Ce réseau s'appuie sur :
 - un centre de coordination assuré par le Centre médical universitaire Erasmus de Rotterdam ;
 - en 2025, une trentaine de partenaires de données (« data partners ») issus de différents pays européens et disposant de données de vie réelle standardisées issues des hôpitaux, de registres, de remboursements ou de biobanques. Ces bases de données sont mises au format OMOP-CDM (« Observational Medical Outcome Partnership – Common Data Model ») afin de favoriser leur interopérabilité et la mise en œuvre d'apprentissage fédéré à l'échelle européenne ;
 - un catalogue d'analyses standardisées des données développé par le centre de coordination (avec des scripts R associés) destiné à permettre la production rapide de résultats à partir des bases de données des partenaires.
4. Le centre de coordination envisage de réaliser annuellement entre trente-cinq et cinquante études nécessitant le requêtage de plusieurs bases de données du réseau DARWIN. En fonction de leur complexité, ces études doivent être mises en œuvre par les partenaires de données du réseau dans des délais allant de trois à douze mois.
5. Chacune de ces études fait l'objet d'un protocole spécifique rédigé et transmis par le réseau DARWIN à chaque partenaire de données européen participant au protocole. Ce protocole permettra à chaque partenaire de calculer les mêmes indicateurs de façon pertinente (critères d'inclusion et d'exclusion, variables, période d'intérêt, *etc.*) et de manière identique pour la ou les pathologies étudiées. Les résultats agrégés de chacune de ces études sont ensuite transmis au centre de coordination du réseau qui réalise les méta-analyses et produit le rapport d'étude.

6. En France, des données du Système national des données de santé (SNDS) sont susceptibles d'être traitées, lorsque cela est pertinent pour les besoins des études commanditées par le réseau DARWIN, par le groupement d'intérêt public « Plateforme des données de santé » (ci-après la PDS) en tant que « partenaire de données ».
7. Dans ce cadre, deux demandes d'autorisation ont été déposées auprès de la CNIL :
 - la première vise à permettre la mise en œuvre d'études destinées à estimer en routine la prévalence et l'incidence de l'utilisation des médicaments et des vaccins en France grâce à des indicateurs standardisés (demande d'autorisation n° 924315) ;
 - la seconde vise à permettre la mise en œuvre d'études destinées à estimer en routine la prévalence et l'incidence des pathologies grâce à des indicateurs standardisés (demande d'autorisation n° 924316).
8. D'après les précisions apportées par l'EMA, d'autres demandes d'autorisation sont susceptibles d'être ultérieurement déposées auprès de la CNIL.
9. A terme, le réseau DARWIN a vocation à utiliser l'infrastructure dédiée qui sera mise en place suite à l'entrée en application du chapitre IV du règlement sur l'espace européen des données de santé (qui devrait avoir lieu en 2028).

Sur le régime juridique applicable :

10. L'EMA étant une agence décentralisée de l'UE, les dispositions du règlement (UE) 2018/1725 régissent les traitements de données personnelles qu'elle met en œuvre.
11. L'accès aux données du SNDS, notamment mises à disposition par la PDS, est, conformément aux dispositions des articles L. 1460-1 et suivants du code de la santé publique, possible pour les traitements respectant les conditions fixées par ces dispositions. Ces dernières prévoient que ces traitements sont soumis aux formalités préalables prévues par les articles 66 et suivants de cette loi, laquelle renvoie également à l'application du RGPD.

Sur le responsable de traitement :

12. L'EMA détermine la finalité et les moyens des traitements de données du SNDS nécessaires aux études qui seront mises en œuvre dans le cadre de la présente autorisation.

Sur les sous-traitants, et notamment le rôle dévolu à la Plateforme des données de santé :

13. La PDS interviendra dans le cadre de l'extraction des données du SNDS depuis le portail de la Caisse nationale de l'assurance maladie (CNAM) en sa qualité de responsable conjoint du SNDS.

