

<p>Comité de sécurité de l'information Chambre sécurité sociale et santé</p>
--

CSI/CSSS/24/358

DÉLIBÉRATION N° 24/172 DU 1^{ER} OCTOBRE 2024 PORTANT SUR LA COMMUNICATION DE DONNÉES À CARACTÈRE PERSONNEL PSEUDONYMISÉES RELATIVES À LA SANTÉ PAR LE CENTRUM VOOR KANKEROPSPORING (CVKO) À L'UNIVERSITAIR ZIEKENHUIS LEUVEN DANS LE CADRE D'UNE ÉTUDE RELATIVE À UN MODÈLE D'ÉVALUATION DU RISQUE DE CANCER DU SEIN

Le Comité de sécurité de l'information, chambre sécurité sociale et santé (dénommé ci-après « le Comité ») ;

Vu le règlement (UE) n° 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 la loi du 30 juillet 2018 *relative à la protection des personnes physiques à l'égard des traitements de données à caractère personnel*;

Vu la loi du 15 janvier 1990 *relative à l'institution et à l'organisation d'une Banque-carrefour de la sécurité sociale* ;

Vu la loi du 21 août 2008 *relative à l'institution et à l'organisation de la plate-forme eHealth* ;

Vu la demande de l'Universitair Ziekenhuis Leuven visant à obtenir une autorisation ;

Vu le rapport d'auditorat de la Plate-forme eHealth du 3 septembre 2024 ;

Vu le rapport de monsieur Michel Deneyer ;

Émet, après délibération, la décision suivante, le 1^{er} octobre 2024 :

I. OBJET DE LA DEMANDE

1. L'Universitair Ziekenhuis Leuven (ci-après l'UZ Leuven) introduit une demande auprès du Comité de sécurité de l'information en vue d'obtenir des données à caractère personnel pseudonymisées relatives à la santé du *centrum voor kankeropsporing* (CvKO) dans le cadre d'une étude relative à un modèle d'évaluation du risque de cancer du sein.
2. Cette étude scientifique vise à évaluer un modèle d'estimation des risques qui calcule la probabilité de développer un cancer du sein par les femmes sur la base de leurs images mammographiques. Dans un premier temps, le risque d'un cancer du sein est calculé pour les femmes âgées de 50 à 69 ans qui ont participé au programme de dépistage du cancer du sein. Étant donné que les informations longitudinales sont aussi incluses dans le modèle, c'est-à-dire les changements dans la structure du sein ou la densité dans les images mammographiques des différents examens de dépistage, il est nécessaire d'avoir recours à des données à caractère personnel pseudonymisées de sorte que les demandeurs puissent relier entre eux les examens de dépistage successifs ainsi que les informations diagnostiques. Les images mammographiques sont pseudonymisées au sein du CvKO et sont partagées avec l'UZ Leuven uniquement sous forme pseudonymisée. Les données diagnostiques et cliniques y afférentes sont codées au sein du CvKO de la même manière de sorte qu'il y ait un lien entre les images codées et les informations diagnostiques.
3. La présente demande concerne un projet de recherche pour lequel une bourse FWO-ARRS a été octroyée et auquel participe non seulement la KU Leuven, mais aussi l'université de Ljubljana (Slovénie). Les données qui sont demandées en l'espèce, ne seront pas partagées avec l'université de Ljubljana. Toutefois, les algorithmes développés par ces derniers seront aussi testés par les chercheurs de l'UZ Leuven sur les données flamandes au moyen d'une 'Federated learning approach'. L'inverse est aussi applicable: les algorithmes développés à la KU Leuven seront aussi testés sur des données de dépistage en Slovénie. Cette approche permet d'entraîner un modèle commun sans toutefois compromettre le caractère privé des données personnelles, étant donné que les données brutes ne quittent jamais le serveur local et que les chercheurs externes n'ont pas accès à la banque de données de l'UZ.

