



**Universidad
Europea**

LAUREATE INTERNATIONAL UNIVERSITIES

TESIS DOCTORAL

SERIE: INGENIERÍA

Modelo de atributos clínicos a anonimizar para el cumplimiento de la Ley Orgánica 15/1999, de 13 de diciembre, de Protección de Datos de Carácter Personal

Doctorado en Tecnologías de la Información Aplicadas
Escuela de Arquitectura, Ingeniería y Diseño

Abel Lozoya de Diego

Dirigida por:

Dra. Dña. María Teresa Villalba de Benito

Co-Dirigida por:

Dra. Dña. María Arias Pou

Madrid, 2017

INFORME Y AUTORIZACIÓN DEL DIRECTOR PARA PRESENTAR LA TESIS DOCTORAL

El doctorando ha realizado una tesis en el área de Informática (Seguridad) y Derecho de Tecnologías de la Información (privacidad) aplicada al área de la Medicina. Se trata, por tanto, de una tesis que encaja por temática en el doctorado de Tecnologías de la Información aplicada, y que además destaca por ser multidisciplinar.

Durante el período de investigación el doctorando ha participado en 2 proyectos: uno interno de la Universidad Europea y otro a nivel nacional. Además, ha publicado 1 artículo en Congreso y 2 en revistas una de ellas indexada en Scopus. Cumple, por tanto, la normativa interna de la universidad para poder presentar su trabajo de tesis.

Su currículum vitae está altamente especializado en el ámbito de la Seguridad informática, habiendo realizado un máster universitario en dicha área y llevando 6 años trabajando en proyectos nacionales e internacionales para el Ministerio de Defensa. Ha recibido, además, el reconocimiento de Auditor Jefe ISO 27001 por AENOR y es Certified Ethical Hacker (CEH) por EC-Council.

Por todo ello, entendemos que el doctorando cumple con los méritos necesarios para la presentación de la tesis doctoral habiendo conseguido el reconocimiento de su trabajo en distintas publicaciones especializadas.

La Dra. Dña. María Teresa Villalba de Benito directora de la tesis y la codirectora Dra. Dña. María Arias Pou, de la que es autor D. Abel Lozoya de Diego

AUTORIZAN la presentación de la referida Tesis para su defensa en cumplimiento del Real Decreto 99/2011, de 28 de enero, por el que se regulan las Enseñanzas Oficiales de Doctorado, y de acuerdo al Reglamento de Enseñanzas Universitarias Oficiales de Doctorado de la Universidad Europea de Madrid RD 1393/2007 y RD99/2011.

Madrid, 24 de abril de 2017



Fdo.: EL DIRECTOR

**ESCUELA DE DOCTORADO E INVESTIGACIÓN
DE LA UNIVERSIDAD EUROPEA DE MADRID**

AGRADECIMIENTOS

En el presente trabajo de tesis han participado numerosas personas que de forma directa o indirecta han realizado sus aportaciones, críticas, comentarios, correcciones, o han aportado paciencia y ánimo en su desarrollo. Además, me ha permitido beneficiarme del conocimiento y la experiencia de muchas personas a las que agradeceré en este apartado sus aportes.

A mis directoras de tesis Dra. María Teresa Villalba de Benito y Dra. María Arias Pou quería darles las gracias por haber depositado su confianza sobre mi persona. Además, gracias también por su paciencia, conocimientos y guía durante el desarrollo de todo el trabajo.

A mi familia, en especial a mis padres Fernando y Marianela, porque siempre me han educado el sacrificio y el esfuerzo. Ellos junto con mi hermano Fernando siempre me han prestado su apoyo y amor incondicional.

A mis amigos, en especial al Dr. D. Álex Pavié Nova, al Dr. D. Yuri Torres de la Sierra y a la Ing. Dña. Sandra Bardón Moral. El primero, plantó la semilla de la investigación, lo que ha culminado en esta tesis doctoral. Tanto Yuri como Sandra, han aportado su conocimiento y se han prestado en ayudar en cualquier momento ante cualquier necesidad.

A mi perro Danko, que siempre me enseñó esa paciencia sin límite en su lucha por comerse un hueso.

A todos ellos, muchas gracias.

“A mis padres Fernando y Marianela, y a mi hermano Fernando que con su amor, trabajo y sacrificios me han permitido llegar hasta aquí”.

ÍNDICE

RESUMEN/ABSTRACT

0. INTRODUCCIÓN	15
0.1 CONTEXTO.....	18
0.1.1 <i>Derechos fundamentales</i>	18
0.1.2 <i>La protección de datos como derecho fundamental</i>	21
0.1.3 <i>Los datos médicos asociados a las nuevas tecnologías</i>	25
0.2 PROPÓSITO DE LA INVESTIGACIÓN	31
0.2.1 <i>Motivación</i>	32
0.2.2 <i>Objetivos</i>	32
0.2.3 <i>Etapas de trabajo</i>	33
0.3 PROYECTOS DE INVESTIGACIÓN EN LOS QUE SE ENMARCA LA TESIS.....	34
0.4 ESTRUCTURA DE LA TESIS.....	36
1. ESTADO DEL ARTE	39
1.1 INTRODUCCIÓN	41
1.2 NORMATIVA APLICABLE A NIVEL NACIONAL Y COMUNITARIO	41
1.2.1 <i>Constitución Española</i>	42
1.2.2 <i>Ley 14/1986</i>	42
1.2.3 <i>Ley Orgánica de Protección de Datos</i>	44
1.2.4 <i>Real Decreto 1720/2007</i>	45
1.2.5 <i>Ley 41/2002</i>	46
1.2.6 <i>Real Decreto 1093/2010</i>	49
1.2.7 <i>Directiva 95/46/CE</i>	50
1.2.8 <i>Reglamento General de Protección de Datos 2016/679</i>	52
1.2.9 <i>Constitución digital</i>	53
1.3 NORMATIVA APLICABLE EN OTROS PAÍSES.....	55

1.3.1	<i>África</i>	55
1.3.2	<i>Europa</i>	61
1.3.3	<i>Oceanía</i>	64
1.3.4	<i>Asia</i>	66
1.3.5	<i>América</i>	67
1.4	ESTUDIOS RELACIONADOS CON LA PROTECCIÓN DE DATOS MÉDICOS	77
1.5	CONCLUSIONES	86
2.	METODOLOGÍA DE INVESTIGACIÓN	91
2.1	INTRODUCCIÓN	93
2.2	METODOLOGÍA PARA LA CREACIÓN DE MODELOS DE DATOS MÉDICOS A ANONIMIZAR.....	95
2.2.1	<i>Fase 1. Desarrollo del modelo de datos médicos</i>	100
2.2.2	<i>Fase 2. Validación del modelo</i>	101
3.	EVALUACIÓN DEL PROCESO Y VALIDACIÓN	121
3.1	INTRODUCCIÓN	123
3.2	APLICACIÓN DE LA METODOLOGÍA AL CASO DE ESTUDIO	123
3.2.1	<i>Fase 1. Desarrollo del modelo de atributos</i>	124
3.2.2	<i>Fase 2. Validación del modelo</i>	129
4.	CONCLUSIONES Y CONTRIBUCIONES.....	199
4.1	CONCLUSIONES	201
4.2	CONTRIBUCIONES	207
5.	LÍNEAS FUTURAS DE INVESTIGACIÓN.....	211
6.	BIBLIOGRAFÍA	215
7.	ANEXOS.....	259

RESUMEN

En el sector sanitario existen grandes volúmenes de conocimiento, almacenados en formato digital. La utilización de estos datos para investigación mediante la aplicación de técnicas de minería de datos (*big data analytics*) permitiría la extracción de nuevo conocimiento, además del avance del área de informática aplicada a la Salud. Dado que estos datos son altamente sensibles, las bases de datos donde se almacenan están protegidas por la Ley Orgánica 15/1999, de 13 de diciembre de Protección de Datos de Carácter Personal (en adelante LOPD). En esta Ley Orgánica se clasifican los datos relacionados con la salud de cada paciente como “datos especialmente protegidos” por lo que el uso para investigación quedaría anulado de no cumplirse determinadas garantías. De esta forma, se garantizan los derechos fundamentales de los pacientes.

Sin embargo, en 2007 se aprobó la Ley 14/2007, de 3 de julio, de Investigación Biomédica en la que se define el término “anonimización” en el sector sanitario en el artículo 3 como el *“proceso por el cual deja de ser posible establecer, por medios razonables, el nexo entre un dato y el sujeto al que se refiere. Es aplicable también a la muestra biológica”*. Además, en el artículo 3 apartado i) se indica que un “dato anonimizado” o disociado es aquel *“que no puede asociarse a una persona identificada o identificable por haberse destruido el nexo con toda la información que identifique al sujeto o porque dicha asociación exige un esfuerzo no*

razonable, entendiéndose por tal el empleo de una cantidad de tiempo, gastos y trabajo desproporcionados”.

Una primera aproximación a esta cuestión en el contexto internacional la encontramos en la Ley estadounidense de Transferencia y Responsabilidad de Seguro Médico (*Health Insurance Portability and Accountability Act*, HIPAA por sus siglas en inglés) de 1996. Esta ley aporta protección a millones de trabajadores y a sus familias en los Estados Unidos que padecen de alguna afección médica.

En lo que a España se refiere, la LOPD no define qué campos de datos deben ser anonimizados, lo que dificulta enormemente los procesos de anonimización, así como otros procesos asociados como es el análisis de riesgos. Además, actualmente hay un creciente interés por estos datos dado que el desarrollo de las tecnologías como las relacionadas con el denominado Internet de la Cosas (IoT, *Internet of Things*), generan una gran cantidad de datos médicos personales que son almacenados en la nube mediante proveedores Cloud y que pueden ser utilizados para la mejora de la calidad de vida de las personas que los utilizan, así como para su salud por parte del personal médico e investigador. Pero de nuevo para poder usar dichos datos, al igual que para almacenarlos, éstos deben ser anonimizados por lo que se hace necesario facilitar estos procesos en la medida de lo posible.

ABSTRACT

Big volumes of knowledge are stored in digital format in the health care sector. Besides, investigators should be able to use these volumes by technologies of data mining for the extraction of a new knowledge. Since these data are highly sensitive, databases where they are stored are protected by the Organic Law 15/1999 of December 13th of Protection of Information of Personal Character (Spanish LOPD). According to the cited law, personal information about the health of patients is classified as "specially protected information", so certain guarantees need to be accomplished in order to use these data with research proposes. In this way the fundamental rights of patients are guaranteed.

Nevertheless, the Law 14/2007 of Biomedical Investigation was approved in 2007. In this Law, the term "anonymization" was defined in the health care sector, specifically in its article 3 as *"action by which it is no longer possible to establish, for reasonable means, the link between personal information and the subject to which it refers. It is applicable also to the biological sample"*. In addition, article 3i) indicates that an anonymised or dissociated data are those which *"cannot be associated with an identified or identifiable person because the link has been destroyed with all the information that identifies the subject or because this association requires an unreasonable effort, meaning the use of a disproportionate amount of time, money and work"*.

A first approximation to this issue in the international context is found in the 1996 Health Insurance Portability and Accountability Act (HIPAA). This law provides protection to millions of workers and their families in the United States who suffer from any medical condition.

As far as Spain is concerned, the LOPD law does not define which data fields should be anonymized, which greatly hinders anonymization processes, as well as other associated processes such as risk analysis. In addition, there is currently a growing interest in these data since technologies such as those related to the so-called Internet of Things generate a large amount of personal medical data that are stored in the cloud using Cloud providers and can be used for the improvement of the quality of life of the people who use them, as well as for their health by doctors and researchers. But again, to be able to use such data, as well as to store them, they must be anonymized so it is necessary to facilitate these processes as much as possible.

0. INTRODUCCIÓN

INTRODUCCIÓN

En este capítulo se exponen las carencias y preocupaciones detectadas que han ayudado en el inicio del camino para el desarrollo de esta tesis. El capítulo está dividido en tres secciones diferentes. La sección 0.1 recoge el contexto en materia de protección de datos desde su reconocimiento como derecho fundamental hasta su aplicación a las nuevas tecnologías. En la sección 0.2 se describe la motivación que ha surgido previa a desarrollar esta investigación, así como, los objetivos que se pretenden conseguir mediante las etapas de trabajo descritas en esta sección. En la sección 0.3, se explican los proyectos en los que se ha enmarcado esta investigación. Finalmente, la sección 0.4 detalla la estructura de este documento de tesis.

Como aclaración previa en relación a la terminología empleada en este trabajo de tesis, y aun disponiendo de un anexo (ANEXO 1) con un glosario de terminología aplicada para facilitar la lectura de este trabajo de tesis, es necesario aclarar que se ha utilizado “atributo médico” como sinónimo a “dato médico”. Esta relación ha sido obtenida a partir del estándar *Digital Imaging and Communications in Medicine* (en adelante DICOM) (*National Electrical Manufacturers Association*, 2016). En lo referente a datos médicos, DICOM los representa como atributos con propiedades y relaciones (Pianykh, 2009) utilizando el modelo entidad-relación de bases de datos (Barker, 1994).

0.1 Contexto

En este apartado se describe la situación a partir de la cual se ha podido extraer una serie de necesidades en protección de datos médicos, que han permitido definir posteriormente el propósito de esta investigación.

0.1.1 *Derechos fundamentales*

Es necesario comenzar con la definición que aporta la Real Academia Española sobre los derechos fundamentales: *“Los que, por ser inherentes a la dignidad humana y por resultar necesarios para el libre desarrollo de la personalidad, suelen ser recogidos por las constituciones modernas asignándoles un valor jurídico superior”* (Real Academia Española, 2014) dado que sientan las bases de todo trabajo en materia de derecho a la protección de datos. Este derecho fundamental, aun siendo inherente a la dignidad humana, no se ha reconocido de manera formal hasta finales de los años 1990 o en la década del 2000 como queda descrito en el Capítulo 1.

En la Declaración Universal de Derechos Humanos (en adelante DUDH) de la Organización de las Naciones Unidas de 1948 (Organización de las Naciones Unidas, 1948), en su artículo 12 se reconoció el respeto a la vida privada y familiar, estableciéndose como un hito inicial el derecho a la protección de la esfera privada de las personas frente a intrusos. Esta fue la primera iniciativa en relación con la protección de la vida privada. A esta DUDH, le siguió el Convenio Europeo de Derechos Humanos (en adelante CEDH) de 1950 (Tribunal Europeo de Derechos Humanos, 1950)

adoptado por el Consejo de Europa (en adelante CdE) que en su artículo 8 garantiza el derecho al respeto de la vida privada y familiar.