14. Seules les données strictement nécessaires et pertinentes au regard des objectifs décrits dans le protocole de recherche commun aux études seront ciblées puis extraites depuis le portail de la CNAM par la PDS. Tous les filtrages de données seront réalisés en amont de la transmission des données vers la solution technique de la PDS. La vérification de la concordance entre les données nécessaires à la mise en œuvre des traitements décrits dans le dossier et les données extraites sera effectuée en lien avec la CNAM.
15. Le dossier de demande mentionne que la PDS interviendra en qualité de sous-traitant pour la mise au format OMOP-CDM des données extraites depuis le portail de la CNAM ainsi que leur analyse. S'agissant des autres opérations de traitement dont l'hébergement des données de l'échantillon de l'EMA, la CNIL estime que la PDS interviendra en qualité de sous-traitant.
16. La répartition des rôles et responsabilités entre l'EMA et la PDS, concernant ces aspects, notamment la conservation des données, la sensibilisation des utilisateurs du projet, la surveillance des traces, la gestion des alertes et des incidents ainsi que les gestions des exports de données anonymes, devra être formalisée par une convention entre les deux parties conformément à l'article 28 du RGPD.
17. Tout traitement des données réalisé à partir de l'échantillon du SNDS est placé sous la responsabilité de traitement exclusive de l'EMA et ne peut avoir lieu que sous couvert de ses instructions documentées.

Sur la finalité des traitements et leur caractère d'intérêt public :

18. Les traitements envisagés ont pour finalité la mise en œuvre d'études destinées à estimer en routine la prévalence et l'incidence des pathologies selon une méthodologie standardisée. Ces estimations pourront notamment être stratifiées selon l'âge, le sexe ou autre sous-groupe d'intérêt. En fonction des pathologies envisagées, elles pourront être également réalisées sur différentes périodes.
19. La finalité des traitements est déterminée, explicite et légitime, conformément au b) du 1 de l'article 5 du RGPD. La CNIL estime que ce traitement présente une finalité d'intérêt public, conformément au I de l'article 66 de la loi « informatique et libertés ».

Sur la base légale des traitements et l'exception permettant de traiter des données sensibles :

20. Les traitements envisagés par l'EMA sont nécessaires à l'exécution de la mission d'intérêt public dont elle est investie. Ces traitements sont, à ce titre, licites au regard de l'article 5 de la loi « informatique et libertés », correspondant au e) du 1 de l'article 6 du RGPD. En outre, ils sont nécessaires à des fins de recherche scientifique et permettent de traiter des données de santé dans le cadre de la présente autorisation prise sur le fondement des articles 66 et suivants de la loi « informatique et libertés et du j) du 2 de l'article 9 du RGPD.

Sur la nature des données traitées :

21. La nature et les catégories de données dont le traitement est envisagé dans ces études ont été justifiées et précisées dans le cadre de l'instruction de la demande d'autorisation.
22. Un échantillon aléatoire de la population française représentatif en termes d'âge, de sexe et de département de résidence sera sélectionné dans la base principale du SNDS. Afin d'évaluer les incidences et prévalences des pathologies en France, à l'exception des maladies rares, cet échantillon comportera les données de dix millions de personnes, sous réserve de l'exercice du droit d'opposition.
23. Certaines données du Système national d'information inter-régimes de l'Assurance maladie (SNIIRAM), du programme de médicalisation des systèmes d'information (PMSI), du système d'information lié à la vaccination contre la COVID-19 (table IR_VAC_F), de la cartographie des pathologies et du référentiel pharmacie correspondant à une profondeur historique de neuf ans plus l'année en cours alimenteront cet échantillon en fenêtre roulante. Les variables concernées ont été listées dans le protocole et dans l'expression de besoins transmis dans le dossier de demande.
24. Les données du système d'information lié à la vaccination contre la COVID-19 (table IR_VAC_F) présentes dans l'échantillon ne seront toutefois pas traitées dans le cadre des études mises en œuvre dans le cadre de la présente autorisation.
25. Aux fins de respect du principe de minimisation et eu égard aux finalités des études envisagées, l'échantillon de données extraites du SNDS dans le cadre de la présente demande d'autorisation sera commun avec celui nécessaire à la mise en œuvre des traitements visés par la demande n°924315.
26. Les données dont le traitement est envisagé sont adéquates, pertinentes et limitées à ce qui est nécessaire au regard des finalités poursuivies, conformément aux dispositions de l'article 4 de la loi « informatique et libertés et du c) du 1 de l'article 5 du RGPD.