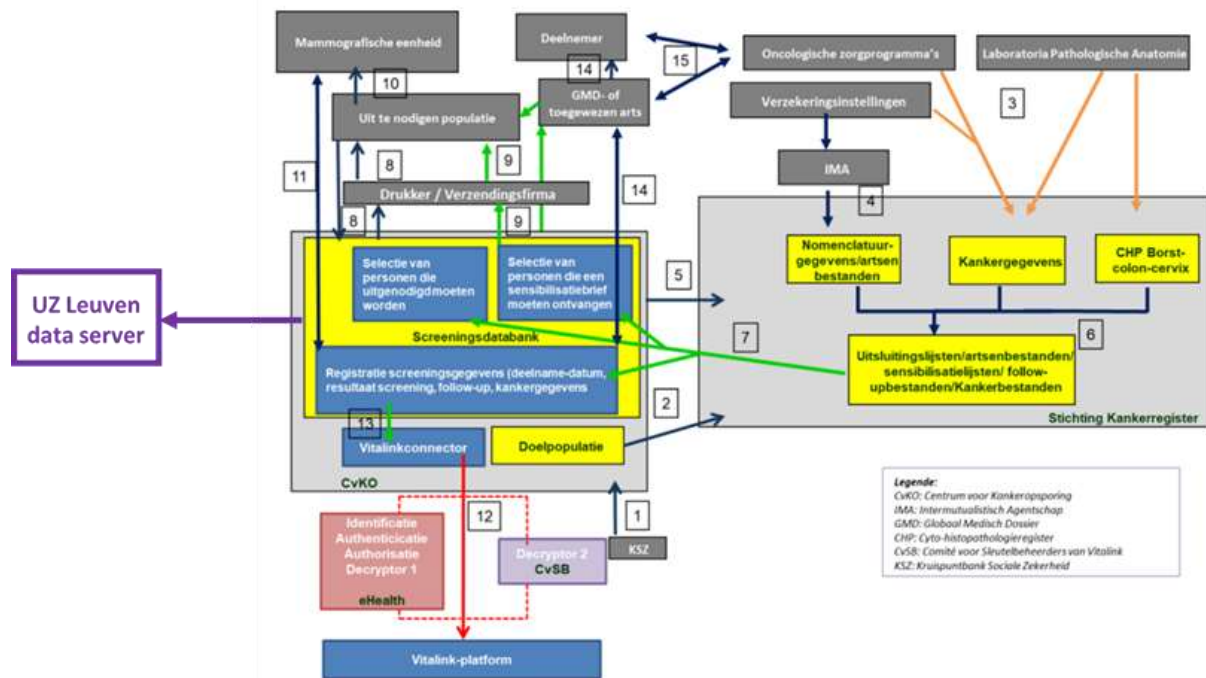
Procédure de sélection

4. La population d'étude se compose de femmes âgées de 50 à 69 ans qui participent à l'examen de dépistage du cancer du sein flamand et qui ont donné leur consentement pour l'utilisation de leurs données. Il y a trois critères de sélection, à savoir l'intervalle de temps entre 2010 et 2019, toutes les femmes participantes ayant rempli le formulaire de consentement et toutes les images mammographiques ainsi que les informations cliniques et diagnostiques de l'ensemble de ces examens de dépistage.
5. Un modèle d'évaluation du risque de cancer du sein basé sur des mammographies a déjà été présenté par Yala et al.¹ Cet algorithme a été entraîné au moyen d'environ 70.000 images provenant de systèmes de mammographie Hologic. La population étudiée comprend tant des

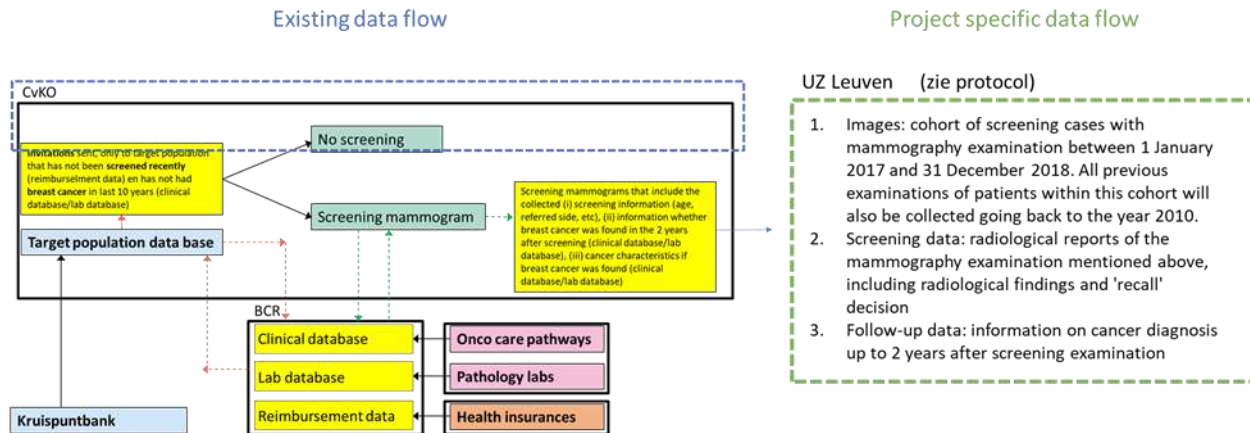
¹ A. Yala et al., "Toward robust mammography-based models for breast cancer risk," *Sci. Transl. Med.*, 2021;13(578):4373.