Con el fin de hacer cumplir el CEDH suscrito por los Estados miembros del CdE se creó el Tribunal Europeo de Derechos Humanos (en adelante TEDH) en 1959 en Estrasburgo, Francia. Desde su creación el TEDH ha examinado situaciones relacionadas con la protección de datos relacionadas con la interceptación de las comunicaciones (Tribunal Europeo de Derechos Humanos, 1984; Tribunal Europeo de Derechos Humanos, 2007), la vigilancia (Tribunal Europeo de Derechos Humanos, 1978; Tribunal Europeo de Derechos Humanos, 2010) y el almacenamiento de los datos (Tribunal Europeo de Derechos Humanos, 1987; Tribunal Europeo de Derechos Humanos, 2008).

En (Arenas Ramiro, 2008), su sección 1, se realiza un análisis sobre el desarrollo de los acontecimientos para lograr el reconocimiento de los derechos fundamentales. En dicho documento, se describe que cada uno de los Estados miembros del CdE de la década de 1950 que estaban en proceso de constitución avanzaban de manera independiente en el desarrollo de cada una de las normativas, existiendo situaciones de conflicto con las disposiciones establecidas por el CdE. Por este motivo, llegó a ser necesario que el Tribunal de Justicia de las Comunidades Europeas (autoridad judicial de la Comunidad Europea que garantiza la aplicación y la interpretación uniforme del Derecho de la Unión, en cooperación con los órganos jurisdiccionales de los Estados miembros del CdE, y en adelante TJCE), se empleara en crear una lista de derechos fundamentales que garantizaran aquellos derechos recogidos en las

Constituciones de cada Estado miembro, y que además fueran de aplicación para todos los países. Esta tarea se fue desarrollando a partir de los casos que se iban aprobando por el TJCE. El TJCE utilizaba el CEDH como herramienta auxiliar para aprobar las diferentes sentencias (Alemania, 1969; Alemania, 1970; Comisión Europea, 1974) con el fin de obtener un catálogo de derechos fundamentales con el valor de “principios generales del Derecho comunitario”.

El Tratado de la Unión Europea firmado en Maastricht, el 7 de febrero de 1992 (Unión Europea, 1992), incorporó la doctrina sobre los derechos fundamentales dictados por el TJCE al Derecho de la Unión Europea. Debido a que esto no era totalmente satisfactorio, el Grupo de Expertos (Comisión Europea, 1999) indicaron la necesidad de reconocer los derechos fundamentales explícitamente en la Unión Europea. Por ello, en (Consejo de la Unión Europea, 2000) quedó establecida la Carta de Derechos Fundamentales de la Unión Europea (en adelante Carta). En ella, queda patente el enfoque hacia la protección del ciudadano frente a la protección de la economía, iniciándose así un proceso de cambio con respecto a los Tratados aprobados hasta el momento. Esta Carta fue insertada en el Tratado por el que se establece una Constitución Europea (en adelante Constitución Europea) (Unión Europea, 2004) con el propósito de otorgarle un carácter jurídico. Sin embargo, para su entrada en vigor era necesaria la firma y ratificación de la Constitución Europea de 2004. Tras un proceso de votación, dicha Constitución no fue aprobada por Francia y Países Bajos, lo que produjo que la Carta de Derechos Fundamentales no tuviera un efecto jurídico en los Estados miembros.

En (Unión Europea, 2007) el 12 de diciembre de 2007 se firmó el Tratado de Lisboa y entró en vigor el 1 de diciembre de 2009. En dicho Tratado se adjuntó la Carta de Derechos Fundamentales y se la dotó a esta de un valor vinculante con los mismos efectos que los Tratados firmados en la Unión Europea. La Carta de Derechos Fundamentales tiene gran importancia porque en su artículo 8 reconoce como derecho fundamental la protección de datos.

0.1.2 La protección de datos como derecho fundamental

Al igual que en el reconocimiento de los derechos fundamentales, el reconocimiento del derecho fundamental a la protección de datos fue evolucionando a lo largo de los años como respuesta al avance tecnológico que afecta a los derechos fundamentales definidos en el pasado. De hecho, tras incorporar la jurisprudencia del TEDH, el TJCE resuelve el primer caso de protección de datos personales conocido como el caso *Stauder* (Alemania, 1969). En este caso se reconocía una protección de datos sobre el campo de información “nombre” debido a que era parte de la vida privada.

De forma paralela, el Consejo de Europa en su Convenio 108 (Tribunal Europeo de Derechos Humanos, 1950) en su artículo 1 hacía hincapié en dos aspectos no contemplados hasta el momento: el derecho a la vida privada de cada una de las personas pertenecientes a los Estados miembros y el tratamiento de forma automatizada de todos los datos de carácter personal correspondientes de dichas personas. A partir de este

momento se inicia un proceso de desarrollo y utilización de las Tecnologías de la Información y las Comunicaciones (en adelante TIC). En lo que concierne a la presente tesis, el Convenio 108 proporciona garantías relacionadas con la recopilación y el tratamiento de datos personales. En el caso de que falten las garantías jurídicas adecuadas, prohíbe el tratamiento de los datos calificados como sensibles, como son los datos relacionados con la salud.

En (Lucas Murillo de la Cueva, 2008) , el autor explica la evolución de la protección de datos desde un estado inicial en el que los estudios se enfocaban a que el derecho a la intimidad englobara toda la información de la persona, hasta un estado final que engloba la vida privada como un bien jurídico autónomo, pero indicando que toda la información de una persona no está incluido en un derecho a la intimidad. Nace así lo que se conoce como el derecho a la protección de datos o a la autodeterminación informativa.

En (Unión Europea, 1995) se desarrolló la Directiva 95/46/CE ante la necesidad de unificar el derecho a la Protección de Datos Personales a nivel europeo. Además, la autoridad responsable de garantizar este derecho a nivel europeo es el Supervisor Europeo de Protección de Datos, y a nivel estatal son las Autoridades de control independientes creadas para tal efecto.

Actualmente en el entorno europeo se ha aprobado el Reglamento (UE) 2016/679 del Parlamento Europeo y del Consejo, de 27 de abril de 2016, 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 y por el que se deroga la Directiva 95/46/CE (Reglamento general de protección de datos o RGPD) (Unión Europea, 2016). Este Reglamento será aplicable a partir del 25 de mayo de 2018.

Para transponer la Directiva 95/46/CE, en España se aprobó la Ley Orgánica 15/1999, de 13 de diciembre, de Protección de Datos de Carácter Personal (España, 1999) (en adelante LOPD). Su artículo 1 indica que tiene por objeto garantizar y proteger el tratamiento de los datos personales asociados al honor y a la intimidad personal y familiar. En cuanto a datos de la salud, la LOPD no define qué campos de información son englobados dentro de este grupo.

El Tribunal Constitucional español reconoció en la Sentencia del Tribunal Constitucional 292/2000, de 30 de noviembre (Tribunal Constitucional de España, 2000), la protección de datos como un derecho autónomo e independiente del derecho a la intimidad, denominado privacidad. Además, se definió el anterior derecho como de autodeterminación informativa, o de libre disponibilidad de los datos de carácter personal. Además, en dicha Sentencia, en su Fundamento Jurídico 6, se expone que tiene como fin garantizar el control sobre sus datos personales, sobre su uso y destino, que impida un uso lesivo de los mismos. En cuanto a su ámbito, la protección de datos incluye además cualquier tipo de dato personal que pueda afectar a sus derechos, sean o no fundamentales, porque su objeto no es solo la intimidad individual, sino los datos de carácter personal.

El derecho fundamental de protección de datos resalta el poder de disposición y control de los datos personales (facultad de obtención, recogida, acceso, almacenamiento y tratamiento de datos personales). El derecho a consentir la obtención y el tratamiento de los datos personales (mediante el uso o no de las TIC), requiere que el individuo esté informado de quién dispone de sus datos personales y la facultad de oponerse a esa posesión y/o usos.

De nuevo, a nivel nacional de España se aprueba el Real Decreto 1720/2007, de 21 de diciembre, por el que se aprueba el Reglamento de desarrollo de la Ley Orgánica 15/1999, de 13 de diciembre, de protección de datos de carácter personal (España, 2007). El R.D. 1720/2007 sí reconoce como derecho fundamental la protección de datos en el tercer párrafo del preámbulo: *“A fin de garantizar la necesaria seguridad jurídica en un ámbito tan sensible para los derechos fundamentales como el de la protección de datos [...]”*. Además, el citado Real Decreto incluye una definición de datos de carácter personal relacionados con la salud: *“las informaciones concernientes a la salud pasada, presente y futura, física o mental, de un individuo. En particular, se consideran datos relacionados con la salud de las personas los referidos a su porcentaje de discapacidad y a su información genética”*. Por su parte el RGPD define los datos relativos a la salud en su artículo 4 como: *“datos personales relativos a la salud física o mental de una persona física, incluida la prestación de servicios de atención sanitaria, que revelen información sobre su estado de salud”*. Al objeto que nos ocupa en la presente tesis, no se definen los atributos clínicos necesarios que deben ser

anonimizados para proteger todos los datos personales relacionados con la salud.

0.1.3 Los datos médicos asociados a las nuevas tecnologías

En (Marimón, 1999) se recoge la evolución de los sistemas de información hospitalaria clasificada en cinco etapas. La etapa inicial se sitúa en los años 70, en la que se empezaban a realizar las primeras implantaciones de aplicaciones informáticas relacionadas con procesos administrativos y productivos, pero sin conexión a otros sistemas externos del centro médico. En las siguientes etapas, los sistemas han ido evolucionando hasta llegar a interconexiones entre centros médicos a través de Internet, e incluso hacia la disponibilidad total del historial del paciente en cualquier lugar donde este se encuentre. Esto ha sido como consecuencia de los cambios culturales que se han ido produciendo en la sociedad, los cuales han permitido la creación de una demanda creciente de nuevos servicios y tecnologías. Uno de los ámbitos donde los ciudadanos han introducido las nuevas tecnologías, de una manera rápida a su vida cotidiana, ha sido la autogestión de su bienestar físico y de su salud. Este cambio beneficia a la sociedad porque ha llevado a sistemas de salud más eficientes en lo que a costes se refiere.

Sin embargo, las nuevas tecnologías implantadas en el ámbito de la salud traen consigo la necesidad de obtener datos individuales sometidos al máximo nivel de protección utilizando técnicas tales como la anonimización/disociación o la privacidad desde el diseño (Fischer & Morte Ferrer, 2013) para la obtención de conocimiento. Un ejemplo de

esto, es el estudio (Amin, Keshishian, Trocio, & Vo, 2017) en el que gracias a una muestra de datos anónimos compuesta por más de 40.000 pacientes se han obtenido resultados médicos sobre un fármaco. Este nivel máximo de protección de los datos médicos protegidos o *Protected Health Information* (denominación creada por la legislación de Estados Unidos (Estados Unidos, 1996), en adelante PHI) ha suscitado la creación de propuestas como la de la Asociación Española de Usuarios de Telecomunicaciones y de la Sociedad de la Información (en adelante AUTEISI) que presentó el estudio titulado *Privacy by Design* (Asociación Española de Usuarios de Telecomunicaciones y de la Sociedad de la Información, 2014) en la sede de Red.es (entidad pública que trabaja para que la sociedad española aproveche al máximo el potencial de Internet y las nuevas tecnologías). En este estudio, se propone aplicar la protección de datos desde el diseño de los sistemas de archivos en los que se guardan los mismos. AUTEISI, para llegar a desarrollar e implementar la propuesta anterior, analiza los riesgos y amenazas asociados al tratamiento de los datos personales creando un catálogo de riesgos y amenazas. Por otro lado, la Asociación Informática Médica Americana (AMIA, *American Medical Informatics Association*), en diciembre de 2015, celebró la séptima Reunión de Política Sanitaria con el propósito de detectar los desafíos que deben afrontar los agentes sanitarios y legislativos. La clasificación y difusión de dichos desafíos fomentarían el uso de datos médicos gestionados por un administrador de datos. Tras la Reunión, se consensó la necesidad de medios económicos que financien estudios sobre nuevas técnicas de anonimización para garantizar el uso de los datos médicos según la normativa vigente (Hripcsak et al., 2014).

Asociada a la evolución de los sistemas de salud, se encuentra la preocupación del usuario por la monitorización de su salud mediante dispositivos del Internet de las Cosas (IoT, *Internet of Things*) con una previsión en el año 2020 de 50 billones de dispositivos (Bradley, Barbier, & Handler, 2013) y la falta de adaptación legislativa en este ámbito con las consiguientes críticas a las instituciones europeas (Avancha, Baxi, & Kotz, 2012; Fernández Barbudo, 2015; Pérez Marzabal, 2014). En (Avancha et al., 2012) los autores realizan una revisión del estado del arte sobre leyes y cuestiones referentes al tratamiento y la privacidad de los datos de los usuarios que hacen uso de dispositivos móviles con fines médicos. La gran cantidad de información generada (datos médicos, localización geográfica, hábitos alimenticios y físicos, incluso horarios en los que se realizan las diferentes mediciones de monitorización) que se transmite y se comparte entre los diferentes dispositivos móviles médicos y sistemas centrales es de carácter privado. Sin embargo, las leyes existentes se centran más en entornos clínicos que en los sistemas móviles/portables como los *wearables*. Los autores introducen un *framework* que actúa como guía de buenas prácticas necesaria en el tratamiento de los datos generados y pertenecientes a los usuarios. Entre ellas, destacan las siguientes necesidades de los usuarios: a) deben conocer todos los datos recogidos por los dispositivos, b) deben saber el propósito específico de cada dato que es recogido, c) pueden limitar la cantidad de datos generados, y d) deben tener el control absoluto de los datos. Sin embargo, el trabajo no estudia ni propone ningún mecanismo que identifique qué datos poseen un carácter privado. Además, dada la gran cantidad de atributos generados, no se hace hincapié en aquellos que son redundantes y que deben de ser eliminados mejorando así, la

legibilidad de la información. Asimismo, plantean preguntas abiertas de investigación entre la que destaca la necesidad de disponer de un sistema que permita adaptarse a diferentes jurisdicciones en términos de privacidad. En este caso, sólo Estados Unidos ha hecho esfuerzos y ha desarrollado un modelo global para los sistemas de información médicos en el que indica los atributos a ser anonimizados. Esto está recogido en *Health Insurance Portability and Accountability Act* (en adelante HIPAA) (Estados Unidos, 1996), pero no ha proporcionado la metodología utilizada para que sea replicada en otros países, por lo que la pregunta de investigación planteada está sin responder. Los consumidores de estos productos en el ámbito sanitario cuentan con una desprotección adicional ya que numerosos estudios han evidenciado el creciente número de ciberataques en los últimos años en el área de Salud (Villalba de Benito, De Buenaga Rodríguez, Gachet Páez, & Aparicio Galisteo, 2015). Estos datos relacionados con la salud son muy atractivos para los ciberdelicuentes por ser muy completos al incluir gran cantidad de datos personales, y se ha evidenciado una falta de recursos y tecnologías para protegerlos (Ponemon Institute LLC, 2015; Symantec, 2015). Para apoyar en esta tarea a las instituciones europeas o a sus participantes, el *Council of European Professional Informatics Societies* (CEPIS) ha creado la red compuesta de expertos en legislación relacionada con las nuevas tecnologías y en temas relacionados con la seguridad denominada “*CEPIS LSI SIN*” (*CEPIS Legal & Security Issues Special Interest Network*) (Villalba de Benito, 2016). Su última declaración de apoyo, por ejemplo, ha estado enfocada en la privacidad de los proveedores de servicios de Internet (Verbeek, 2016).