Sur les catégories d'accédants et de destinataires des données :

27. L'EMA n'accèdera pas aux données individuelles du SNDS.
28. Auront accès aux données dans le cadre de la présente autorisation :
 - les personnels habilités de la PDS pour les opérations d'extraction, de mise en forme et de stockage des données, ainsi que leur préparation et versement dans les espaces d'analyse ;
 - des personnels de la PDS, distincts des précédents, spécifiquement habilités pour l'analyse des données ;

- des organismes s'étant déclarés conformes au référentiel déterminant les critères de confidentialité, d'expertise et d'indépendance pour les laboratoires de recherche et bureaux d'études prévu par l'arrêté du 17 juillet 2017, agissant comme sous-traitants de la PDS pour la mise en œuvre de certaines études.

29. La PDS tient à jour des documents indiquant la ou les personnes compétentes en son sein pour délivrer l'habilitation à accéder aux données, la liste des personnes habilitées, leurs profils d'accès respectifs et les modalités d'attribution, de gestion et de contrôle des habilitations.

30. La qualification des personnes habilitées et leurs droits d'accès doivent être régulièrement réévalués, conformément aux modalités décrites dans la procédure d'habilitation établie par le responsable de traitement.

Sur l'opportunité du recours au mécanisme de la décision unique :

31. Outre le fait de poursuivre une même finalité, les traitements envisagés par l'EMA présentent des caractéristiques similaires concernant :

- les catégories de destinataires ;
- la nature des données traitées.

32. L'EMA envisage de mettre en œuvre, dans ce cadre, une dizaine de traitements de données du SNDS annuellement à moyen terme. A plus long terme, une vingtaine de traitements sont susceptibles d'être mis en œuvre annuellement.

33. Ces traitements relevant du régime prévu par les dispositions des articles 66, 72 et suivants de la loi « informatique et libertés », la CNIL estime opportun, au vu des éléments présentés dans le dossier de demande, d'autoriser la mise en œuvre de ces traitements dans le cadre d'une décision unique.

Sur l'information et les droits des personnes :

S'agissant des modalités d'information :

34. En application de l'article 69 de la loi « informatique et libertés » et du b) du 5 de l'article 14 du RGPD, l'obligation d'information individuelle de la personne concernée peut faire l'objet d'exceptions, notamment dans l'hypothèse où la fourniture d'une telle information exigerait des efforts disproportionnés. En pareil cas, 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.

35. En l'espèce, il sera fait exception au principe d'information individuelle des personnes et des mesures appropriées devront être mises en œuvre à travers la diffusion sur le site web du responsable de traitement :
- d'une information relative aux traitements de données du SNDS susceptibles d'être mis en œuvre dans le cadre de la présente autorisation un mois avant le tirage au sort de l'échantillon de données sur le portail de la CNAM ;
 - d'une information relative à chaque étude qui sera effectivement mise en œuvre.
36. Ces documents d'information rédigés en français et en anglais devront comporter l'ensemble des mentions prévues par l'article 14 du RGPD.

S'agissant de l'exercice des droits :

37. Les personnes concernées pourront exercer leurs droits d'accès, de rectification, de limitation, d'opposition et d'effacement à tout moment auprès de l'EMA.
38. La CNIL considère que ces modalités d'information et d'exercice des droits des personnes sont satisfaisantes au regard des dispositions du RGPD et de la loi « informatique et libertés ».