cancers détectés pendant le dépistage que des cancers d'intervalle. Outre Hologic, il y a en Flandre d'autres fabricants de systèmes de mammographie. Pour ce projet, les chercheurs souhaitent créer un algorithme qui s'applique aux systèmes de mammographie de l'ensemble des fabricants. Sur la base des données de Yala et al., les demandeurs estiment qu'ils ont besoin de données d'environ 700.000 femmes. Il ressort d'une autre étude² qu'environ 200.000 femmes participent annuellement à l'examen de dépistage du cancer du sein en Flandre. Pour obtenir le nombre de 700.000 femmes, les demandeurs incluront des données de la période prenant cours au moment du lancement de la mammographie numérique dans toute la Flandre (2010) jusqu'à l'année la plus récente pour laquelle l'ensemble des données cliniques et diagnostiques et de suivi pendant 5 ans sont disponibles (2019). Cela implique donc que les informations *ground truth*, à savoir le fait qu'un cancer du sein a été diagnostiqué ou non 5 ans après l'examen de dépistage, sont déjà présentes en 2024.

6. Le flux de données se présente comme suit:



² M. Goossens et al., “Flemish breast cancer screening programme: 15 years of key performance indicators (2002–2016),” *BMC Cancer*, 2019;19(1):1012.



Outre les flux de données existant déjà entre le CvKO et la BCR et les différents acteurs impliqués par les différents examens de dépistage, une branche supplémentaire au flux de données sera ajoutée, ce qui permettra à l'UZ Leuven d'envoyer des données sous la forme d'images mammographiques codées et des informations cliniques et diagnostiques provenant de la banque de données des examens de dépistage du CVKO (flèche mauve sur la figure ci-dessus).

II. COMPÉTENCE

7. Conformément à l'article 42, § 2, 3°, de la loi du 13 décembre 2006 *portant dispositions diverses en matière de santé*, toute communication de données à caractère personnel relatives à la santé requiert, sauf les exceptions prévues, une délibération de principe de la chambre sécurité sociale et santé du Comité de sécurité de l'information.
8. Le Comité estime dès lors qu'il est compétent pour se prononcer sur cette communication de données à caractère personnel relatives à la santé.

III. EXAMEN DE LA DEMANDE

A. ADMISSIBILITÉ

9. Le traitement de données à caractère personnel relatives à la santé est en principe interdit, et ce conformément au prescrit de l'article 9, §1^{er} du 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 (RGPD).
10. L'interdiction du traitement de données à caractère personnel relatives à la santé ne s'applique pas lorsque le traitement est nécessaire à des fins archivistiques dans l'intérêt public, à des fins de recherche scientifique ou historique ou à des fins statistiques, conformément à l'article 89, paragraphe 1, sur la base du droit de l'Union ou du droit d'un État membre, qui doit être proportionné à l'objectif poursuivi, respecter l'essence du droit à

la protection des données et prévoir des mesures appropriées et spécifiques pour la sauvegarde des droits fondamentaux et des intérêts de la personne concernée³.

11. Le Comité prend acte du fait que le Comité éthique de l'UZ Leuven a formulé un avis positif pour cette étude.
12. A la lumière de ce qui précède, le Comité est d'avis qu'il existe un fondement admissible pour le traitement des données à caractère personnel relatives à la santé envisagé.