Se requiere, por tanto, un mayor desarrollo en diferentes ámbitos de la privacidad de los datos para obtener una protección adecuada tal y como indica el Grupo de Trabajo Internacional sobre la Protección de Datos en las Telecomunicaciones (en adelante IWGDPT), creado en 1983 en la Conferencia Internacional de Privacidad y Protección de Datos de los Comisionados a iniciativa del Comisionado de Berlín para la Protección de datos. El IWGDPT, el 6 de mayo de 2014, publicó el trabajo (International Working Group on Data Protection in Telecommunications, 2014) donde se aportan nuevas propuestas que intentan solventar los problemas derivados de la protección de datos surgidos a partir de nuevos aspectos (datos usados para nuevos propósitos, maximización de datos y la falta de transparencia en su tratamiento, información confidencial, riesgo de reidentificación, implicaciones de seguridad, etc.), haciendo uso de la tecnología Big Data (Agencia Española de Protección de Datos & ISMS Forum Spain, 2017; González Pedraz, 2014) utilizada como motor económico en la explotación de los datos y sus ventajas en la reidentificación de los datos (President's Council of Advisors on Science and Technology, 2014; Recio Gayo, 2015; Supervisor Europeo de Protección de Datos, 2014). Por ello, el IWGDPT para garantizar una anonimización robusta propone evaluaciones de impacto en la privacidad sobre los datos que han sido anonimizados. Con ello, se pretende obtener un nivel de riesgo frente a un proceso de reidentificación de los mismos datos. Además, a esto hay que sumar la necesidad de garantizar que el proceso de anonimización se ha realizado correctamente para que los datos cumplan el fin para el que fueron recogidos. Para ayudar en el proceso de toma de decisiones en la elección de algoritmos de anonimización, IWGDPT propone la creación de una red u órgano donde

se intercambie información sobre los algoritmos de anonimización más adecuados en cada caso. En Reino Unido se ha creado la Red de Anonimato (UKAN, *UK Anonymisation Network*), coordinada por las universidades de Manchester y Southampton, el Instituto Open Data y la Oficina Nacional de Estadística. Además, el IWGDPT (Grupo de Trabajo sobre Protección de Datos del Artículo 29, 2014) indica que *“han de definirse con claridad los requisitos previos (el contexto) y los objetivos del proceso para obtener la anonimización deseada al mismo tiempo que se generan datos útiles”* y recomienda que *“la solución óptima debe decidirse caso por caso y puede conllevar la combinación de diversas técnicas”*. Como se puede ver, cada proceso de anonimización es un proceso arduo que debería incluir la Evaluación de Impacto en la Privacidad (o más conocida como PIA del inglés *Privacy Impact Assessments*) (Agencia Española de Protección de Datos, 2014), la anonimización y la definición de los requisitos previos. Sumado a esto último, incluso se debe contemplar el concepto de dato cuasi-identificador (El Emam, Jabbouri, Sams, Drouet, & Power, 2006) que añade aún más complejidad en los procesos de anonimización/disociación y es un factor a tener en cuenta en los procesos de Evaluación de Impacto en la Privacidad (en adelante EIPD) (Agencia Española de Protección de Datos & ISMS Forum Spain, 2017). Por ello, con el presente trabajo de tesis se desarrolla un modelo de atributos médicos a anonimizar para procesos de anonimización bajo la legislación aplicable en España, con la intención de mejorar y agilizar procesos de anonimización repetitivos.

En relación con la anonimización señalar que la Agencia Española de Protección de Datos publicó el 13 de octubre de 2016 unas “*Orientaciones y garantías en los procedimientos de anonimización de datos personales*” (Agencia Española de Protección de Datos, 2016) a la que después haremos referencia.

En este capítulo se han expuesto una serie de carencias y debilidades existentes en la comunidad científica actual sobre la falta de protección aplicada en los datos médicos. Cada una de estas carencias se centra en áreas de conocimiento diferentes, tales como medicina, derecho y seguridad informática. Además, el proceso de creación de modelos propuesto en este trabajo surge como una adaptación del germen presentado en (Villalba, Fernández-Sanz, & Martínez, 2010), donde se define el proceso DuMoD del inglés *Domain-oriented qQuality MOdels Development* basado en el área de la ingeniería del software. Por ello, queda presente el carácter interdisciplinar del presente trabajo de tesis que ofrece un proceso de creación de modelos de datos médicos para ayudar en la anonimización de atributos médicos.

0.2 Propósito de la investigación

El propósito de la investigación recoge la motivación que ha surgido al detectar las deficiencias y necesidades reseñadas en el punto anterior, el objetivo de esta tesis y las etapas de trabajo para su consecución.

0.2.1 Motivación

La finalidad que se pretenden lograr con la realización de esta tesis es desarrollar un primer paso hacia una propuesta que sea utilizada en una futura ley de protección de datos o como complemento a la actual, con el fin de definir los campos de datos médicos mínimos a ser anonimizados. Con ello, se pretende garantizar que, tras un proceso de anonimización de los datos personales que permiten identificarle, el paciente pase a estar en un estado “no identificable” ante técnicas de identificación del individuo mediante el tratamiento informático de los datos personales. Además, esto contribuirá a que las investigaciones médicas puedan obtener el conocimiento almacenado en las bases de datos médicas siempre manteniendo el nexo de unión entre los datos identificativos y los datos clínico-asistenciales (Igalada Menor, 2003).

0.2.2 Objetivos

El objetivo principal a formular en este trabajo de tesis es comprobar si:
Es viable desarrollar un modelo de atributos clínicos a anonimizar que garanticen el cumplimiento de la Ley Orgánica 15/1999, de 13 de diciembre, de Protección de Datos de Carácter Personal

Este objetivo general puede desglosarse en los siguientes objetivos específicos:

- Revisión sistemática de las normativas disponibles en protección de datos para extraer todos los campos médicos que permitan obtener un listado de datos médicos protegidos.

- Adaptación de una metodología de creación de modelos validada al dominio de protección de datos médicos.
- Definición de una taxonomía de atributos clínicos con las relaciones y pesos entre ellos, validada a nivel nacional.
- Validación de la metodología adaptada mediante la definición de un modelo formal empíricamente validado para el caso español, de forma que pueda ser aplicada en otros países.

0.2.3 *Etapas de trabajo*

A continuación, se presenta un esquema de trabajo:

1. Estudio de la normativa nacional e internacional sobre protección de datos con el propósito de obtener un conjunto inicial de campos médicos protegidos por la legislación vigente.
2. Análisis de artículos y bases de datos científicas relacionadas para la extracción de nuevos atributos o encontrar la justificación de la existencia de casos especiales.
3. Definición de un catálogo preliminar de atributos según el análisis del estado de la cuestión obtenido en los pasos anteriores.
4. Validación empírica de la taxonomía obtenida a través de la utilización de metodología cuantitativa para la obtención de un modelo válido en España:
 - a. Realización de entrevistas a un conjunto reducido de personal relacionado con el ámbito sanitario para

eliminar campos de información médica no utilizada en España.

- b. Construcción del constructo a utilizar para la recogida de datos.
- c. Configuración y uso de dos herramientas web de terceros para facilitar el proceso de recogida de datos de expertos.
- d. Realización de cuestionarios piloto con personal relacionado con el ámbito sanitario para adaptar las herramientas web a las necesidades de los encuestados para la validación del constructo.
- e. Validación empírica del modelo mediante análisis estadístico multivariante, en concreto análisis de componentes principales, análisis factorial exploratorio y el análisis factorial confirmatorio del modelo.
- f. Desarrollo de las conclusiones, aportaciones del trabajo de tesis y el planteamiento de posibles trabajos futuros.

0.3 Proyectos de investigación en los que se enmarca la tesis

Esta tesis doctoral se enmarca en los siguientes proyectos de investigación (ver Tabla 1 y Tabla 2).

Proyecto privado de investigación	
Título del contrato/proyecto:	Sistema inteligente de abstracción de información y soporte al diagnóstico de embarazos de riesgo, basado en parámetros de biometría fetal y conocimiento del dominio (OTRI 2014 UEM 20).
Tipo de contrato:	Privado.

Proyecto privado de investigación	
Empresa/Administración financiadora:	Universidad Europea.
Entidades participantes:	Hospital Gregorio Marañón.
Duración:	Enero 2014 - Enero 2015
Investigador responsable:	M ^ª Teresa Villalba de Benito.
Número de investigadores participantes:	8
Precio total del proyecto:	6.000 euros.
Descripción:	En este proyecto se anonimizaron datos médicos utilizando una versión inicial del proceso lo que permitió recoger datos sobre las necesidades de los procesos de anonimización que posteriormente se incluirían en este trabajo de tesis.

Tabla 1 Proyecto de carácter privado en el que se enmarca esta investigación

Proyecto público de investigación	
Título del contrato/proyecto:	IPHealth: Plataforma Inteligente basada en Open, Linked y Big Data para la toma de decisiones y aprendizaje en el ámbito de la Salud.
Tipo de contrato:	Público
Empresa/Administración financiadora:	Ministerio de Economía y Competitividad. Plan Estatal de Investigación Científica y Técnica y de Innovación (2013-2016). Programa Estatal de Fomento de la Investigación Científica y Técnica de Excelencia. TIN-2013-47153-C03-01.
Entidades participantes:	Universidad Europea de Madrid (coordinador), Universidad de Huelva, Universidad de Vigo.
Duración:	01/01/2014 - 31/12/2017
Investigador responsable:	Manuel de Buenaga Rodríguez y Diego Gachet Páez.
Número de investigadores participantes:	11(UEM), 11(UVIGO), 4(UHU).
Subvención total del proyecto:	281.484 euros.
Subvención del proyecto UEM:	115.954 euros.
Descripción:	Este proyecto está aún en desarrollo, y las aportaciones de este trabajo de tesis al proyecto son las de facilitar los procesos de anonimización en las arquitecturas Big Data para el ámbito de la Salud y el bienestar.

Tabla 2 Proyecto de carácter público en el que se enmarca esta investigación

0.4 Estructura de la tesis

El presente trabajo de tesis está estructurado en siete capítulos como se describe a continuación:

- 0. Introducción: Se presenta una introducción del problema, describe los objetivos que persigue el presente trabajo, así como las aportaciones que se esperan alcanzar; muestra el problema y expone las etapas de trabajo para llevar a cabo la solución propuesta, describe los proyectos de investigación en los que se enmarca este trabajo y, por último, detalla la estructura de este documento.
- 1. Estado del arte: En este capítulo se describen los conceptos básicos en protección de datos personales, la normativa aplicable a nivel nacional sobre protección de datos personales relacionados con la salud que permita saber el estado del arte que afecta a España, la normativa que, no siendo de aplicación a España puede aportar información acerca del modelo de atributos sanitarios utilizados en otros países y, por último, una breve conclusión.
- 2. Metodología de investigación: Se desarrollan las fases del método de investigación utilizado para la obtención de la metodología de desarrollo del modelo de atributos clínicos presentado en esta tesis.

- 3. Evaluación del proceso y validación: Se describen tanto la validación de la aplicabilidad del proceso definido en el anterior capítulo, como la corrección y validación del modelo teórico de atributos.
- 4. Conclusiones y contribuciones: Este capítulo está dedicado a las conclusiones extraídas y a las contribuciones de la tesis.
- 5. Líneas futuras de investigación: Se detallan las líneas de investigación para la mejora en la definición de atributos clínicos en las normativas futuras sobre protección de datos relativos a la salud.
- 6. Bibliografía: En este capítulo se recogen las referencias bibliográficas utilizadas como recursos en el presente trabajo de investigación.
- 7. Anexos: En este capítulo se recogen los anexos que complementan la información del trabajo de tesis.

1. ESTADO DEL ARTE

ESTADO DEL ARTE

1.1 Introducción

Este capítulo expone y justifica los trabajos relacionados que actúan como estado de la ciencia, los cuales sustentan los cimientos del presente trabajo. El capítulo se divide en cuatro secciones diferentes. La sección 1.2 analiza la normativa aplicable a nivel nacional y comunitario en protección de datos médicos. La sección 1.3 describe la normativa aplicable en otros países, así como normativa internacional. La sección 1.4 recoge los estudios y propuestas realizadas por autores expertos en la materia relacionada. Esta sección aporta conocimiento en el tratamiento de atributos clínicos que deben ser protegidos para evitar la reidentificación del paciente. Finalmente, la sección 1.5 recoge las conclusiones que merece el capítulo.

1.2 Normativa aplicable a nivel nacional y comunitario

En este apartado se explica cronológicamente y desde la legislación nacional a la europea aplicable a nivel nacional sobre protección de datos. Esta normativa puede ser a nivel interno o nacional (por ejemplo la Constitución Española) o a un nivel superior (por ejemplo, nivel europeo como el Reglamento General de Protección de Datos 2016/679).

1.2.1 Constitución Española

A nivel nacional, la Constitución de 1978 (España, 1978) fue la primera norma en la que se incluyó el derecho a la protección de datos mediante la protección a la intimidad.