Sur la transparence :

39. Lorsque les résultats des traitements de données seront rendus publics, l'identification directe ou indirecte des personnes concernées doit être impossible, conformément à l'article 68 de la loi « informatique et libertés ».
40. Le protocole de chaque étude ainsi que les résultats de chacune d'entre elles seront publiés dans le catalogue des données de santé et des études en vie réelle du réseau des agences réglementaires européennes.
41. Ces traitements seront enregistrés dans le répertoire public mis à disposition par la PDS.
42. A l'issue du délai de validité de cette décision unique, un bilan contenant notamment la liste des analyses réalisées dans le cadre de la décision unique devra être adressé à la CNIL.

Sur la sécurité des données :

43. Le traitement envisagé comportant des données du SNDS, les mesures de sécurité doivent être conformes au référentiel de sécurité prévu par l'arrêté du 6 mai 2024 pendant toute la durée du traitement.
44. Des pseudonymes seront calculés par des méthodes à l'état de l'art lors de la mise à disposition des données du SNDS. La génération de ces pseudonymes et la gestion des tables de correspondance seront sous le contrôle de la CNAM et de la PDS.

45. La sécurité des données du SNDS dans les espaces d'ingestion, de stockage, de préparation et d'analyse dépend essentiellement de la solution technique de la PDS, qui a fait l'objet d'une analyse globale des risques et de l'impact sur la vie privée, suivie d'une homologation selon le référentiel de sécurité du SNDS en vigueur. Cette homologation, incluant les espaces d'analyse, a été réalisée par la PDS le 22 juillet 2024, pour une durée de trois ans, sous réserve de la mise en œuvre du plan d'actions qu'elle a défini.

Concernant le cycle de vie des données du SNDS :

46. En premier lieu, les données brutes du SNDS extraites depuis le portail de la CNAM seront transférées dans l'espace d'ingestion de la plateforme technologique de la PDS, où elles feront l'objet d'une pseudonymisation spécifique, puis déplacées temporairement dans la zone de stockage attaché à l'espace d'ingestion. Elles seront alors traitées afin de générer une copie au format OMOP-CDM dans un espace de préparation temporaire. Cette copie constitue « l'échantillon EMA » qui sera déplacé dans un espace de stockage spécifique au projet DARWIN, placé sous la responsabilité de l'EMA.
47. En principe, dès la fin du contrôle qualité de cet échantillon, toutes les données du SNDS mobilisées pour la génération de l'échantillon EMA devraient être supprimées de la plateforme de la PDS et de ses sauvegardes, afin de limiter leur exposition aux risques. La mise à jour des données pourrait ensuite avoir lieu selon une périodicité annuelle en suivant ces mêmes modalités.
48. Cependant, dans le cadre de la finalisation de l'instruction de la demande d'autorisation, la PDS a précisé avoir besoin d'une éventuelle conservation plus longue des données brutes extraites du SNDS, dans le cas d'une évolution du format OMOP-CDM qui nécessiterait le retraitement de l'ensemble des données.
49. La CNIL relève que ces évolutions étant nécessairement planifiées à l'avance avec la communauté OMOP-CDM, elles pourront être anticipées et prises en compte par la PDS lors de la mise à jour annuelle des données. Elle estime, dans ces conditions, que conserver en permanence deux échantillons aussi larges du SNDS au sein de sa solution technique n'est pas acceptable en termes de sécurité. L'échantillon devra donc être supprimé une fois que la copie au format OMOP-CDM aura été réalisée et contrôlée, dans un délai qui ne devrait en principe pas dépasser trois mois à compter de l'extraction du portail de la CNAM.
50. Pour faire face à la difficulté soulignée par la PDS, une nouvelle extraction des données brutes de l'intégralité de l'échantillon à l'occasion de la mise à jour pourrait, le cas échéant, être réalisée. Par ailleurs, la CNIL invite la PDS à réfléchir à d'autres solutions. Il pourrait notamment être possible de conserver l'échantillon sur un support de stockage indépendant et non connecté, en dehors de la solution technique de la PDS s'il respecte les procédures et outils homologués par la CNAM à cet égard.