B. PRINCIPES RELATIFS AU TRAITEMENT DES DONNÉES À CARACTÈRE PERSONNEL

1. FINALITÉS

13. En vertu de l'article 5 du RGPD, les données à caractère personnel doivent être traitées de manière licite, loyale et transparente au regard de la personne concernée. Elles doivent être collectées pour des finalités déterminées, explicites et légitimes, et ne pas être traitées ultérieurement d'une manière incompatible avec ces finalités.
14. La présente étude vise à évaluer le modèle d'estimation des risques qui calcule la probabilité de développer un cancer du sein par les femmes sur la base de leurs images mammographiques. Les résultats permettront de contribuer à une meilleure prédiction du risque de cancer du sein afin de finalement améliorer l'examen de dépistage du cancer du sein flamand au moyen d'une procédure de stratification objective, basée sur des risques, sur la base d'un profil de risque de cancer du sein individuel. Les résultats de cette étude permettent de préciser le bilan des avantages et inconvénients de certaines modalités de dépistage supplémentaires ou d'intervalles de dépistage adaptés.
15. La présente demande concerne un projet de recherche pour lequel une bourse FWO-ARRS a été octroyée et auquel participe non seulement la KU Leuven, mais aussi l'université de Ljubljana (Slovénie). Les données qui sont demandées en l'espèce, ne seront pas partagées avec l'université de Ljubljana. Toutefois, les algorithmes développés par ces derniers seront aussi testés par les chercheurs de l'UZ Leuven sur les données flamandes au moyen d'une 'Federated learning approach'. L'inverse est aussi applicable: les algorithmes développés à la KU Leuven seront aussi testés sur des données de dépistage en Slovénie. Cette approche permet d'entraîner un modèle commun sans toutefois compromettre le caractère privé des données à caractère personnel, étant donné que les données brutes ne quittent jamais le serveur local et que les chercheurs externes n'ont pas accès à la banque de données de l'UZ.
16. Étant donné que la présente étude vise à évaluer le risque de développer un cancer du sein par la population du dépistage, l'établissement de ce modèle requiert un très grand nombre de données. D'une part, la population des femmes qui n'ont à ce jour pas encore souffert d'un cancer du sein, d'autre part, la population des femmes ayant reçu le diagnostic d'un cancer du sein (risque de la population versus risque accru). Outre le grand nombre de données qui sont nécessaires à l'établissement du modèle, les demandeurs souhaitent faire

³ Art. 9, alinéa 2, j) du RGPD.

en sorte que le modèle soit suffisamment performant pour chacun des fabricants de mammographie auxquels l'examen de dépistage du cancer du sein flamand a recours. En outre, les chercheurs souhaitent aussi réaliser une analyse des effets fixes par centre de dépistage. L'ensemble de données demandé a été établi de la sorte que les demandeurs ne demandent pas davantage de variables ou de données que celles nécessaires pour la présente étude.

17. Compte tenu des finalités du traitement décrites ci-avant, le Comité de sécurité de l'information considère que le traitement des données à caractère personnel envisagé poursuit bien des finalités déterminées, explicites et légitimes.

2. MINIMISATION DES DONNÉES

18. Selon l'article 5 du RGPD, les données à caractère personnel doivent être adéquates, pertinentes et limitées à ce qui est nécessaire au regard des finalités pour lesquelles elles sont traitées.
19. Il s'agit d'un traitement ultérieur. La présente étude concerne l'usage secondaire de données collectées par le CvKO dans le cadre de l'examen de dépistage du cancer du sein flamand. Par ailleurs, les femmes ont donné leur consentement pour l'utilisation de leurs données à des fins scientifiques dont fait partie ce projet.
20. Il s'agit d'une consultation ponctuelle des données. Toutes les images et informations diagnostiques sont demandées à titre unique, étant donné que l'ensemble de données complet est nécessaire pour mettre au point, entraîner et valider le modèle de risques.
21. Les images mammographiques sont pseudonymisées au sein du CvKO et sont uniquement partagées sous forme pseudonymisée avec l'UZ Leuven. Les données cliniques et diagnostiques y afférentes sont codées de la même manière au sein du CvKO de sorte qu'il y ait un lien entre les images codées et les informations diagnostiques.
22. Aucun lien avec l'individu n'est conservé, ni au sein du CvKO, ni de l'UZ Leuven.
23. Les images mammographiques DICOM ne contiennent pas de données d'identification du patient dans les images mêmes; ces données sont disponibles dans le dicomheader/les métadonnées. La pseudonymisation des images mammographiques implique donc la suppression de ces données dans le header. Ceci se fera à l'aide du RSNA DICOM Anonymizer. Il s'agit d'un logiciel qui est gratuitement mis à la disposition par la Radiological Society of North America.⁴ Le logiciel est installé localement et celui-ci supprime ou remplace les données du patient dans les métadonnées des images par un code. Le document contenant le lien entre ce code et l'identification du patient peut être conservé et reste au CvKO et n'est pas partagé avec l'UZ Leuven. Les chercheurs ne sont pas en mesure de retrouver l'identité de la personne.