Los autores de la Constitución de 1978 haciendo uso de la Constitución de Portugal proclamada dos años antes, fijaron en su artículo 18.4 el derecho a la intimidad en el campo de la informática dado que se preveía el uso de las tecnologías en la protección de datos.

En el artículo 43.1 de la misma Constitución se expone el derecho a la protección de la salud de los ciudadanos impulsada por la administración pública. En dicho artículo se delimita la frontera entre el interés general de la salud y el interés particular de la intimidad de datos médicos. Este aspecto está muy presente en todo el desarrollo legislativo llevado a cabo hasta el momento, siendo el interés general el que prevalece al interés particular debido a que es necesario para el avance del conocimiento a partir de los datos de carácter médico (datos generados por la población y almacenados en los historiales clínicos).

1.2.2 Ley 14/1986

La Ley 14/1986 General de Sanidad (Ley 14/1986), de 25 de abril, se promulgó el 25 de abril de 1986 por las Cortes Generales (España, 1986). Su finalidad es establecer las acciones relativas a la protección de la salud

que establece la Constitución Española en su artículo 43.1 *“Se reconoce el derecho a la protección de la salud”*.

En el artículo (Criado Álvarez, Repullo Labrador, & García Millán, 2011) se realiza un análisis de la Ley 14/1986 en el que se denominan como “artículos inmaduros” los relacionados con la información sanitaria ya que no define el tipo de información que se requiere aunque incluye la información como un bien a proteger definiendo los actores implicados. Sin embargo, no se refleja el análisis del artículo 61 de la Ley 14/1986 que fue derogado por la Ley 41/2002, de 14 de noviembre.

El artículo 61 contiene definiciones sobre los límites físicos dentro de los cuales no deben salir los datos de cada paciente. Además, refleja el personal al que estarán accesibles los datos del paciente y se hace cargo del derecho a la intimidad del paciente, aunque sin definir los mecanismos que garanticen dicho derecho. Además, expone como los poderes públicos son los encargados de garantizar los derechos y deberes tanto del paciente, como del personal sanitario.

En conclusión, se puede destacar que la conciencia referente a la protección de datos personales del paciente actual difiere, de forma significativamente con respecto a los enfoques de protección de datos a los que se enfrenta la sociedad.

1.2.3 Ley Orgánica de Protección de Datos

La Ley Orgánica de Protección de Datos aprobada por la Cortes el 13 de diciembre de 1999 (LOPD) (España, 1999), sustituyó a la Ley Orgánica de Regulación del Tratamiento Automatizado de los Datos (LORTAD) (España, 1992).

En la LOPD se trató de adaptar el entorno jurídico en protección de datos al nuevo panorama tecnológico que nos rodea. Para ello, se defiende el derecho a la intimidad y a la privacidad de los ciudadanos en todos sus ámbitos tanto para personas o empresas, como a organismos de carácter público o privado que dispongan de datos de carácter personal. Todo ello, basándose en el artículo 18 de la Constitución Española de 1978.

La LOPD afecta a todos los datos que hacen referencia a personas físicas registradas sobre cualquier soporte. Por ello, quedan incluidos los datos clínicos independientemente del soporte en el que sean almacenados. Así mismo, podrán ser sometidos a tratamiento tal y como se indica en el artículo 4.2 *“Los datos de carácter personal objeto de tratamiento no podrán usarse para finalidades incompatibles con aquellas para las que los datos hubieran sido recogidos. No se considerará incompatible el tratamiento posterior de éstos con fines históricos, estadísticos o científicos”*. De esta forma, se legitima el tratamiento de los mismos en estudios de investigación.

La LOPD, denomina como datos de carácter personal aquellos referentes a los datos de la salud tal y como se indica en el artículo 7.3 *“Los datos de*

carácter personal que hagan referencia al origen racial, a la salud y a la vida sexual sólo podrán ser recabados, tratados y cedidos cuando, por razones de interés general, así lo disponga una ley o el afectado consienta expresamente". En la Disposición adicional sexta se explica la necesidad de consentimiento por parte del paciente en aquellos casos en los se manipulen datos personales relativos a la salud.

Aunque en la LOPD se hace referencia al tratamiento sanitario de los datos personales relativos a la salud (art. 8), no se indica en ningún momento aquello que se considera como datos de la salud.

1.2.4 Real Decreto 1720/2007

El Real Decreto 1720/2007 de 21 de diciembre por el que se desarrolla la Ley Orgánica de Protección de Datos (RDLOPD) y que deroga el Real Decreto 994/1999 de 11 de junio por el que se aprobaban el Reglamento de Medidas de Seguridad de los ficheros automatizados que contengan datos de carácter personal (España, 2007), supone un gran avance con respecto a la protección de datos ampliando en 129 artículos más el Real Decreto derogado.

El RDLOPD, el artículo 5.1.g define "*Datos de carácter personal relacionados con la salud: las informaciones concernientes a la salud pasada, presente y futura, física o mental, de un individuo. En particular, se consideran datos relacionados con la salud de las personas los referidos a su porcentaje de discapacidad y a su información genética*".

Como se puede observar la definición de datos de la salud no termina de indicar qué tipo de campos de información son considerados datos relacionados con la salud, aunque cita dos posibles categorías: los atributos relacionados con la discapacidad o la información genética.

El artículo 81.3.a define los datos a los que se deben aplicar medidas de seguridad de nivel alto: *“Los que se refieran a datos de ideología, afiliación sindical, religión, creencias, origen racial, salud o vida sexual.”*. Quedan incluidos los datos de la salud como aquellos en los que se deben aplicar medidas de seguridad de alto nivel, sin embargo, no se especifican a que datos propios de la salud se refiere.

1.2.5 Ley 41/2002

Durante la exposición, tanto de la LOPD, como del Reglamento 1720/2007 presentado anteriormente, se crea la necesidad de ahondar en la definición de datos que se refieren a la salud. Por ello, se hace necesario analizar la Ley 41/2002, de 14 de noviembre (España, 2002), básica reguladora de la autonomía del paciente y de derechos y obligaciones en materia de información y documentación clínica (en adelante Ley 41/2002), en la que se determinan los Derechos y Obligaciones en Materia de Información y Documentación Clínica.

La Ley 41/2002 legitima la recogida de la información siempre que se garantice la intimidad y la dignidad humana del paciente tal y como se refleja en el artículo 2.1: *“La dignidad de la persona humana, el respeto a*

la autonomía de su voluntad y a su intimidad orientarán toda la actividad encaminada a obtener, utilizar, archivar, custodiar y transmitir la información y la documentación clínica". La Ley garantiza la confidencialidad de los datos como se indica en su artículo 7: "[...] *el carácter confidencial de los datos referentes a su salud [...]*". En su artículo 16.3 hace mención a la protección del acceso a los datos clínicos almacenados en la historia clínica: "*El acceso a la historia clínica con fines judiciales, epidemiológicos, de salud pública, de investigación o de docencia, se rige por lo dispuesto en la Ley Orgánica 15/1999, de Protección de Datos de Carácter Personal, y en la Ley 14/1986, General de Sanidad [...]*".

El artículo 3 de la Ley 41/2002, define la información clínica de la siguiente manera: "*todo dato, cualquiera que sea su forma, clase o tipo, que permite adquirir o ampliar conocimientos sobre el estado físico y la salud de una persona, o la forma de preservarla, cuidarla, mejorarla o recuperarla.*" Esta definición es de gran utilidad con respecto a la definición del concepto, pero no aporta ningún dato concreto que ayude a definir los campos que deben ser anonimizados, evitando así, la identificación del paciente.

En el artículo 15 apartado 2 de esta misma ley, se detalla el contenido de la historia clínica de cada paciente tal y como se recoge en la Tabla 3, aunque no se concretan los datos de los que consta cada informe, autorización o documentación necesaria en estos documentos.

Contenido de la historia clínica recogida en Ley 41/2002	
Documentación relativa a la hoja clínico-estadística.	Consentimiento informado. ¹
Autorización de ingreso. ¹	Informe de anestesia. ¹
Evolución y planificación de cuidados de enfermería.	Informe de quirófano o de registro del parto. ¹
Anamnesis y la exploración física.	Informe de urgencia. ¹
Evolución.	Informe de anatomía patológica. ¹
Informes de exploraciones complementarias.	Aplicación terapéutica de enfermería.
Hoja de interconsulta.	Gráfico de constantes. ¹
Órdenes médicas.	Informe clínico de alta. ¹
¹ Necesarios para cumplimentar la historia clínica cuando se requiera o en procesos de hospitalización.	

Tabla 3 Contenido de la historia clínica extraída del artículo 15 apartado 2

En la Ley 41/2002 en el artículo 16.3, existe una definición encubierta: “[...] *El acceso a la historia clínica con estos fines obliga a preservar los datos de identificación personal del paciente, separados de los de carácter clínico-asistencial, de manera que como regla general quede asegurado el anonimato, salvo que el propio paciente haya dado su consentimiento para no separarlos. Se exceptúan los supuestos de investigación de la autoridad judicial en los que se considere imprescindible la unificación de los datos identificativos con los clínico-asistenciales [...]*”. En esta definición se contemplan datos de identificación personal del paciente y de carácter clínico-asistencial. No se hace mención sobre qué datos son agrupados en cada categoría. En el mismo artículo se indica la necesidad de un proceso contrario a la anonimización para aquellos casos en los que se hiciera uso de la historia clínica por parte de la autoridad judicial.

1.2.6 Real Decreto 1093/2010

El Real Decreto 1093/2010, de 3 de septiembre, por el que se aprueba el conjunto mínimo de datos de los informes clínicos en el Sistema Nacional de Salud (España, 2010); establece el conjunto mínimo de datos que deben contener una serie de documentos clínicos (quedan enumerados en su artículo 3). Este conjunto de datos debe permitir homogeneizar dichos datos por todos los centros y dispositivos existentes que integran el Sistema Nacional de Salud independientemente del soporte, electrónico o papel en el que se encuentre.

Los documentos y/o informes donde quedan reflejados los campos médicos que deben ser incluidos como mínimo, son los siguientes: informe clínico de alta –tras hospitalización–, de consulta externa, de urgencias, de Atención Primaria (AP), de resultados de pruebas de laboratorio, de resultados de pruebas de imagen, de cuidados de enfermería y, finalmente, de la historia clínica resumida.

La Tabla 4 especifica una muestra de los campos médicos mínimos que son identificados en el Real Decreto como datos del paciente. Nótese que los campos recogidos en la Tabla 4, se tratan de un conjunto mínimo de atributos que al menos debe contener, pero cada centro asistencial podrá añadir los que considere oportunos.

Extracto de atributos médicos recogidos en R.D. 1093/2010	
CIP de la C. Autónoma	NASS
CIP Europeo	Nombre
Código Postal	Nombre de la vía
Código SNS	País

Extracto de atributos médicos recogidos en R.D. 1093/2010	
Cuidador Principal	Persona de Referencia
DNI/T. Residencia/Pasaporte	Piso
Domicilio	Primer Apellido
Fecha de nacimiento	Provincia
Letra	Segundo Apellido
Municipio	Sexo
N.º Cama / N.º Consulta	Teléfono
N.º Historia Clínica	Tipo de vía

Tabla 4 Muestra de atributos médicos del paciente recogidos en el R.D. 1093/2010 (España, 2010)

Los campos de información que componen estos documentos clínicos, se han utilizado para validar los campos de información que pertenecerán al conjunto de campos válidos para el caso de España. Este proceso se explica en Capítulo 3 de la presente tesis.

1.2.7 Directiva 95/46/CE

La Directiva 95/46/CE (Unión Europea, 1995) nació para homogeneizar la protección de datos en todos los países miembros, y así garantizar, los derechos fundamentales por igual en todas las naciones que componen la Unión Europea. Además, los ciudadanos cuentan con una protección equivalente.

En la “Sección III Categorías Especiales de Tratamientos” en el que en el artículo 8.1 “Tratamiento de categorías especiales de datos” se indica: *“Los Estados miembros prohibirán el tratamiento de datos personales que revelen el origen racial o étnico, las opiniones políticas, las convicciones religiosas o filosóficas, la pertenencia a sindicatos, así como el tratamiento de los datos relativos a la salud o a la sexualidad.”* De esta

manera, quedan protegidos los datos de la salud. Sin embargo, no se contemplan casos extremos en los que sea necesaria la utilización de dichos datos por el bien común o del país.

El Considerando 26 de la Directiva dispone: *“Considerando que los principios de la protección deberán aplicarse a cualquier información relativa a una persona identificada o identificable; que, para determinar si una persona es identificable, hay que considerar el conjunto de los medios que puedan ser razonablemente utilizados por el responsable del tratamiento o por cualquier otra persona, para identificar a dicha persona; que los principios de la protección no se aplicarán a aquellos datos hechos anónimos de manera tal que ya no sea posible identificar al interesado; que los códigos de conducta con arreglo al artículo 27 pueden constituir un elemento útil para proporcionar indicaciones sobre los medios gracias a los cuales los datos pueden hacerse anónimos y conservarse de forma tal que impida identificar al interesado”*. Es la primera legislación a nivel europeo que se introduce el concepto de datos anónimos para que se impida la reidentificación del individuo.

Por lo tanto, esta directiva crea un marco regulador que pretende establecer un nivel homogéneo de protección de los datos personales en la Unión Europea garantizando, además, la libre circulación de los mismos.

1.2.8 Reglamento General de Protección de Datos 2016/679

El Reglamento General de Protección de Datos, de 27 de abril de 2016, al que ya se ha hecho referencia, (Unión Europea, 2016) define en su artículo 4, los datos relativos a la salud como *“los datos personales relativos a la salud física o mental de una persona, incluida la prestación de servicios de atención sanitaria, que revelen información sobre su estado de salud”*. Además, el RGPD refleja, en ese mismo artículo, el concepto de seudonimización como *“el tratamiento de datos personales de manera tal que ya no puedan atribuirse a un interesado sin utilizar información adicional, siempre que dicha información adicional figure por separado y esté sujeta a medidas técnicas y organizativas destinadas a garantizar que los datos personales no se atribuyan a una persona física identificada o identificable”*. También indica que si tras realizar un proceso de seudonimización, los datos pueden atribuirse a una persona física mediante el uso de información adicional, se considerarán estos datos como identificables y deberán seguir la legislación de protección de datos. Además, el RGPD indica que se deben tener en cuenta *“todos los factores objetivos, como los costes y el tiempo necesarios para la identificación”* en el proceso de reidentificación con el propósito de evaluar la probabilidad de identificar directa o indirectamente a una persona. También, se deben considerar los medios tecnológicos disponibles y los nuevos avances en esta materia para evitar poner en riesgo la protección de los datos seudonimizados. Y a continuación dispone que el RGPD no es de aplicación a los datos anónimos, pero sí a los datos seudonimizados que son los que regula.