51. En deuxième lieu, l'échantillon EMA sera utilisé pour produire des extractions ciblées en fonction des besoins des différentes études commanditées par l'EMA. L'accès à l'échantillon sera désactivé entre deux besoins d'extractions. L'EMA devra s'assurer que toute utilisation ou réutilisation des données de son échantillon et des données de ses études respecte les conditions de la présente autorisation.
52. Chaque étude disposera d'un espace temporaire de préparation et d'un espace d'analyse qui seront isolés de l'échantillon EMA complet, afin notamment d'assurer la minimisation des données propres à l'étude, une pseudonymisation distincte, la gestion des habilitations aux seules personnes dûment habilitées pour l'étude et une durée de conservation adaptée aux finalités de celle-ci. L'extraction ciblée sera générée dans l'espace de préparation dédié à l'étude, lequel sera désactivé dès que l'extraction aura été transmise à l'espace d'analyse. Les opérateurs techniques et référents données de la PDS, chargés des opérations jusqu'à ce stade, n'auront pas accès aux espaces d'analyse.
53. Les analyses seront conduites par des équipes spécifiques de la PDS, dans l'espace propre à chaque étude, sous la responsabilité exclusive de l'EMA et selon le protocole de chaque étude entrant dans le cadre de la présente autorisation. Le cas échéant, les sous-traitants de la PDS intervenant dans l'analyse des données devront être détaillés dans le protocole des études concernées.
54. Enfin, les résultats transmis au centre de coordination devront être fournis uniquement sous la forme d'extractions anonymes. A cet égard, la CNIL rappelle que des règles comme l'agrégation simple (k-anonymat) sont aujourd'hui insuffisantes pour assurer l'anonymisation des données dans tous les cas. Chaque export devra donc faire l'objet d'une analyse de risques en réidentification spécifique au contexte de l'étude et au jeu de données extrait.
55. Plus globalement, la CNIL invite l'EMA à mettre en place une gouvernance aux fins de s'assurer que tout traitement des données du SNDS sera conforme au cadre de la présente autorisation, au préalable, durant la conduite et à la fin de chaque étude. Cette gouvernance pourrait constituer un outil pertinent pour s'assurer du respect du protocole cadre validé par le CESREES et des termes de l'autorisation de la CNIL, et notamment du pilotage et de la surveillance des sous-traitants mobilisés pour chaque étude.

Concernant les autres mesures de sécurité :

56. Le responsable de traitement a réalisé et transmis à l'appui de la demande d'autorisation une analyse d'impact relative à la protection des données spécifique aux traitements envisagés.
57. Les droits d'accès seront accordés par l'opérateur projet de la PDS, qui effectuera une revue trimestrielle des habilitations.

58. L'EMA déléguera à la PDS l'analyse des traces d'utilisation des données et la transmission d'un tableau de bord d'indicateurs de sécurité défini conjointement. La sensibilisation des utilisateurs devra tenir compte de la surveillance des indicateurs de sécurité, afin de corriger d'éventuelles dérives de comportement et prendre en compte l'apparition de nouveaux risques.
59. Enfin, la gestion des incidents et des violations interviendra selon les procédures en vigueur au sein de l'EMA et de la PDS.
60. Les mesures de sécurité mises en œuvre par le responsable de traitement apparaissent proportionnées aux risques présentés par le traitement.

