GUÍA DE DISOCIACIÓN Y ANONIMIZACIÓN DE DATOS PERSONALES EN EL ÁMBITO DE LA SALUD

Montevideo, diciembre de 2017

Índice

1. Introducción

El impacto de las Tecnologías de la Información constituye una realidad a la cual no es ajeno el ámbito de la salud y se manifiesta a través de nuevas herramientas que tienen como objetivo mejorar la calidad de la atención de los pacientes.

En general, sus principales manifestaciones están dadas por la historia clínica electrónica, la gestión administrativa electrónica de los pacientes, la receta médica electrónica, la gestión electrónica de imágenes.

Es importante resaltar que la información de salud recabada en este escenario puede ser utilizada en determinados supuestos para fines estadísticos o de investigación siempre y cuando se apliquen técnicas que permitan la disociación de la información.

El objetivo de esta guía es brindar al sector asistencial herramientas que permitan la utilización de la información con fines de investigación médica respetando la normativa vigente y garantizando los derechos de todos los implicados.

2. Marco normativo

Como punto de partida debemos tener en cuenta que la información de los pacientes recabada en el ámbito de la salud se encuentra amparada por un conjunto normativo orientado a brindar las mayores garantías en su uso y acceso, siendo la disociación una herramienta legalmente reconocida a los efectos de posibilitar el uso de la información sin vulnerar la Protección de Datos Personales.

En lo que refiere a la investigación médica el Decreto N° 379/008 de 4 de agosto de 2008 resalta la necesidad de proteger en forma integral a los seres humanos sujetos de una investigación, con especial consideración de su dignidad e integridad. Además, destaca la libertad para llevar a cabo dichas investigaciones siempre y cuando se respeten los derechos y libertades esenciales que emanan de la

personalidad humana y se hallan reconocidos en la Constitución de la República y en los Tratados Internacionales ratificados por Uruguay.

La Ley N° 18.331 de 11 de agosto de 2008 en su artículo 4° literal g) define a la disociación como todo tratamiento de datos personales de manera que la información obtenida no pueda vincularse a persona determinada o determinable, lo cual cobra especial importancia en el ámbito de la salud si se pretende utilizar la información con fines estadísticos o para estudios epidemiológicos.

Debe tenerse presente que de principio y con carácter general la Historia Clínica es reservada y solamente pueden acceder a ella el personal de salud, personal administrativo, el paciente o en su caso la familia y el Ministerio de Salud Pública cuando lo considere pertinente. Debe remarcarse que la reserva es un principio general que se aplica a cualquier tratamiento de datos personales, debiéndose recabar el consentimiento del titular del dato en forma previa, expresa e informada, documentándose por escrito (artículo 18 de la Ley N° 18.331). Excepcionalmente estos datos pueden tratarse sin consentimiento por razones de salud e higiene públicas, de emergencia, o en forma disociada para la realización de estudios epidemiológicos.

El personal de la institución de salud que acceda a información personal en el ejercicio de sus funciones, se encuentra obligado a utilizarla respetando el principio de reserva y con la finalidad específica para la cual fue recabada, prohibiéndose la difusión a terceros.

Se debe diferenciar el consentimiento para el tratamiento de los datos de salud de aquel dado para el tratamiento médico en el ámbito de la consulta o investigación.

Los datos relativos a la salud de los individuos son datos personales especialmente protegidos conforme a lo dispuesto por el artículo 4° literal d) de la Ley N° 18.331, la cual define a los datos de salud como datos sensibles, dotándoles de una protección especial. En éstos quedan incluidas todas aquellas informaciones concernientes a

la salud pasada, presente, y futura, física o mental, de una persona.

Otro punto a considerar es la previsión efectuada en el artículo 19 de la Ley N° 18.331. Los establecimientos sanitarios, públicos y privados, y los profesionales vinculados a las ciencias de la salud, pueden recolectar y tratar los datos personales relativos a la salud física o mental de los pacientes que acudan a éstos o que estén o hubieren estado bajo tratamiento de aquéllos, respetando los principios del secreto profesional, la normativa específica y lo establecido en la ley.