⁴ <https://www.rsna.org/research/imaging-research-tools>.

24. Les images des femmes ayant subi un examen de dépistage à l'UZ Leuven sont aussi disponibles dans les archives cliniques (PACS de l'hôpital) dans lesquelles l'identification du patient est disponible. Il s'agit uniquement d'un nombre limité de mammographies de dépistage flamands. Pour la présente étude, les chercheurs utilisent un autre serveur pour l'archivage et l'utilisation de données mammographiques pseudonymisées. Il n'y a donc pas d'interaction avec le PACS de l'hôpital. Par ailleurs, les demandeurs demandent uniquement la date exprimée en mois/années d'acquisition des images et ne conservent pas la date exacte des images mammographiques. Ce qui complexifie fortement la recherche d'une image spécifique dans le PACS. L'accès au système PACS de l'hôpital sera interrompu pour les chercheurs pendant la durée du présent projet.
25. Une liste des données demandées ainsi qu'une motivation circonstanciée par variable sont jointes ci-après.

3. LIMITATION DE LA CONSERVATION

26. Selon l'article 5, §1^{er}, e) du RGPD, les données à caractère personnel doivent être conservées sous une forme permettant l'identification des personnes concernées pendant une durée n'excédant pas celle nécessaire au regard des finalités pour lesquelles elles sont traitées; les données à caractère personnel peuvent être conservées pour des durées plus longues dans la mesure où elles seront traitées exclusivement à des fins archivistiques dans l'intérêt public, à des fins de recherche scientifique ou historique ou à des fins statistiques conformément à l'article 89, §1^{er}, pour autant que soient mises en œuvre les mesures techniques et organisationnelles appropriées requises par le règlement afin de garantir les droits et libertés de la personne concernée (limitation de la conservation).
27. En ce qui concerne le délai de conservation des images et des informations diagnostiques, un délai de 5 ans à compter de l'accès aux données est prévu. Étant donné que le présent projet de recherche constitue une première étude des modèles de risque de développer un cancer du sein et que seule une première version de ces algorithmes sera développée, les demandeurs s'attendent, à bref délai, à ce qu'ils doivent aussi réaliser des optimisations, une validation et des contrôles de qualité sur ces algorithmes. Par ailleurs, les demandeurs prévoient qu'après 5 ans, la qualité des images mammographiques aura tellement changé que les images mammographiques actuelles ne seront plus utiles pour tester des algorithmes nouveaux ou optimisés de prédiction du risque de développer un cancer du sein.
28. Le Comité estime que le délai de conservation est raisonnable.

4. TRANSPARENCE

29. Lorsque les données à caractère personnel n'ont pas été collectées auprès de la personne concernée, le responsable du traitement est tenu de fournir les informations visées à l'article 14, §§1 et 2 du RGPD à la personne concernée.
30. Cette obligation ne s'applique cependant pas lorsque la fourniture de telles informations se révèle impossible ou exigerait des efforts disproportionnés, en particulier pour le traitement

à des fins archivistiques dans l'intérêt public, à des fins de recherche scientifique ou historique ou à des fins statistiques sous réserve des conditions et garanties visées à l'article 89, alinéa 1^{er}, ou dans la mesure où l'obligation visée à l'alinéa 1^{er} du présent article est susceptible de rendre impossible ou de compromettre gravement la réalisation des objectifs dudit traitement. En pareils cas, le responsable du 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.