Sur les transferts de données en dehors de l'Union européenne :

61. La CNIL rappelle que les dispositions de l'article R. 1461-1 du CSP prévoient qu'aucun transfert de données à caractère personnel ne peut être réalisé en dehors de l'Union européenne, sauf dans le cas d'accès ponctuels aux données par des personnes situées en dehors de l'Union européenne, pour une finalité relevant du 1° du I de l'article L. 1461-3 du CSP.
- En l'espèce, le dossier de demande ne fait pas état de tels accès à distance par les personnels de la PDS et ses sous-traitants ultérieurs chargés de l'analyse des données du SNDS.

Sur le recours à la société Microsoft en tant qu'hébergeur :

62. La PDS a fait le choix de recourir à la société Microsoft Ireland Operations Ltd (Microsoft) en qualité d'hébergeur des données. Les données de l'entrepôt seront conservées dans les centres de données de Microsoft situés en France. En application de l'article 28 du RGPD, les sous-traitants doivent présenter des garanties suffisantes, en fonction des caractéristiques du traitement et des risques pour les personnes. Ces garanties doivent en particulier minimiser les risques de transferts illégaux ou de rupture de la confidentialité des données traitées.

S'agissant des transferts de données liés à l'administration de la plateforme :

63. Il résulte des informations communiquées que, compte tenu du contrat passé avec Microsoft et du fonctionnement des opérations d'administration de la plateforme technique, il est possible que des données techniques d'usage de la plateforme (qui ne révèlent aucune information de santé) soient transférées vers des administrateurs situés aux États-Unis.
64. Ces transferts de données vers les États-Unis sont encadrés par les clauses contractuelles types de la Commission européenne. Les personnes devront être spécifiquement informées de ces transferts.

Sur les risques d'accès par les autorités étasuniennes :

65. L'hébergeur retenu par le GIP PDS appartient à un groupe dont la société mère est située aux États-Unis et soumise au droit de cet État. De ce fait, en application de cette législation, les autorités états-uniennes sont susceptibles d'adresser à Microsoft des injonctions de communication des données qu'il héberge.
66. La législation des États-Unis, qui prévoit ces injonctions de communication de données, a été révisée, à la suite de l'arrêt de la Cour de justice de l'Union européenne du 16 juillet 2020 (C-311/18) dit « Schrems II », pour mieux les encadrer et prévoir des voies de recours plus effectives. Cette législation a été jugée conforme aux standards de protection européens, ce qui a permis l'adoption par la Commission européenne d'une nouvelle décision d'adéquation le 10 juillet 2023, reconnaissant que le nouveau cadre de transferts de données à caractère personnel entre les États-Unis et l'Union européenne, le *Data Privacy Framework*, assure un niveau de protection adéquat.
67. Il n'en reste pas moins que les données stockées par un hébergeur soumis à un droit extra-européen peuvent être exposées à un risque de communication à des puissances étrangères. Ce risque doit être apprécié au regard, d'une part, du RGPD et, d'autre part, de la loi française.
68. D'une part, s'agissant du RGPD, l'article 28 du RGPD prévoit que « *lorsqu'un traitement doit être effectué pour le compte d'un responsable du traitement, celui-ci fait uniquement appel à des sous-traitants qui présentent des garanties suffisantes quant à la mise en œuvre de mesures techniques et organisationnelles appropriées de manière à ce que le traitement réponde aux exigences du présent règlement et garantisse la protection des droits de la personne concernée* ». Il résulte de ces dispositions que le responsable de traitement doit, dans le choix de son sous-traitant et les mesures de sécurité mises en place, protéger la confidentialité des données, y compris en s'efforçant de diminuer le risque d'accès illégal par des Etats étrangers à l'Union européenne.
69. Si ce risque est souvent acceptable, notamment s'agissant de pays adéquats, la CNIL recommande, pour les bases de données les plus sensibles, d'assurer une protection contre les possibilités de divulgation à des autorités publiques de pays tiers. Cette protection implique que, d'une part, sauf exception ponctuelle (par exemple dans le cadre d'un projet de recherche international), les données hébergées dans l'Union européenne ne soient pas transférées en dehors de celle-ci et, d'autre part, de recourir à un prestataire exclusivement soumis au droit de l'Union européenne et offrant le niveau de protection adéquat, tel que prévu par le référentiel SecNumCloud de l'Agence nationale de la sécurité des systèmes d'information (ANSSI).