En especial deberán respetarse los principios consagrados en la ley mencionada de protección de datos personales, en los que se incluyen el principio de legalidad, veracidad o exactitud de los datos, finalidad, previo consentimiento informado, seguridad de los datos, reserva y responsabilidad.

Debemos tener en cuenta que la protección de datos personales no es contraria a la utilización de la información, puesto que existen criterios de disociación que pueden ser implementados a los efectos de eliminar aquella información sensible que afecta a personas u organizaciones y cuya identidad debe protegerse legalmente, permitiendo su utilización. De esta manera, previo a publicar cualquier información que sea sensible, deben aplicarse las recomendaciones realizadas por la Unidad Reguladora y de Control de Datos Personales (URCDP) por Resolución N° 68/2017 de 26 de abril de 2017.

Se adiciona como medida la consideración de otras fuentes de información disponibles y que por combinación puedan presentar algún riesgo por su cruzamiento (ya que si bien la información de una base de datos por sí sola no genera un riesgo en la identificación de la persona, al combinarse con otra información, se vuelve identificable).

3. Consideraciones generales

En la Ley N° 18.331 (artículo 17) se prevé que los datos personales objeto de tratamiento sólo podrán ser comunicados para el cumplimiento de los fines directamente relacionados con el interés legítimo del emisor y del destinatario y con el previo consentimiento del titular de los datos, al que se le debe informar sobre la finalidad de la comunicación e identificar al destinatario o los elementos que permitan hacerlo. Finaliza el artículo estableciendo que el previo consentimiento para la comunicación es revocable.

En cuanto a los datos de salud, el mismo artículo 17 (literales c y d) establece como excepción, que no será necesario el consentimiento para su comunicación por razones sanitarias, de emergencia o para la realización de estudios epidemiológicos, preservando la identidad de los titulares de los datos mediante mecanismos de disociación adecuados (cuando ello sea pertinente). Establece además que se podrán comunicar si se hubiera aplicado un procedimiento de disociación de la información, de modo que los titulares de los datos no sean identificables.

A los efectos de comprender a cabalidad el alcance de los términos referidos la URCDP define la disociación como "todo tratamiento de datos personales de manera que la información obtenida no pueda vincularse a persona determinada o determinable". Siguiendo el Dictamen Nº 05/2014, de 10 de abril de 2014, adoptado por el Grupo de Trabajo del Artículo 29 de la Directiva 95/46/CE define la anonimización como "el resultado de un tratamiento de los datos personales realizado para evitar de forma irreversible su identificación" y la seudonimización como aquella operación que "reduce el vínculo de un conjunto de datos con la identidad original del interesado", adiciona además a esta última la definición dada por Reglamento (UE) Nº 679/2016 del Parlamento Europeo y del Consejo relativo a la protección de las personas físicas en lo que respecta al tratamiento de datos personales y a la libre circulación de estos datos (Reglamento General de Protección de Datos), que entiende a la seudonimización como "tratamiento de

datos personales de manera tal que ya no puedan atribuirse a un interesado en particular sin recurrir a información adicional, siempre que dicha información adicional se mantenga separada y sujeta a medidas técnicas y organizativas destinadas a garantizar que los datos no se atribuyan a personas identificadas o identificables".

4. Procedimientos y técnicas de disociación

Conforme lo dispuesto por la URCDP a los efectos de desarrollar una correcta disociación y llegar a anonimizar los datos de salud, se deben implementar ciertas medidas como buena práctica. La anonimización implica obtener como resultado de los procedimientos o técnicas de disociación, la imposibilidad de identificar o re identificar en forma irreversible al titular del dato.

4.1. Procedimiento

El responsable de la base de datos deberá seguir tres etapas denominadas preanonimización, anonimización y control.