Otra referencia existente en el RGPD a seudonimización se realiza en el artículo 32, referente a la seguridad de los datos. Éste, dispone que para garantizar un nivel de seguridad adecuado al riesgo, el responsable y el encargado del tratamiento deberán aplicar medidas técnicas (seudonimización o el cifrado) y organizativas atendiendo a los aspectos recogidos en la Tabla 5.

Aspectos relacionados con la seudonimización	
Estado de la técnica de seudonimización.	Contexto de los datos personales.
Costes de aplicación del proceso.	Fines de tratamiento de seudonimización.
Naturaleza de los datos.	Probabilidad y gravedad de las variables.
Alcance del proceso.	Derechos y libertades de las personas físicas.

Tabla 5 Aspectos a tener en cuenta en un proceso de seudonimización (Unión Europea, 2016)

Por su parte, el artículo 40 del RGPD dedicado a los códigos de conducta, recoge que podrán crear y/o modificar códigos de conducta por parte de las asociaciones y organismos que tengan el rol de responsables o encargados del tratamiento de datos personales, con el propósito de especificar la aplicación de disposiciones del Reglamento, por ejemplo, la seudonimización de datos personales.

1.2.9 Constitución digital

El 14 de marzo de 2017 la Fundación Ramón Areces junto con la Organización para la Cooperación y el Desarrollo Económicos (en adelante OCDE) organizaron el seminario titulado “La nueva revolución de la producción: la transformación digital” (Fundación Ramón Areces & OCDE, 2017) en el que participó el ministro de Energía, Turismo y Agenda

Digital y puso de manifiesto la necesidad de convertir la Constitución Española de 1978 denominada por él como “analógica” en una “Constitución digital” que tenga en cuenta la evolución tecnológica de la sociedad y proteja a los ciudadanos en la red mediante los derechos digitales. La principal conclusión compartida entre el Gobierno y la OCDE fue que la transformación digital viene asociada a la velocidad de cambio por las tecnologías disruptivas (nube, big data, IoT, ciberseguridad, etc.). En relación a la necesidad de una normativa referente al mundo digital, Carlos López Blanco, director general de Asuntos públicos y regulación de Telefónica, apoyó la creación de normativa para la nueva economía digital con el propósito de garantizar la privacidad, fiscalidad y protección a los usuarios.

El 28 de marzo de 2017, el ministro de Energía, Turismo y Agenda Digital, Álvaro Nadal, presentó el Grupo de Trabajo sobre Derechos Digitales de los Ciudadanos con el cometido de *“realizar un diagnóstico de la situación actual sobre la protección de los derechos y obligaciones de los ciudadanos, promover un debate neutral, transparente, abierto e informado sobre los derechos digitales y contribuir a la agenda europea en esta materia”* (España, 2017). El Grupo de Trabajo se compone de expertos del ámbito académico y de la sociedad civil: José Luis Piñar, Tomás de la Quadra-Salcedo, Ricard Martínez, Diego Naranjo, Miguel Pérez Subías, Christoph Steck y Ofelia Tejerina; y abordará entre otras cuestiones la privacidad y los datos abiertos de los ciudadanos como unas de las preocupaciones en el ámbito digital.

Aunque la “Constitución digital” es un nuevo concepto en desarrollo, es importante tenerlo en cuenta para aportar el modelo resultante con el propósito de ayudar en la protección de los datos médicos.

1.3 Normativa aplicable en otros países

En esta sección se realiza un análisis, a nivel internacional, sobre protección de datos personales referentes a la salud. Esta protección de datos debe permitir identificar aquellos campos necesarios en los procesos de anonimización garantizando el anonimato del paciente en el momento del tratamiento de los datos clínicos. A lo largo de la sección, se agrupa cada país según el continente al que pertenezca, consiguiendo así, una lectura más cómoda y legible sobre la situación internacional.

En esta sección, se han realizado las traducciones oportunas de la legislación analizada con el fin de garantizar la comprensión del lector.

1.3.1 *África*

En la legislación del continente africano, existe una carencia significativa sobre la protección de datos personales ya que únicamente 20 países de los 54 que componen el continente africano, tienen leyes de protección de datos (constituye un 37,04% del total de países). En la Tabla 6 se recogen los países y la normativa analizada de protección de datos que no han aportado más atributos médicos de los descritos en los países detallados en esta sección.

País	Normativa
Benin	<i>Loi n° 2009-09 portant protection des données à caractère personnel en République du Bénin</i> (Benin, 2009)
Burkina Faso	<i>Loi n° 010-2004/AN Portant Protection des Donnees a Caractere Personnel</i> (Burkina Faso, 2004)
Cabo Verde	<i>Law 133/V/2001 Data Protection Act</i> (Cabo Verde, 2001)
Costa de Marfil	<i>Loi n° 2013 -450 du 19 juin 2013 relative à la protection des données à caractère personnel</i> (Costa de Marfil, 2013)
Egipto	Constitución aprobada el 15 de enero de 2014 (Egipto, 2014) (Reconocido solo el derecho a la intimidad).
Emiratos Árabes Unidos: Dubai	<i>Data Protection Law</i> (Dubai, 2007)
Israel	<i>Protection of Privacy Law, 5741-1981</i> (Israel, 1981)
Mali	<i>Loi n° 2013-015 Portant Protection des Donnees a Caractere Personnel en Republique du Mali</i> (Mali, 2013)
Nigeria	<i>Guidelines on Data Protection</i> (Nigeria, 2013)
Qatar	<i>Regulation N°. 6 of 2005 QFC Data Protection Regulations</i> (Qatar, 2005)
Senegal	<i>Loi n° 2008-12 du 25 janvier 2008 portant sur la Protection des données à caractère personnel</i> (Senegal, 2008)
Seychelles	<i>Data Protection Act, 2003 (No. 9 of 2003)</i> (Seychelles, 2003)
Tunisia	<i>Organic Act n°2004-63 of July 27th 2004 on the protection of personal data</i> (Tunisia, 2004)
Zimbabwe	<i>Access to Information and Protection of Privacy Act</i> (Zimbabwe, 2002)

Tabla 6 Relación de países de África y su normativa en protección de datos

1.3.1.1 Angola

La ley de protección de datos de Angola es *Data Protection Law* (Angola, 2011) y aporta dos conceptos: (a) Define los datos sensibles relacionados con la salud, incluyendo los datos genéticos; y (b) Identifica la combinación de elementos específicos de identidad física, fisiológica, psíquica, económica, cultural o social. Esta especificación mantendría la misma tesis que la definición de los cuasi-identificadores.

1.3.1.2 Ghana

En el año 2012 se aprobó la *Data Protection Act, 2012* (Ghana, 2012). En su artículo 96 se definen los datos personales y además de lo indicado referente a campos de atributos como color o raza, también se definen campos de atributos como estado mental, condición mental, estado físico, estado médico, ADN y orientación sexual.

1.3.1.3 Madagascar

La ley de protección de datos *Loi n° 2014-038 Sur la protection des données à caractère personne* (Madagascar, 2014), define los datos de la salud dentro de la definición de datos personales, aunque incluye de manera particular de datos personales los datos genéticos y biométricos.

1.3.1.4 Mauricio

La ley de protección de datos de Mauricio *The Data Protection Act 2004* (Mauricio, 2004) o también denominada *Act 13/2004*, contiene dos definiciones relativas a la protección de datos. La primera define los datos personales y la segunda describe los datos personales sensibles. La definición relativa a datos personales únicamente hace referencia a los datos de un individuo que le permitan identificar exceptuando los relacionados con la salud: "Datos personales" significa: (a) Los datos personales que se refieren a una persona que puede ser identificado a partir de estos datos; o (b) Datos y otra información, incluyendo un dictamen que forme parte de una base de datos, grabados o no en una

forma material, acerca de un individuo cuya identidad es aparente o razonablemente puede determinarse a partir de los datos, información u opinión.

En cambio, los datos personales sensibles se definen como *“la información personal relativa a un titular de los datos que consiste en información sobre: [...] (e) La salud física o mental”*. Esta definición, aunque intenta abarcar la salud física o mental como una definición de información personal, no aporta atributos médicos que puedan ser anonimizados debido al nivel de abstracción de la definición.

1.3.1.5 Sudáfrica

En Sudáfrica, la ley que protege los datos personales es *Electronic Communications and Transactions Act, 2002* (Sudáfrica, 2002) aprobada el 31 de julio de 2002 y denominada como No. 25 of 2002. En esta ley se definen los siguientes aspectos:

"Información personal significa información sobre un individuo identificable, incluyendo, pero no limitado a:

- (a) Información relativa a la raza, género, sexo, embarazo, estado civil, nacionalidad, origen étnico o social, color, orientación sexual, edad, salud física o mental, el bienestar, discapacidad, religión, conciencia, creencia, cultura, idioma y fecha de nacimiento;*

- (b) Información relativa a la educación o la historia médica, penal o laboral de la persona o la información relativa a las transacciones financieras en la que el individuo ha participado;*
- (c) Cualquier número, símbolo identificativo, u otro dato asignado a la persona;*
- (d) La dirección, las huellas digitales o el tipo sanguíneo del individuo;*
- (e) Las opiniones personales, opiniones o preferencias del individuo, salvo si son de otra persona o sobre una propuesta de subvención, una gratificación o un premio que se ha hecho a otra persona;*
- (f) La correspondencia enviada por la persona que es implícita o explícitamente de naturaleza privada o confidencial o más correspondencia que pudiera revelar el contenido de la correspondencia inicial;*
- (g) Los puntos de vista u opiniones de otra persona sobre el individuo;*
- (h) Los puntos de vista y opiniones de otra persona acerca de una propuesta de subvención, gratificación o premio que se hagan a la persona, pero sin incluir el nombre de la persona en la que aparece con los puntos de vista u opiniones de la otra persona; y*
- (i) El nombre de la persona en la que aparece con otra información personal relacionada con el individuo o en que la divulgación del nombre propio revelaría información sobre el individuo,*

Pero excluye la información sobre un individuo que ha muerto hace más de 20 años;”.

La Ley (Sudáfrica, 2002) especifica un conjunto de campos que deben ser protegidos (véase Tabla 7).

Información personal en Sudáfrica		
Raza	Nacionalidad	Educación
Género	Origen étnico o social	Información laboral
Sexo	Color	Dirección
Embarazo	Orientación sexual	Huellas dactilares
Estado civil	Edad	Opiniones
Religión	Salud física o mental	Información penal
Discapacidad	Bienestar	Número identificativo
Conciencia	Idioma	Símbolo identificativo
Creencia	Cultura	Tipo sanguíneo
Fecha de nacimiento	Historia médica	Preferencias

Tabla 7 Campos de información personal protegidos en Sudáfrica (Sudáfrica, 2002)

Durante el presente periodo de investigación, Sudáfrica ha desarrollado la ley *Protection of Personal Information Act, 2013* (Sudáfrica, 2013). Con respecto a la ley de 2002 anteriormente analizada se ha incluido la definición de datos biométricos en los que ha englobado los campos de datos médicos: tipo sanguíneo, huella dactilar, análisis de ADN, escáner de retina y reconocimiento de voz. Además, en la definición de datos personales ha incluido en el apartado C de la definición los campos de información: número de teléfono, información de ubicación e identificador de línea. En el apartado D se ha incluido la información biométrica como dato personal. Por lo tanto, se han incluido nuevos atributos de información como se muestran en la Tabla 8.

Nueva información personal en Sudáfrica	
Escáner de retina	Reconocimiento de voz
Análisis de ADN	Número de teléfono
Identificador de línea	Información de ubicación

Tabla 8 Nuevos campos de información personal protegidos en Sudáfrica (Sudáfrica, 2013)

1.3.1.6 Marruecos

La Ley nº 09-08, de 18 de febrero de 2009 (Marruecos, 2009b) y su desarrollo con el Decreto nº 2-09-165 de 21 de mayo de 2009 (Marruecos, 2009a) es la que protege los datos personales.

En el artículo 1.3 de dicha ley, se hace referencia a los datos personales sensibles relacionados con la salud, tal y como se indica a continuación: *“Datos sensibles: los datos personales que revelen el origen racial o étnico, opiniones políticas, creencias religiosas o filosóficas, la afiliación sindical de la persona interesada o que se refieren a la salud incluyendo los datos genéticos;”* (ver Tabla 9). Sin embargo, no quedan definidos qué datos están relacionados con el ámbito sanitario.

Datos sensibles de Marruecos	
Origen racial	Creencias religiosas
Origen étnico	Creencias filosóficas
Opiniones políticas	Afiliación sindical
Datos genéticos	

Tabla 9 Datos sensibles según Marruecos (Marruecos, 2009b)

1.3.2 Europa

En este apartado se describen los países del continente europeo. Como se puede apreciar que hay dos grupos diferentes de países: los que han transpuesto directamente la Directiva 95/46/CE (Tabla 10), y los que han modificado y adaptado la Directiva 95/46/CE a sus necesidades nacionales, definiendo campos de información relacionados con la salud.