31. Il s'agit d'une très grande population et donc l'envoi d'une lettre d'information à toutes les femmes individuelles exigerait des efforts disproportionnés ou signifierait un coût trop élevé. Les demandeurs préfèrent ne pas informer les femmes sur leur profil de risque de développer un cancer du sein obtenu au cours de la présente étude. Ceci n'est pas non plus possible étant donné que l'UZ Leuven n'aura pas accès aux données à caractère personnel (nom et adresse). Par ailleurs, la période de 5 ans concernée par l'estimation des risques est déjà échue, étant donné que les données les plus récentes qui seront collectées datent de 2019. D'autre part, le partage de ce profil de risque inquiéterait inutilement les femmes. De plus, aucun examen de dépistage du cancer du sein adapté pour la population générale n'a dans l'intervalle été élaboré sur la base du profil de risque des femmes auquel les femmes pourraient participer. Ce projet de recherche établira un profil de risque sur la base des images mammographiques. Cependant, les facteurs de risques déjà connus tels des seins extrêmement denses, une prédisposition familiale, des mutations génétiques sont maintenus et ces femmes peuvent avoir recours aux dispositifs de dépistage diagnostique dans les hôpitaux selon le standard des soins.
32. Le Comité est d'avis qu'il existe suffisamment de transparence quant au traitement envisagé.

5. SÉCURITÉ DE L'INFORMATION

33. Selon l'article 5, §1^{er}, f) du RGPD, les données à caractère personnel doivent être traitées de façon à garantir une sécurité appropriée des données à caractère personnel, y compris la protection contre le traitement non autorisé ou illicite et contre la perte, la destruction ou les dégâts d'origine accidentelle, à l'aide de mesures techniques ou organisationnelles appropriées (intégrité et confidentialité).
34. Le Comité constate qu'une analyse d'impact relative à la protection des données a eu lieu, conformément à l'article 35 du RGPD, et qu'elle lui a été transmise.
35. Le Comité constate qu'aucune analyse de risque *small cell* sera réalisée. Les demandeurs estiment que ceci n'est pas nécessaire étant donné que les demandeurs ont, pour cette étude, recours à de grands nombres d'images mammographiques et d'informations diagnostiques y afférentes, et ce pour les groupes de femmes sans et avec un diagnostic de cancer du sein. La stratification des images a lieu sur la base de la catégorie d'âge, du système ou du fabricant de mammographies et du centre de dépistage. Le nombre de femmes par cellule est donc encore très grand. La stratification n'a jamais lieu sur base de la femme individuelle/au niveau du patient.

36. Il constate en outre que l'UZ Leuven a désigné un médecin responsable du traitement des données à caractère personnel relatives à la santé, ainsi qu'un délégué à la protection des données.
37. Le Comité constate que les collaborateurs de l'UZ Leuven ont un devoir de confidentialité légal⁵ et contractuel par rapport aux données qu'ils traitent dans le cadre de leur fonction.
38. Le Comité rappelle que ni l'UZ Leuven, ni ses collaborateurs ne peuvent entreprendre des démarches pour réidentifier les personnes concernées. Les résultats de l'étude doivent être publiés sous forme anonyme.
39. Le Comité rappelle qu'en vertu de l'article 9 de la loi du 30 juillet 2018 *relative à la protection des personnes physiques à l'égard des traitements de données à caractère personnel*, le responsable du traitement prend les mesures supplémentaires suivantes lors du traitement de données génétiques, biométriques ou des données concernant la santé :
- 1° les catégories de personnes ayant accès aux données à caractère personnel, sont désignées par le responsable du traitement ou, le cas échéant, par le sous-traitant, avec une description précise de leur fonction par rapport au traitement des données visées;
- 2° la liste des catégories des personnes ainsi désignées est tenue à la disposition de l'autorité de contrôle compétente par le responsable du traitement ou, le cas échéant, par le sous-traitant;
- 3° il veille à ce que les personnes désignées soient tenues, par une obligation légale ou statutaire, ou par une disposition contractuelle équivalente, au respect du caractère confidentiel des données visées.
40. Le Comité estime nécessaire de rappeler que depuis le 25 mai 2018, l'UZ Leuven est tenu de respecter les dispositions et les principes du 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). Ces instances sont également tenues de respecter les dispositions de la loi du 30 juillet 2018 *relative à la protection des personnes physiques à l'égard des traitements de données à caractère personnel*.

⁵ Article 458 du Code pénal.

Par ces motifs,

la chambre sécurité sociale et santé du comité de sécurité de l'information

conclut que :

la communication des données à caractère personnel telle que décrite dans la présente délibération est autorisée moyennant le respect des mesures de protection des données qui ont été définies, en particulier les mesures en matière de limitation de la finalité, de minimisation des données, de limitation de la durée de conservation des données et de sécurité de l'information.