70. Le Conseil d'Etat a jugé, s'agissant d'un entrepôt de données de santé hébergé sur la solution technique de la PDS, que les traitements de pseudonymisation multiples des données et les garanties apportées par la société Microsoft, qui bénéficie de la certification « hébergeur de données de santé » prévue par l'article L. 1111-8 du CSP, et est soumise à ce titre à un audit régulier par un organisme accrédité, suffisaient en l'espèce à apporter des garanties suffisantes pour l'application de l'article 28 du RGPD malgré la persistance d'un risque d'accès aux données par les administrations des Etats-Unis (CE, 19 nov. 2024, Ass. Internet Society France, n° 491644, inédit).
71. En l'espèce, les mêmes garanties sont apportées par la PDS et son sous-traitant. Toutefois, l'échantillon de données concernera un plus grand nombre de personnes concernées. En considération de ces éléments et des modalités d'hébergement envisagées, la CNIL relève que l'EMA a proposé au cours de l'instruction de réduire le volume de données traitées tant s'agissant du nombre de personnes concernées que de la profondeur historique sollicitée sans que cela soit de nature à remettre en cause la pertinence scientifique des études qu'elle commande.
72. Par ailleurs, il y a un fort intérêt public à autoriser les traitements en cause, en raison de l'importance du projet DARWIN et de ce que ce projet s'inscrit dans le cadre de la préfiguration de l'espace européen des données de santé.
73. Dans ces conditions, eu égard à l'intérêt public des recherches envisagées et afin de ne pas retarder leur mise en œuvre, la CNIL estime possible d'autoriser le recours à la solution technique de la PDS, qu'elle considère comme conforme à l'article 28 du RGPD, à condition de réduire la durée demandée (*cf.* ci-dessous).
74. D'autre part, s'agissant des exigences propres à la législation française, la CNIL relève que l'article 31 de la loi n° 2024-449 du 21 mai 2024 visant à sécuriser et à réguler l'espace numérique (loi « SREN »), qui est applicable à la PDS, vise à assurer la protection des données stratégiques et sensibles sur le marché de l'informatique en nuage (*cloud*) dans les cas où ces organismes ont recours à un service de *cloud* fourni par un prestataire privé pour le traitement de données d'une sensibilité particulière dont la violation est susceptible d'engendrer une atteinte à l'ordre public, à la sécurité publique, à la santé ou à la vie des personnes ou à la protection de la propriété intellectuelle. Dans cette hypothèse, les organismes doivent veiller à ce que le service de *cloud* fourni par le prestataire privé mette en œuvre des critères de sécurité et de protection des données garantissant notamment leur protection contre tout accès par des autorités publiques d'Etats tiers non autorisé par le droit de l'Union européenne ou d'un Etat membre. La loi prévoit la possibilité de délais dérogatoires de migration pour les projets déjà engagés nécessitant le recours à un service de *cloud*.
75. Les modalités d'application de cet article devaient être précisées par un décret en Conseil d'Etat dans un délai de six mois à compter de la promulgation de la loi « SREN ». Ce décret, notifié à la Commission européenne à la fin du mois de janvier 2025 en application de la directive (UE) 2015/1535, ce qui implique un *statu quo* de trois mois, n'est pas paru à ce jour.