País	Normativa
Alemania	<i>Federal Data Protection Act</i> (Alemania, 2003)
Andorra	<i>LLei 15/2003</i> (Andorra, 2004)
Austria	<i>Federal Act concerning the Protection of Personal Data (DSG2000)</i> (Austria, 1999)
Bélgica	<i>Act of 8 December 1992 on the protection of privacy in relation to the processing of personal data</i> (Bélgica, 1992)
	<i>Royal Decree implementing the Act of 8 December 1992 on the protection of privacy in relation to the processing of personal data</i> (Bélgica, 2001)
Bielorrusia	<i>Law on Information, Informatization and Protection of Information</i> (Bielorrusia, 2008)
Bosnia y Herzegovina	<i>Law on the Protection of Personal Data</i> (Bosnia and Herzegovina, 2006)
Bulgaria	<i>Law For Protection of Personal Data</i> (Bulgaria, 2002)
Chipre	<i>The processing of personal data (protection of individuals) 138/2001</i> (Chipre, 2001)
Croacia	<i>The Act on Personal Data Protection</i> (Croacia, 2003)
Mónaco	<i>Data Protection Law n° 1.165 of 23 December 1993</i> (Mónaco, 1993)
República Checa	<i>Personal Data Protection Act 101</i> (República Checa, 2000)
Dinamarca	<i>The Act on Processing of Personal Data N° 429</i> (Dinamarca, 2000)
Estonia	<i>Personal Data Protection Act</i> (Estonia, 2007)
Francia	<i>Loi n° 2004-801</i> (Francia, 2004)
Gibraltar	<i>Data Protection Act No. 2004-01</i> (Gibraltar, 2004)
Grecia	<i>Law 2472/1997 on the Protection of Individuals with regard to the Processing of Personal Data</i> (Grecia, 1997)
Guernsey	<i>The Data Protection (Bailiwick of Guernsey)</i> (Guernsey, 2002)
Irlanda	<i>Data Protection Act 1988 y Data Protection (Amendment) Act 2003</i> (Irlanda, 1988; Irlanda, 2003)
Italia	<i>Italian personal Data Protection Code Legislative Decree n°. 196</i> (Italia, 2003)
Jersey	<i>Data Protection (Jersey) Law 2005</i> (Jersey, 2005)
Letonia	<i>Protection of Personal Data of Natural Persons</i> (Letonia, 2000)
Luxemburgo	<i>Protection des personnes à l'égard du traitement des données à caractère personnel</i> (Luxemburgo, 2002)
Eslovaquia	<i>Act No. 428/2002 Coll. on Protection of Personal Data</i> (Eslovaquia, 2002)
República de Eslovenia	<i>Personal data protection act of the Republic of Slovenia</i> (República de Eslovenia, 2004)
Holanda	<i>Personal Data Protection Act</i> (Holanda, 1999)
Hungría	<i>Act CXII of 2011</i> (Hungría, 2011)
Malta	<i>Chapter 440 Data Protection Act</i> (Malta, 2001)
Macedonia	<i>Law on Personal Data Protection</i> (Macedonia, 2005)
Montenegro	<i>Personal Data Protection Law</i> (Montenegro, 2008)

País	Normativa
Noruega	<i>Personal Data Act y Personal Health Data Filing System Act</i> (Noruega, 2000; Noruega, 2001)
Polonia	<i>Act of August 29, 1997 on the Protection of Personal Data</i> (Polonia, 1996)
Portugal	<i>Act 67/98</i> (Portugal, 1998)
Reino unido	<i>Data Protection Act 1998</i> (Reino Unido, 1998)
Rumania	<i>Law No. 677/2001 on the Protection of Individuals with Regard to the Processing of Personal Data and the Free Movement of Such Data</i> (Rumania, 2001)
Serbia	<i>Law on Personal Data Protection</i> (Serbia, 2008)
Suecia	<i>Personal Data Act 1998</i> (Suecia, 1998)
Suiza	<i>Federal Act on Data Protection</i> (Suiza, 1992)
Ucrania	<i>The Law of Ukraine No. 2297 VI "On Personal Data Protection" ("Data Protection Law")</i> (Ucrania, 2011)

Tabla 10 Relación de países de Europa y normativa nacional de protección de datos

A continuación, se describen los países que han aportado datos médicos a este trabajo de tesis.

1.3.2.1 Finlandia

En Finlandia la ley *Personal Data Act 523/1999* (Finlandia, 1999) en su capítulo 3 sección 11, prohíbe el procesado de datos sensibles. Además, indica el tipo de datos: el estado de la salud, enfermedad o discapacidad y tratamientos (Tabla 11). Los campos a proteger están definidos con un alto grado de abstracción sobre el conjunto de datos.

Datos sensibles relacionados con la salud de Finlandia	
Estado de salud	Enfermedad
Discapacidad	Tratamientos

Tabla 11 Datos sensibles aportados por Finlandia relacionados con datos la salud (Finlandia, 1999)

1.3.2.2 Lituania

En la ley *Law on Legal Protection of Personal Data No. I-1374* (Lituania, 1996), en su artículo 10.1, hace hincapié sobre los datos de la salud, descomponiendo éstos en el estado, diagnóstico, pronóstico y tratamiento entre otros campos (Tabla 12).

Datos personales sobre la salud de Lituania	
Estado	Diagnóstico
Tratamiento	Pronóstico

Tabla 12 Datos sanitarios definidos por Lituania (Lituania, 1996)

1.3.2.3 Islandia

En Islandia la ley *Act on the Protection of Privacy as regards the Processing of Personal Data* (Islandia, 2000) en su artículo 2 definición 8 sobre datos sensibles, aporta tres campos de datos englobados dentro de los datos de la salud (ver Tabla 13).

Datos sensibles definidos por Islandia	
Datos de uso de alcohol	Datos de consumo de narcóticos
Datos de consumo de drogas médicas	

Tabla 13 Datos nuevos aportados por Islandia (Islandia, 2000)

1.3.3 Oceanía

El continente de Oceanía se caracteriza por tener solo dos leyes nacionales. Australia contiene una legislación relacionada con la

protección de datos donde quedan definidos, de forma específica, aquellos datos identificables dentro de los personales. Sin embargo, en Nueva Zelanda no se define ningún tipo de datos con el mismo nivel de detalle. Tampoco se contempla la clasificación de datos sensibles para salvaguardar la protección de los datos sanitarios.

1.3.3.1 Australia

La ley de protección de datos es *Privacy Act No. 119*, de 1988, (Australia, 1988). En la parte 2 división 1 de definiciones generales y sección 6, se definen los campos que pertenecen a la información identificable sobre un individuo (Tabla 14).

Información identificable de Australia	
Nombre	Fecha de nacimiento
Alias o nombre anterior	Sexo
Dirección o direcciones anteriores	Nombre del empleador del individuo
Nº de licencia de conducir	

Tabla 14 Atributos identificables categorizados por Australia (Australia, 1988)

En la Tabla 15, se muestran los campos que se pueden obtener de la definición de información sanitaria (Australia, 1988):

Información médica de Australia	
Estado de salud	Discapacidad
Deseos expresados sobre el servicio de salud futuro	Otra información recopilada
Donación de partes de cuerpo, órganos o sustancias coporales	Información genética

Tabla 15 Campos médicos definidos por Australia (Australia, 1988)

En la definición de información sensible se podría equiparar a la definición propuesta por la Directiva Europea 95/46/CE por lo que no se va a entrar en detalle por su falta de concreción.

1.3.3.2 Nueva Zelanda

La ley de protección de datos es *Privacy Act 1993*, de 17 de mayo de 1993 (Nueva Zelanda, 1993) y se caracteriza por no definir los datos personales propios del campo de la salud como datos personales ni como datos sensibles.

1.3.4 Asia

En el continente asiático, la situación es muy distinta a los continentes analizados hasta el momento. Esto se debe a la existencia, por ejemplo, de países como China, que tienen ley de protección de datos sin clasificación alguna sobre los mismos, y además, no profundizan sobre su protección.

Los países se han agrupado, en dos categorías, atendiendo a características comunes en protección de datos:

- Grupo 1: Los países que teniendo ley de protección de datos hacen referencia a los datos relativos a la salud. Sin embargo, estos países no especifican los datos considerados como sanitarios (Tabla 16).

- Grupo 2: Se componen de los países con ley de protección de datos sin contemplar los datos de la salud como datos personales (Tabla 17).

País	Normativa
Corea del Sur	<i>Act on Promotion of Information and Communication Network Utilization and Information Protection</i> (Corea del Sur, 2011)
Hong Kong	<i>Personal data (privacy) ordinance</i> (Hong Kong, 1995)
India	<i>The Information Technology Act 2000</i> (India, 2000)
Japón	<i>Act on the Protection of Personal Information (Act No. 57 of 2003)</i> (Japón, 2003)
Filipinas	<i>Republic Act No. 10173</i> (Filipinas, 2011)
Macao	<i>Lei n.º 8/2005 da Protecção de Dados Pessoais</i> (Macao, 2005)
Malasia	<i>Personal Data protection act 2010</i> (Malasia, 2010)
Rusia	<i>Federal Law of 27 July 2006 N 152-FZ on Personal Data</i> (Rusia, 2006)
Taiwán	<i>Personal Information Protection Act</i> (Taiwán, 2010)
Tailandia	<i>Official Information Act, B.E. 2540 (1997) for state agencies</i> (Tailandia, 1997)
Turquía	<i>Data Protection Law number 6698</i> (Turquía, 2016)

Tabla 16 Grupo 1 de países asiáticos

País	Normativa
China	<i>National Standard Guiding Technical Document of the People's Republic of China</i> (China, 2012)
Indonesia	<i>Regulation of the government of the republic of Indonesia number 82 of 2012</i> (Indonesia, 2012)
Singapur	<i>Personal Data Protection Act 2012 (No. 26 of 2012)</i> (Singapur, 2012)
Vietnam	<i>Law on Electronic Transactions (Draft: Finalized in 2006)</i> (Vietnam, 2005)

Tabla 17 Grupo 2 de países asiáticos

1.3.5 América

En el continente americano se ha detectado una gran diferencia entre la normativa de América del Norte y el resto partes de América. Para demostrar esta afirmación es necesario analizar la ley de protección de

datos de Canadá (Canadá, 1985; Canadá, 2000) y Estados Unidos (Estados Unidos, 1996), donde se protegen los datos personales relativos a la salud y se definen aquellos datos que son englobados en esa categoría de datos personales. En Estados Unidos se proporciona un listado detallado de los datos que deben ser protegidos.

En América Central y América del Sur el ámbito de definición va subiendo en el nivel de abstracción como ha sucedido en Europa con la Directiva 95/46/CE, aunque hay alguna excepción como el país Trinidad y Tobago o el estado mexicano de Nuevo León en los que incluyen unos campos dentro de la definición de datos personales de la salud.

La Tabla 18 muestra los países que tienen ley de protección del continente americano y además, protegen los datos relativos a la salud. Sin embargo, no definen la información que componen los datos.

País, Estado o Ciudad	Normativa
Argentina	Ley 25.326 de Protección de Datos Personales (Argentina, 2000)
Buenos Aires (Argentina)	Ley 1.845 de Protección de Datos Personales o Ley Caba (Buenos Aires, 2006)
Chile	Ley 19.628 de Protección de Datos de Carácter Personal (Chile, 1999)
Colombia	Ley Estatutaria 1581 y el Decreto 1377 (Colombia, 2012; Colombia, 2013)
Costa Rica	Ley Nº 8968 de Protección de la Persona Frente al Tratamiento de sus Datos Personales y el Reglamento a la Ley de Protección de la Persona frente al Tratamiento de sus Datos Personales (Costa Rica, 2011; Costa Rica, 2013)
Honduras	Ley de transparencia y acceso a la información pública Decreto No. 170-2006 (Honduras, 2006)
México	Ley Federal de Protección de Datos Personales en Posesión de los Particulares (México, 2010)

País, Estado o Ciudad	Normativa
Aguascalientes (México)	Ley de Transparencia e Información Pública del Estado de Aguascalientes (Aguascalientes, 2006)
Baja California (México)	Ley de Acceso a la Información Pública para el Estado de Baja California (Baja California, 2010)
Baja California Sur (México)	Ley de Transparencia y Acceso a la Información Pública para el estado de Baja California Sur (Baja California Sur, 2010)
Campeche (México)	Ley de Transparencia y Acceso a la Información Pública del Estado de Campeche (Campeche, 2012)
Chihuahua (México)	Ley de Transparencia y Acceso a la Información Pública del Estado de Chihuahua (Chihuahua, 2005)
Coahuila (México)	Ley de Acceso a la Información Pública y Protección de Datos Personales para el Estado de Coahuila (Coahuila de Zaragoza, 2014)
Colima (México)	Ley de Transparencia y Acceso a la Información Pública del Estado de Colima y Ley de Protección de Datos del Estado de Colima (Colima, 2003; Colima, 2013)
Distrito Federal (México)	Ley de Transparencia y Acceso a la Información Pública del Distrito Federal derogada por Ley de Transparencia y Acceso a la Información Pública del Distrito Federal durante el desarrollo de esta tesis sin aportar cambios en la misma en lo que afecta al presente trabajo de tesis, y Ley de Protección de Datos del Distrito Federal (México, 2008a; México, 2008b; México, 2016).
Durango (México)	Ley de Transparencia y Acceso a la Información Pública del Estado de Durango derogada por la ley del mismo nombre pero de 4 mayo de 2016 publicada durante el desarrollo de esta tesis sin aportar cambios en la misma en lo que afecta al presente trabajo de tesis, y Ley de Protección de Datos del Estado de Durango (Durango, 2008; Durango, 2013; Durango, 2016)
Guanajuato (México)	Ley de Acceso a la Información Pública para el Estado y los Municipios del Estado de Guanajuato derogada por la Ley de Transparencia y Acceso a la Información Pública para el Estado de Guanajuato y la Ley de Protección de Datos Personales para el Estado y los Municipios de Guanajuato (Guanajuato, 2012; Guanajuato, 2016a; Guanajuato, 2016b)
Guerrero (México)	Ley de Acceso a la Información Pública del Estado de Guerrero (Guerrero, 2010)
Hidalgo (México)	Ley de Transparencia y Acceso a la Información Pública Gubernamental del Estado de Hidalgo (Hidalgo, 2006)
Jalisco (México)	Ley de Transparencia e Información Pública del Estado de Jalisco (Jalisco, 2013)
Michoacán (México)	Ley de Transparencia y Acceso a la Información Pública del Estado de Michoacán de Ocampo derogada por la Ley identificada con el Decreto N° 144 (Michoacán, 2008;