A. Durante la primera etapa, denominada "etapa de preanonimización" se debe diseñar el proyecto de anonimización, que permitirá identificar con claridad qué información de salud podría llegar a utilizarse en cumplimiento de la normativa antes señalada. Se deberán identificar:

   i. Las variables. Por ejemplo nombres, números de teléfono, correos electrónicos, fotografías e imágenes similares, datos biométricos, dirección, domicilio, celulares y sus números de serie o cualquier otro número de identificación.

ii. Identificadores directos e indirectos.
iii. Datos confidenciales.
iv. Las técnicas adecuadas de anonimización, según el conjunto de datos de que se trate.

Además, deberá determinarse el riesgo de reidentificación asociado.

Se planifica cuáles son las técnicas adecuadas que permitan la ruptura de los eslabones de la relación de identificación-información.

En esta etapa también se deben definir los perfiles necesarios de los actores, que llevarán a cabo el proceso de anonimización y la responsabilidad de cada uno de ellos (técnicos informáticos, usuarios, entre otros) y trazar un plan de contingencia para el caso que se verifique riesgo de o una reidentificación del titular del dato.

Esta etapa finaliza al iniciar la etapa de ejecución del proyecto denominada etapa de anonimización.

B. Durante la "etapa de anonimización" se pondrá en práctica el proyecto efectuado en la primera etapa. Si se aplican correctamente las técnicas y se llega al punto máximo (anonimización de los datos) se habrá conseguido la ruptura de la cadena identificación-información y no será posible la reversión al estado anterior, es decir, impedir que se vuelva atrás con lo efectuado y se identifique al interesado, en la

práctica equivale al borrado permanente. Debe tomarse en consideración que esto no implica borrar la información que se encuentra en una Historia Clínica sino que implica que el conjunto de información que será utilizada por ejemplo para investigación con fines epidemiológicos, se entregará al investigador luego de aplicársele las técnicas de disociación como muestra parcial. Esto se vincula directamente con la definición de perfiles referida en la primera etapa.

Tal como señala la URCDP en los criterios de disociación referidos, siguiendo el plan elaborado por el responsable del proceso de anonimización, durante la primera etapa, "los técnicos deberán aplicar las técnicas seleccionadas, los algoritmos necesarios, realizar pruebas de calidad y entregar el resultado al responsable para su aprobación. El objetivo final de la anonimización es proveer los datos desagregados para que el público en general pueda utilizarlos, sin generar conflictos con los titulares de los datos".

C. Durante la tercera etapa del proceso, denominada "etapa de control" se deberán efectuar controles periódicos por parte de los técnicos en contraposición con la aparición de las nuevas tecnologías y métodos para prevenir y evitar los posibles riesgos de reidentificación. Implica además la adopción de las medidas necesarias, según el plan de contingencia proyectado, en el caso de verificarse un riesgo o en caso que el titular del dato informe su posible reidentificación. Este es un ciclo continuo que debe verificarse en forma periódica por los avances tecnológicos constantes.

4.2. Técnicas de anonimización

A continuación se enumeran algunos de los posibles conjuntos de técnicas de anonimización a ser utilizadas por los responsables del proceso y técnicos implicados, según lo indica la URCDP en los criterios referidos.

Estos conjuntos de técnicas se dividen en generales y particulares. Dentro de los generales encontramos la aleatorización y la generalización. A su vez, en particular, la aleatorización se subdivide en adición de ruido, permutación, y privacidad diferencial; y la generalización se subdivide en agregación, anonimato k, diversidad l y proximidad t.

4.2.1. Aleatorización

La aleatorización consiste en un conjunto de técnicas que tienen como objetivo modificar la veracidad del dato con la finalidad de suprimir el vínculo que existe entre aquél y el titular. Al volver los datos lo suficientemente ambiguos, esto evitará que se identifique a una persona concreta. Dentro de la aleatorización encontramos la adición de ruido, permutación y privacidad diferencial.