La présente délibération entre en vigueur le 16 octobre 2024.

Michel DENEYER
Président

Le siège de la chambre sécurité sociale et de la santé du comité de sécurité de l'information est établi dans les bureaux de la Banque Carrefour de la sécurité sociale, à l'adresse suivante: Quai de Willebroeck 38 - 1000 Bruxelles (tél. 32-2-741 83 11).

Bijlage 2: lijst van gegevens met verantwoording

Screening informatie

Naam	Type	Herkomst	Uitleg
Anoniem patient ID	String	CvKO dataset	Longitudinale opvolging van screeningsbeelden en kankerdetectie informatie van de vrouwen vereist codering tot op individueel niveau
Leeftijd	Integer	CvKO dataset	Belangrijke inputfactor voor borstkankerrisico bepaling
Onderzoeksdatum (maand-jaar)	Date format	CvKO dataset	Opdat de tijd (in maanden) tussen het mammografisch beeld en kankerdetectie gekend zou zijn.
Fabrikant en model	String	DICOM (doet Leuven zelf, moet dus niet uit DICOM gehaald bij anonimisatie)	Om een risicomodel op te stellen dat relevant is voor de verschillende mammografie fabrikanten aanwezig in screening
Software versie mammografie systeem	String	DICOM (doet Leuven zelf, moet dus niet uit DICOM gehaald bij anonimisatie)	Om een fixed effect analyse uit te voeren zodat invloed van systeem en systeemversie kan onderzocht worden
Projectierichting	CC of MLO	DICOM (doet Leuven zelf, moet dus niet uit DICOM gehaald bij anonimisatie)	Het model analyseert beide projectierichtingen.
compressiekracht	Float	DICOM (doet Leuven zelf, moet dus niet uit DICOM gehaald bij anonimisatie)	Als kwaliteitsparameter voor de beeldacquisitie
Gecomprimeerde borstdikte	Integer	DICOM (doet Leuven zelf, moet dus niet uit DICOM gehaald bij anonimisatie)	In mm
Technische exposie parameters	String	DICOM (doet Leuven zelf, moet dus niet uit DICOM gehaald bij anonimisatie)	Anode/filter, kV, mAs, Organ dose
Lateraliteit doorverwijzing	L of R	CvKO dataset	Opdat lateraliteit van het beeld gekend is zodat bepaalde regio's die leiden tot een verhoogd risico schatting kunnen gelinkt worden aan eventuele kankerdetectie in dezelfde borst (explainable AI)
Code voor mammografische eenheid	String	CvKO dataset	Codenummer voor de mammografische eenheden om fixed effect analyse uit te voeren zodat invloed van positionering of technologenteam kan uitgesloten worden
BI-RADS borstdensiteit categorie (visueel bepaald door radioloog tijdens screening)	A, B, C of D	CvKO dataset	Reeds gekende onafhankelijk risicofactor voor borstkanker
Reden voor weerhouden van screeningsonderzoek	String	CvKO dataset	Weerhouden omwille van technische problemen, microcalcificaties, opaciteit of distortie.

Follow-up informatie

Naam	Type	Herkomst	Uitleg
Anoniem patient ID	String	BCR dataset van CvKO	Om een link met de mammografische beelden mogelijk te maken
Borstkanker diagnose datum (maand-jaar)	Date format	BCR dataset van CvKO	Opdat de tijd (in maanden) tussen het mammografisch beeld en kankerdetectie gekend zou zijn.
Kanker type	Invasief of in situ	BCR dataset van CvKO	Subanalyse (mogelijks andere performantie of relevantie van borstkankerrisico predictie model voor beide subgroepen). Om vergelijking met reeds gepubliceerde performantie indicatoren voor beide subgroepen mogelijk te maken.
Lateraliteit kanker	L, R of bilateraal	BCR dataset van CvKO	Opdat lateraliteit van het beeld gekend is zodat bepaalde regio's die leiden tot een verhoogd risico schatting kunnen gelinkt worden aan eventuele kankerdetectie in dezelfde borst (explainable AI)
kankerdetectietype	Screening of interval	BCR dataset van CvKO	Subanalyse (mogelijks andere performantie of relevantie van borstkankerrisico predictie model voor beide subgroepen). Om vergelijking met reeds gepubliceerde performantie indicatoren voor beide subgroepen mogelijk te maken.