País, Estado o Ciudad	Normativa
	Michoacán, 2016)
Morelos (México)	Ley de Información Pública, Estadística y Protección de Datos Personales del Estado de Morelos (Morelos, 2003)
Nayarit (México)	Ley de Transparencia y Acceso a la Información Pública del Estado de Nayarit derogada por la Ley del mismo nombre de 2016 (Nayarit, 2007; Nayarit, 2016)
Oaxaca (México)	Ley de Transparencia y Acceso a la Información Pública del Estado de Oaxaca derogada por la Ley del mismo nombre en 2016 y la Ley de Protección de Datos Personales del Estado de Oaxaca (Oaxaca, 2008a; Oaxaca, 2008b; Oaxaca, 2016)
Puebla (México)	Ley de Transparencia y Acceso a la Información Pública del Estado de Puebla (Puebla, 2011)
Querétaro (México)	Ley Estatal de Acceso a la Información Gubernamental en el Estado de Querétaro (Querétaro, 2002)
Quintana Roo (México)	Ley de Transparencia y Acceso a la Información Pública del Estado de Quintana Roo derogada por la Ley del mismo nombre del año 2016 (Quintana Roo, 2008; Quintana Roo, 2016)
San Luis Potosí (México)	Ley de Transparencia y Acceso a la Información Pública del Estado de San Luis Potosí derogada por la Ley del mismo nombre de 2016 (San Luis de Potosí, 2007; San Luis de Potosí, 2016)
Sinaloa (México)	Ley de Acceso a la Información Pública del Estado de Sinaloa (Sinaloa, 2008)
Sonora (México)	Ley de Acceso a la Información Pública del Estado de Sonora derogada por la Ley de Transparencia y Acceso a la Información Pública del Estado de Sonora (Sonora, 2005; Sonora, 2016)
Tabasco (México)	Ley de Transparencia y Acceso a la Información Pública del Estado de Tabasco (Tabasco, 2015)
Tlaxcala (México)	Ley de Acceso a la Información Pública y Protección de Datos Personales para el Estado de Tlaxcala (Tlaxcala, 2012)
Tamaulipas (México)	Ley de Información Pública del Estado de Tamaulipas derogada por la Ley de Transparencia y Acceso a la Información Pública del Estado de Tamaulipas (Tamaulipas, 2013; Tamaulipas, 2016)
Veracruz (México)	Ley de Transparencia y Acceso a la Información Pública para el Estado de Veracruz de Ignacio de la Llave (Veracruz, 2007)
Zacatecas (México)	Ley de Acceso a la Información Pública del Estado de Zacatecas (Zacatecas, 2011)
Nicaragua	Ley de Protección de Datos Personales (Nicaragua, 2012)
Paraguay	Ley N° 1682 y su modificación por la Ley N° 1969 (Paraguay, 2001; Paraguay, 2002)
Perú	Ley N° 29733 y el Reglamento de la Ley N° 29733 (Perú, 2011; Perú, 2013)
República	Ley 126-02 sobre protección de datos de carácter personal por

País, Estado o Ciudad	Normativa
Dominicana	los sujetos regulados (República Dominicana, 2013)
Uruguay	Protección de Datos Personales y Acción Habeas Data y el Decreto nº 414/009 (Uruguay, 2008; Uruguay, 2009)

Tabla 18 Relación de países de América y normativa de protección de datos

Durante el desarrollo de la presente tesis, el estado de México realizó la Ley General de Transparencia y Acceso a la Información Pública (México, 2015). Esta ley es de aplicación a las entidades federadas definidas en el artículo 3 definición VIII: Aguascalientes, Baja California, Baja California Sur, Campeche, Coahuila de Zaragoza, Colima, Chiapas, Chihuahua, Durango, Guanajuato, Guerrero, Hidalgo, Jalisco, México, Michoacán, Morelos, Nayarit, Nuevo León, Oaxaca, Puebla, Querétaro, Quintana Roo, San Luis Potosí, Sinaloa, Sonora, Tabasco, Tamaulipas, Tlaxcala, Veracruz, Yucatán, Zacatecas y el Distrito Federal. Sin embargo, esta ley no incluye en su definición los datos de la salud, por lo que se ha optado por mantener las leyes de cada Estado previas a la (México, 2015).

1.3.5.1 Canadá

La ley de protección de datos es *Privacy Act de 1985* (Canadá, 1985) en la que en su apartado “3. Définitions” incluye la definición de información personal. Esta definición incluye algunos campos de información que aunque no todos son aplicables al campo sanitario, sí que pueden ser posibles candidatos a ser incluidos (Tabla 19).

Información personal en Canadá		
Raza	Origen étnico	Correspondencia
Nacionalidad	Color	Opiniones personales
Religión	Edad	Tipo sanguíneo

Información personal en Canadá		
Estado civil	Educación	Huellas dactilares
Historia clínica	Penal	Dirección
Laboral	Financiera	Número de identificación
Símbolo de identificación		

Tabla 19 Información personal definida por Canadá (Canadá, 1985)

La ley *Personal Information Protection and Electronic Documents Act*, de 1 de junio de 2009 (Canadá, 2000), en su artículo 2 apartado 1 define la información personal relativa a la salud (Tabla 20):

- (a) Información relativa a la salud física o mental de la persona;
- (b) Información relativa a cualquier servicio de Salud proporcionado al individuo;
- (c) Información relativa a la donación por parte del individuo de cualquier parte del cuerpo o cualquier sustancia corporal del individuo o de la información derivada de las pruebas o exámenes de una parte del cuerpo o sustancia corporal del individuo;
- (d) Información que se recoge en el curso de la prestación de servicios de salud a la persona; o
- (e) Información que se recaba de paso a la prestación de servicios de salud a la persona.

Atributos médicos de Canadá	
Salud física	Salud mental
Servicio proporcionado al individuo	Donación de parte de cuerpo o sustancia corporal
Información derivada de pruebas o exámenes de parte de cuerpo o sustancia corporal	Información de una prestación en curso al individuo
Información posterior a la prestación del servicio de salud	

Tabla 20 Información relativa con la salud según Canadá (Canadá, 2000)

1.3.5.2 Estados Unidos

La ley de protección de datos sanitarios de Estados Unidos es HIPAA, de 21 de agosto de 1996 (Estados Unidos, 1996). Esta ley incluye la definición de “información de la salud” e “información identificable de un individuo”. Además, define “información de la salud protegida” como aquellos campos que permiten identificar a un individuo y que deben ser anonimizados para garantizar la privacidad y protección del mismo. El listado de atributos definidos por HIPAA son los siguientes:

- (a) Nombres;
- (b) Todas las subdivisiones geográficas menores que un Estado, incluyendo dirección, ciudad, condado, distrito, código postal y sus equivalentes en códigos de geolocalización, excepto para los tres dígitos iniciales de un código postal si, de acuerdo a los datos disponibles al público en la Oficina del Censo:
 1. La unidad geográfica formada por la combinación de todos los códigos postales con los mismos tres dígitos iniciales contiene más de 20.000 personas; y
 2. Los tres dígitos iniciales de un código postal para todas estas unidades geográficas que contengan 20.000 o menos personas se cambia a 000.
- (c) Todas las fechas (excepto el año) para las fechas directamente relacionadas a un individuo, incluyendo fecha de nacimiento, fecha de admisión, fecha de alta, fecha de la muerte; y todas las edades mayores de 89 y todos los elementos de indicativo (incluyendo año) las fechas de dicha edad, salvo que tales

elementos y edades pueden agruparse en una sola categoría de mayores de 90 años;

- (d) Números de teléfono;
- (e) Números de fax;
- (f) Correo electrónico;
- (g) Número de Seguro Social;
- (h) Números de registros médicos;
- (i) Números de los beneficiarios del plan de salud;
- (j) Números de cuenta;
- (k) Números certificado/licencia;
- (l) Identificadores de vehículo y números de serie, incluidos los números de matrícula;
- (m) Dispositivos identificadores y números de serie;
- (n) Localizadores de recursos universales web (URL);
- (o) Números de dirección *Internet Protocol* (IP);
- (p) Biométricos, como huellas dactilares y voz;
- (q) Imágenes fotográficas con el rostro completo y cualquier imagen comparable; y
- (r) Cualquier otro número de identificación, característica o código.

Tras analizar el listado de información a anonimizar por HIPAA, se concluye que HIPAA aporta un amplio conjunto de atributos médicos que deben ser protegidos mediante anonimización.

Además, se han analizado los estándares de la *Health Level Seven International* (en adelante HL7) (Health Level Seven International, 2017) y DICOM centrados en el manejo de información. HL7 define la

estructura de los documentos clínicos para el intercambio de información médica entre aplicaciones. Por su parte DICOM ha sido desarrollado para controlar la manipulación de información relacionada con imágenes médicas. Aunque ambos estándares aportan un listado ingente de datos médicos, no posee ninguna clasificación de protección de datos. Por ello, no se han incluido todos los datos médicos que aportan, aunque podría ser de utilidad la aplicación de la metodología desarrollada al caso práctico de Estados Unidos.

1.3.5.3 *Trinidad y Tobago*

La ley *Act No. 13 of 2011*, aprobada el 3 de junio de 2011 (Trinidad y Tobago, 2011), además de proteger los datos médicos en la definición de *personal information* (apartado b), también indica (apartado h de la misma definición) la necesidad de conocer los datos de huellas dactilares, ADN, grupo sanguíneo y datos biométricos (Tabla 21).

Datos relativos con la salud de Trinidad y Tobago	
Huellas dactilares	ADN
Grupo sanguíneo	Datos biométricos

Tabla 21 Datos relativos con la salud según Trinidad y Tobago (Trinidad y Tobago, 2011)

1.3.5.4 *Nuevo León*

La Ley de Transparencia y Acceso a la Información del Estado de Nuevo León, cuya última reforma se realizó el 17 de septiembre de 2012 (Nuevo León, 2008). Esta ley sigue un cauce diferente al resto de países mexicanos y a muchos países de Sudamérica. Además, dicha ley refleja

los campos contenidos en el país de Trinidad y Tobago en su artículo 79 sección III y define, datos de alto nivel. Queda definidos los datos de la salud como sigue: *“Estado de salud, historial clínico, alergias, enfermedades, información relacionada con cuestiones de carácter psicológico y/o psiquiátrico, incapacidades médicas, intervenciones quirúrgicas, vacunas, consumo de sustancias tóxicas, uso de aparatos oftalmológicos, ortopédicos, auditivos, prótesis, entre otros.”* (Tabla 22).

Atributos médicos de Nuevo León	
Estado de salud	Historial clínico
Alergias	Enfermedades
Datos psicológico	Datos psiquiátricos
Incapacidades médicas	Intervenciones quirúrgicas
Vacunas	Consumo de sustancias tóxicas
Aparatos oftalmológicos	Aparatos ortopédicos
Aparatos auditivos	Prótesis

Tabla 22 Datos relativos con la salud según Nuevo León (Nuevo León, 2008)

En esta Ley se añaden nuevos campos que deben ser protegidos dentro del ámbito de la salud. Hasta el momento dichos campos no habían sido protegidos de manera concreta. Algunos ejemplos característicos de estos campos son: alergias, incapacidades médicas, intervenciones quirúrgicas, vacunas, consumo de sustancias tóxicas, uso de aparatos oftalmológicos, ortopédicos, auditivos y prótesis entre otros.

En la nueva ley del estado de Nuevo León (Nuevo León, 2016) se han definido nuevos campos de información relativos a la salud (Tabla 23).

Nuevos atributos médicos de Nuevo León	
Salud física	Salud mental
Preferencias sexuales	Huella digital
Ácido desoxirribonucleico (ADN)	Fotografía
Número de la seguridad social	

Tabla 23 Datos relativos con la salud con la nueva normativa de Nuevo León (Nuevo León, 2016)

1.4 Estudios relacionados con la protección de datos médicos

Con el significativo avance de las TIC en la última década se ha producido un cambio en el modelo de gestión de los datos, de tal envergadura, que ha pasado de ser almacenada en papel a ser almacenada en bases de datos informatizadas. Además, el modelo de gestión contiene acceso a nivel global para que toda la información esté disponible en el momento que sea requerida (Castilla y León, 2014). Esto, requiere cumplir tanto la normativa aplicable en cada momento como las medidas de seguridad necesarias que permitan garantizar un nivel de seguridad suficiente, y además, garantice la protección del paciente. No siempre se consigue la protección de la información; por ejemplo, se ha dado algún caso en el que un impostor se hace pasar por un paciente con los datos conseguidos en ficheros de pacientes (previamente publicados) con el fin de conseguir medicamentos (Dixon, 2006). La información sanitaria protegida, o *protected health information* (PHI), es una fuente principal para el robo de identidad del paciente (Dadisho, 2005), siendo utilizada en delitos tipificados en las leyes.

Para prevenir posibles fugas de información (por ejemplo, información médica sobre datos sanitarios), es necesario anonimizar todo tratamiento de datos fuera del entorno seguro en el que deben estar almacenados. Esta anonimización es de gran importancia por ejemplo, para permitir la utilización de los datos con fines de investigación, tal y como queda recogido en los siguientes ejemplos:

- 750.000 datos anónimos han sido utilizados para concluir que las dietas “milagro” perjudican la salud (Lang, De Sterck, & Abrams, 2016).
- 451.232 datos anónimos han sido estudiados para obtener una relación entre la obesidad y la demencia (Wotton & Goldacre, 2014).
- 40.000 datos anónimos se han estudiado para concluir que un aumento de plaquetas puede ser un predictor de cáncer (Bailey, Ukoumunne, Shephard, & Hamilton, 2017).
- A partir de los datos anónimos recogidos durante las visitas médicas en Reino Unido, almacenados en “*Health Improvement Network (THIN) database*” se ha creado un algoritmo que permite predecir el riesgo de demencia en las visitas rutinarias al médico de cabecera (Walters et al., 2016).
- 129 datos médicos anónimos han permitido investigar los efectos adversos de la ingesta de hierro (Plasencia Cano, 2014).
- 40.000 datos médicos han permitido demostrar la efectividad del medicamento apixaban en problemas cardiovasculares (Amin et al., 2017).

Pero la anonimización no es utilizada solo en el ámbito sanitario, sino que se presenta como una solución de seguridad en protección de datos en Big Data o IoT aunque aún está lejos de ser aplicada por la falta de consenso entre los distintos actores implicados, como los fabricantes o los legisladores (INCIBE, Red.es, & Huawei, 2017). Dicho informe recoge distintas iniciativas de estandarización en aspectos de seguridad en IoT y señala la urgencia de consenso entre fabricantes y legisladores para ayudar a los usuarios en su protección de datos ante el rápido crecimiento de esta nueva industria. A esto, además se suman la evolución de las aplicaciones en Big Data (Sirera, 2015) que predicen del comportamiento de las personas desde distintos enfoques y que se clasifican en (Morte Ferrer, 2017):

- *Scoring*: Extraer el comportamiento ante diferentes tipos de contratos bancarios.
- *Profiling*: Cruzar datos no relacionados para obtener perfiles detallados de personas o grupos de personas.
- *Personalizing*: Obtener valoraciones sobre características de una persona.
- *Tracking*: Trazar el rastro de una persona a partir de las evidencias que ha dejado en sus comunicaciones, como por ejemplo por Internet.