76. Au jour de l'adoption de la présente délibération, l'article 31 de la loi « SREN » ne s'oppose donc pas à l'autorisation des traitements envisagés. La CNIL souligne toutefois que dès que le décret d'application permettant l'entrée en vigueur de la loi « SREN » sera paru, et le cas échéant à l'expiration d'un éventuel délai dérogatoire de mise en conformité, la PDS devra, nonobstant la présente autorisation, disposer d'un hébergeur protégeant les données contre tout accès par des autorités publiques d'Etat tiers non autorisé.
77. Enfin, la CNIL réitère ses regrets que la PDS ne dispose toujours pas à ce jour d'un prestataire susceptible de répondre à ses besoins tout en protégeant les données du SNDS contre les accès des autorités publiques d'Etat tiers. Elle rappelle avoir autorisé le 21 décembre 2023 l'entrepôt « EMC2 » hébergé via la solution technique de la PDS en considération des assurances de cette dernière et du gouvernement qu'une solution, fût-elle temporaire, serait rapidement trouvée pour assurer un hébergement souverain des données du SNDS (offre intercalaire). Si le projet de stratégie nationale pour l'utilisation secondaire des données de santé prévoit un plan d'actions à ce sujet, elle relève qu'aucun appel d'offre n'a été publié en ce sens. Elle prend note du fait que le Gouvernement a indiqué qu'il est en préparation.
78. La CNIL ne peut qu'inciter à nouveau le Gouvernement et la PDS à travailler activement à la mise en place d'une solution souveraine permettant aux organismes d'avoir accès rapidement aux données du SNDS à des fins de recherche scientifique.

Sur la durée de conservation des données :

Sur la durée de conservation de l'échantillon du SNDS :

79. L'EMA a initialement sollicité une autorisation pour mettre en œuvre des traitements de données du SNDS pendant une durée de dix ans, en demandant notamment de pouvoir conserver et réutiliser l'échantillon du SNDS présenté ci-dessus pendant dix ans via la solution technique de la PDS. La CNIL relève qu'une autorisation pour six ans a ensuite été demandée, l'EMA proposant de transmettre à l'issue d'un délai de trois ans un bilan sur le fonctionnement du dispositif et l'utilisation des données de la base principale du SNDS.
80. Au regard d'une part, des réserves sur les modalités d'hébergement des données décrites ci-dessus et, d'autre part, de la possibilité pour l'EMA d'utiliser à terme l'infrastructure dédiée qui sera mise en place suite à l'entrée en application du chapitre IV du règlement sur l'espace européen des données de santé, la CNIL estime opportun de limiter la durée de la présente autorisation à trois ans à compter de la présente autorisation.
81. Les données de l'échantillon de données du SNDS mises au format OMOP-CDM pour le compte de l'EMA seront donc conservées pendant trois ans.

Sur la durée de conservation des données de chaque espace d'analyse correspondant à une recherche :

82. Dans chaque espace d'analyse, les données seront mises à disposition pendant la durée nécessaire à la réalisation des objectifs des recherches tels que décrits dans chaque protocole spécifique. Sur ce point, le dossier de demande mentionne que les résultats de la plupart des études seront publiés dans un délai de trois à douze mois à compter de leur lancement. Les données pourront ensuite être archivées pendant un an lorsque cela est nécessaire.

83. Ces durées de conservation des données n'excèdent pas celles nécessaires aux finalités pour lesquelles elles sont collectées et traitées, conformément aux dispositions du e) du 1 de l'article 5 du RGPD.

Dans ces conditions, la CNIL autorise l'Agence européenne des médicaments à mettre en œuvre des traitements de données à caractère personnel ayant pour finalité des études portant sur l'estimation d'incidence et de prévalence des pathologies dans la population générale en France dans le cadre du projet DARWIN EU pendant une durée de trois ans.

Dans l'hypothèse où, durant cette période, une migration serait effectuée afin d'être conforme à l'article 31 de la loi « SREN », la CNIL souligne, que toutes choses égales par ailleurs, il n'y aurait pas d'obstacle de principe à ce que la durée de l'autorisation soit substantiellement allongée.

La présidente,

M.-L. Denis