4.2.1.1. Adición de ruido

Esta técnica modifica los atributos del conjunto de datos para que sean menos exactos, pero conserva su distribución general. Los datos son ciertos hasta determinado punto, por ejemplo en el caso del atributo altura donde se cuenta con un sujeto de 1,60 centímetros se

introducirán valores + 10 centímetros, es decir entre 1,50 a 1,70 centímetros en la tabla.

4.2.1.2. Permutación

La URCDP indica que: "La técnica de permutación consiste en mezclar los valores de los atributos en una tabla para que algunos de ellos puedan vincularse artificialmente a distintos interesados", es decir, se intercambian algunos valores contenidos en un conjunto de datos, con los de otro registro, cuidando no romper la relación lógica.

4.2.1.3. Privacidad diferencial

En cuanto a la privacidad diferencial si bien se encuentra dentro de las técnicas de anonimización por aleatorización, es diferente, ya que se recogen datos del global de usuarios sin saber a quién corresponde cada dato, es decir, el responsable del tratamiento de datos genera vistas anonimizadas del conjunto de datos, pero al mismo tiempo almacena copia de los originales.

4.2.2. Generalización

Es la segunda familia de técnicas de anonimización que detalla la URCDP. A través de este enfoque se generaliza o se diluyen los atributos de los interesados. Se modifican las respectivas escalas u órdenes de magnitud, así por ejemplo se puede sustituir una ciudad por una región, o una semana por un mes, etc.). La generalización puede ser efectiva para descartar la singularización, pero no permite obtener una anonimización eficaz para todos los casos; siendo es necesario aplicar otros enfoques que impidan la vinculabilidad y la inferencia.

4.2.2.1. Agregación y Anonimato k

Estas técnicas tienen por objetivo impedir que un interesado sea singularizado cuando se le agrupa junto con, al menos, un número k de personas. La URCDP ejemplifica esta técnica de la siguiente forma: "cuando se toma un atributo que equivale a la edad de los funcionarios, formando grupos de intervalos de valores, es decir entre 30 a 40 años, entre 40 y 50, haciendo franjas". Introduce además dos métodos que son aplicables a estas técnicas, a saber, supresión y generalización. En el primero se reemplaza algún atributo por asteriscos y en el segundo los valores individuales de algún atributo son reemplazados por categorías más amplias.

4.2.2.2. Diversidad l y Proximidad t

Tal como lo indica la URCDP la diversidad l extiende el anonimato k para garantizar que no se puedan realizar ataques por inferencia deterministas. Ello se logra previniendo que en cada clase de equivalencia, todos los atributos tengan al menos l valores diferentes. La proximidad t perfecciona la técnica de la diversidad l. En la primera se crean clases equivalentes que sean parecidas a la distribución inicial de los atributos en la tabla. Es una técnica útil para el caso que haya que conservar los datos lo más próximo posible a los originales. Se añade una nueva restricción a la clase de equivalencia, pues ya no basta que exista al menos l valores diferentes en cada clase, sino que, al mismo tiempo, cada valor debe representarse tantas veces como sea necesario a fin de reflejar la distribución inicial de cada atributo.

5. Consideraciones finales

En conclusión, debe tenerse presente que el proceso de anonimización implica una constante labor de control y revisión ante la aparición de nuevas tecnologías que pongan en riesgo la reidentificación de los titulares de los datos.

Además es importante evitar la confusión en la metodología, es decir, tomar en cuenta que uno de los principales riesgos es confundir y pensar que la seudonimización es lo mismo que la anonimización, lo que no es así.

Los datos seudonimizados se utilizan para ocultar identidades y casi siempre queda un rastro entre el seudónimo y la identidad del titular del dato que corresponde, de manera que permite establecer quién es la persona y vincularlo con otros conjuntos de datos.

Por último, vale resaltar que en el ámbito de la salud la normativa permite la utilización de la información siempre y cuando se respeten los lineamientos básicos en cuanto al consentimiento y la disociación.