A modo de ejemplo existen dos casos en los que mediante la aplicación de técnicas de minería de datos predictivas en Big Data, se ha podido extraer un perfil de una persona o grupo de personas a partir de datos no relacionados:

- Un grupo de científicos han intentado identificar al artista contemporáneo con seudónimo Banksy mediante la aplicación del modelo matemático de Dirichlet a partir de los datos geográficos (*geographic profiling*) de sus obras, obteniendo el nombre de Robin Gunningham como posible artista. También se propone la utilización de estos modelos para predecir bases terroristas y prevenir atentados (Hauge, Stevenson, Kim Rossmo, & Le Comber, 2016).
- Desarrollo de un algoritmo predictivo para identificar a los pacientes con alto riesgo de readmisión en los próximos 12 meses a partir de los datos de los años 1999-2000. Para realizar esta predicción se utilizaron los siguientes campos de información: edad, sexo, origen étnico, número de ingresos previo y condición clínica (Billings, Dixon, Mijanovich, & Wennberg, 2006).

Como se puede ver, el análisis predictivo en Big Data identifica patrones de comportamiento sobre un grupo de personas que puede ser aplicados a individuos concretos y aunque haya situaciones en las que no se incurren en prácticas ilegales en protección de datos, pueden llegar a producir una situación discriminatoria hacia una persona (Espino Timón, 2017).

Un primer intento para contemplar aquellos campos que deben de ser anonimizados es elegir todos los existentes. Sin embargo, esta aproximación conlleva una pérdida importante de información relevante en la extracción de conocimiento a partir de los datos recogidos por los

pacientes. Por este motivo, una gran cantidad de leyes, que se han expuesto con anterioridad, no recogen un listado de campos obligatorios a ser anonimizados.

Existen diferentes métodos para anonimizar atributos PHI como, por ejemplo, el uso de diccionarios. Los algoritmos que hacen uso de diccionarios alcanzan un notable rendimiento en su aplicación a atributos PHI. Sin embargo, la utilización de este método conlleva un gran esfuerzo dado que es necesario una continua actualización de diccionarios y su generalización es clasificada como compleja (Meystre, Friedlin, South, Shen, & Samore, 2010), por ello, es recomendable la utilización de más métodos. Existen también, métodos basados en aprendizaje automático con expresiones regulares sin incluir categorías de PHI. Con posterioridad, se desarrollaron sistemas que mejoraron el rendimiento mediante combinación de reglas para algunos atributos PHI y aprendizaje automático para otros. Otro intento por la mejora de rendimiento se centraba en algoritmos basados, únicamente, en aprendizaje automático sin expresiones regulares y también aquellos basados en reglas (Uzuner, Luo, & Szolovits, 2007). Sin embargo, todos estos métodos necesitan una definición de atributos que deben ser anonimizados, motivo por el cual se justifica esta sección.

En el año 2002, uno de los primeros estudios en anonimización de atributos (Thomas, Mamlin, Schadow, & McDonald, 2002) identifica el problema en el que los algoritmos que utilizan diccionarios, no poseen un amplio listado de nombres de personas que deben ser anonimizados. La solución que proponen los autores, es la localización de sufijos tales

como Mrs., Dr. o PhD. Estos sufijos ayudan en la localización de pares (sufijo, nombre) que permitan la anonimización de los nombres o apellidos de personas, debido a que podrían ser identificables. Este grupo de personas incluye tanto a pacientes como a todo el personal sanitario incluido en los datos clínicos.

La aportación que realiza otro investigador en este campo (El Emam et al., 2006) es la distinción de tres tipos de variables que agrupan los campos de datos. Los tipos de datos considerados son los siguientes:

- Identificadores: atributos que pueden identificar directamente a los individuos, tales como nombre, dirección de correo electrónico, número de teléfono, dirección postal, número de seguro social y número de tarjeta sanitaria.
- Cuasi-identificadores: Identificadores que no identifican directamente a una persona, pero pueden jugar un papel importante cuando se combinan entre ellos mediante la vinculación de las bases de datos externas que contienen variables de identificación.
- No identificadores: Son atributos que no identifican por sí solos ni en combinación con otros atributos a un individuo.

En otros estudios presentados (El Emam et al., 2006; Sweeney, 2002), se afirma que si se posee el siguiente escenario: (a) una base de datos con atributos identificables (también denominados “microdatos” (Agencia Española de Protección de Datos, 2016)) tales como el nombre, la dirección, el número de teléfono, las iniciales, el género, el código postal o la fecha de nacimiento; (b) creación una base de datos de investigación

en la que se eliminen los datos identificables, por ejemplo, nombre, dirección y número de teléfono; (c) una tercera base de datos en la que se pueda buscar una combinación de atributos como, por ejemplo, el género, el código postal y la fecha de nacimiento; entonces se obtendría un riesgo alto de identificación de la persona. Razón por la cual, estos atributos, que en un principio eran “no identificadores”, se denominan cuasi-identificadores (también denominados “datos de identificación indirecta” (Agencia Española de Protección de Datos, 2016)) como se puede ver en la representación gráfica de la Figura 1.

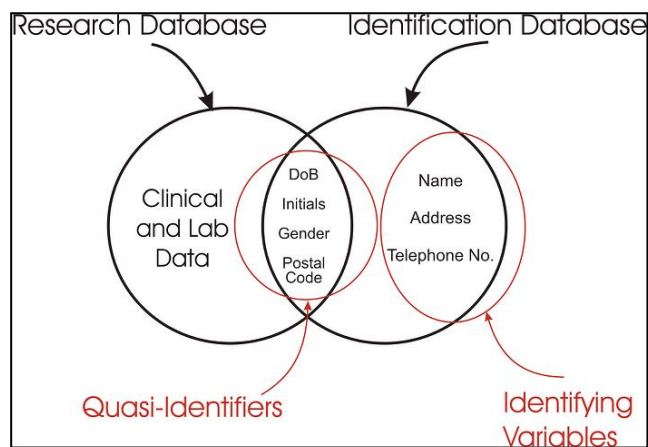


Figura 1 Primeros cuasi-identificadores identificados (El Emam et al., 2006)

Por tanto, además de los identificadores, es necesario considerar los cuasi-identificadores en el momento de obtener el listado de atributos a anonimizar que se deben anonimizar sin perder el conocimiento en el posterior análisis de los datos.

En (*Institute of Medicine (IOM), 2015*) se define un listado más completo de cuasi-identificadores tal como se muestra en la Tabla 24. Además, en (*Institute of Medicine (IOM), 2015*) queda recogida la demostración de reidentificación de usuarios mediante cuasi-identificadores.

Clasificación de cuasi-identificadores	
Sexo	Ingresos totales
Fecha de nacimiento / Edad	Condición de minoría visible
Códigos de geolocalización (códigos postales, censo geográfico, información sobre la proximidad a puntos de referencia conocidos o únicos u otra información de localización)	Fechas de los eventos (admisión, alta, tratamiento, muerte, recogida de muestras, visita)
Idioma materno	Profesión
Origen étnico	Dificultades de actividad y reducciones
Identidad aborígen	Códigos (diagnos, tratamientos y eventos adversos)
Años de escolaridad	Religión
Estado civil (Legal)	País de nacimiento
Historial criminal	Pluralidad de nacimiento

Tabla 24 Cuasi-identificadores comunes (Institute of Medicine (IOM), 2015)

En otro estudio (El Emam, 2014), los autores aportan un listado de atributos identificables a ser anonimizados por su riesgo de reidentificación (véase la Tabla 25):

Clasificación de atributos identificables	
Nombre	Fotografía
Dirección	Número de placa médica
Número de teléfono	Dirección de correo electrónico
Número de fax	Datos biométricos
Número de registro médico	Número de tarjeta sanitaria
Número de los dispositivos implantados	Número del beneficiario del seguro médico
Número de la Seguridad/seguro social	Número de registros de ensayos clínicos

Tabla 25 Posibles atributos identificables (El Emam, 2014)

En (Sweeney, 1996), se presenta un sistema, denominado *Scrub*, para localizar y reemplazar la información personal situada en los registros médicos. En el artículo se describen una serie de atributos que están compuestos, a su vez, de otros (véase la Tabla 26). Las relaciones entre atributos pueden ayudar a obtener una lista de atributos médicos compuestos aunque no es el objetivo principal del trabajo, pero podrá justificar la modificación del modelo en el análisis factorial confirmatorio.

Entidades Scrub		
Atributos compuestos	Bloque de identificación	{1, 2, 3, 4, 5, 10, 6, 7, 8, 9, 20, 11, 12, 13, 15, 16, 20}
	Etiqueta de correo	{1, 2, 3, 4, 10, 6, 7, 8, 9, 12, 13}
	Bloque de direcciones	{1, 2, 3, 4, 10}
	Nombre completo	{6, 7, 8, 9, 12}
	Localización	{2, 3, 10}
Atributos simples	1. Calle	10. País
	2. Ciudad	11. Seguridad social
	3. Estado	12. Título
	4. Código postal	13. Organización
	5. Teléfono	14. Medidas
	6. Primer apellido	15. Edad
	7. Segundo nombre	16. Fecha
	8. Apellido	17. Términos médicos
	9. Apodo	18. Número de referencia

Tabla 26 Algunas de las entidades reconocidas por Scrub se enumeran más arriba en orden relativo de prioridad (Sweeney, 1996)

Por último, en el artículo (Meystre et al., 2010) se indican los PHI que anonimizan los algoritmos más representativos utilizados en anonimización automática de documentos de texto. En una primera instancia los autores clasifican los atributos a anonimizar en siete categorías de atributos:

- Nombres de personas: nombres del paciente, miembro de la familia, representante del paciente, y proveedor de atención médica.
- Edades > 89: edad mayor de 89.
- Localizaciones geográficas: lugares tales como direcciones, nombres de calles, códigos postales y ciudades.
- Hospitales y organizaciones de la salud: centros de salud, laboratorios, y residencias de ancianos.
- Fechas: todos los elementos relacionados con las fechas menores de un año.
- Información de contacto: números de teléfono, números de contacto, números de fax, y direcciones de email.
- IDs: números de la seguridad social, número de historia clínica, número de licencia de conducir, y otros identificadores.

1.5 Conclusiones

Tras realizar un análisis exhaustivo de la normativa nacional e internacional sobre protección de datos, se han obtenido las siguientes conclusiones:

1. Existen países con leyes de protección de datos personales que no contemplan los datos relativos a la salud como datos que deban ser protegidos (por ejemplo Singapur). Sin embargo, son una minoría a nivel mundial.

2. Hay una gran mayoría de países que han reutilizado el concepto de protección de datos de la salud en sus leyes de protección de datos. Sin embargo, no han definido qué datos personales se consideran como datos personales sanitarios. Entre ellos está España con su LOPD. Por esto, se hace necesario la creación de una metodología que permita generar modelos de atributos médicos para su protección aplicables a la normativa de cada país.
3. Existen un número reducido de países que incluyen en su ley de protección de datos aquellos datos personales relativos al campo de la salud. Además, incluyen una serie de campos que deben ser protegidos dentro del campo sanitario para garantizar la protección del individuo frente a su identificación.

Además, analizando los informes anuales presentados por el *Identity Theft Resource Center* de 2014 (Identity Theft Resource Center, 2015b) y de 2015 (Identity Theft Resource Center, 2015a) en los que divide los incidentes de seguridad en cinco sectores (Medicina/Salud, Banca/Financiero/Tarjetas de crédito, Negocios, Educativo y Gobierno/Militar) se obtiene lo siguiente: (a) Se produjeron 333 incidentes de seguridad en el campo de la Medicina/Salud (42,5% del total anual) con un número de robo de registros de 8.277.991 (9,7% del total anual) y (b) En el año 2015 y a fecha de 30 de marzo de 2015, el número de incidentes de robo de información en el campo de la Medicina/Salud ha sido de 68 (32,7% del total anual) con 99.335.375 de registros robados (97,9% del total anual). Esto es debido al incidente

provocado en la aseguradora Anthem, donde se produjo un robo de 80 millones de registros. Por ello, surgió la necesidad de protección de los datos médicos como propósito del año 2015.

Por último, se recoge en la Tabla 27 el listado con los 122 datos médicos después de extraerlos de la normativa analizada y eliminar los atributos duplicados.

Listado de datos médicos protegidos a nivel internacional	
Nombre	Ingresos totales
Primer apellido	Proveedor de atención médica
Segundo apellido	Transacciones financieras
Apodo	Números de los beneficiarios del plan de salud
Persona de contacto del paciente	Número de la Seguridad Social
Fecha de nacimiento	Datos financieros-clínicos
Estado civil	Representante del paciente
Género	Número del beneficiario del seguro médico
Nacionalidad	Números de cuenta
Idioma	Número de tarjeta sanitaria
Teléfono	Datos biométricos de voz
Fax	Otros datos biométricos (iris, firma, ...)
Correo electrónico	Uso de aparatos auditivos
Calle	Uso de aparatos oftalmológicos
Ciudad	Uso de aparatos ortopédicos
Municipio	Prótesis
Distrito	Número identificativo de los dispositivos implantados
Dirección	Tipo sanguíneo
Comunidad autónoma	Salud física
Estado	Salud mental
País de nacimiento	Servicio prestado
Código postal	Tratamiento
Otros códigos de localización	Ácido desoxirribonucleico
Matrícula del vehículo	Alergias
Bastidor del vehículo	Centros de salud
Número de licencia de conducir	Consumo de sustancias tóxicas
Números certificado/licencia	Diagnosis
Otros identificadores de vehículo	Elemento indicativo de edad
Origen étnico	Enfermedades
Religión	Estado embarazo
Creencias	Estado de enfermedad
Opiniones personales	
Color	

Listado de datos médicos protegidos a nivel internacional	
Condición de minoría visible	Historial clínico
Cultura	Información psicológica
Necesidades personales	Información psiquiátrica
Opinión política	Información recolectada durante un servicio sanitario
Opiniones, comentarios o preferencias	Intervenciones quirúrgicas
Orientación sexual	Laboratorios que analizaron las pruebas médicas
Origen filosófico	Medidas corporales (talla, peso)
Origen religioso	Prognosis
Origen social	Residencias de ancianos
Pluralidad de nacimiento	Términos médicos
Historial criminal	Vacunas
Símbolos (tatuajes, piercings, escarificaciones, ...)	Incapacidades médicas
Identidad aborigen	Imágenes fotográficas identificables
Dirección IP (Dirección de red o Internet)	Estado de discapacidad
Enlaces a páginas o usuarios en Internet	Genética familiar cercano
Afiliación sindical	Genética propia
Años de escolaridad	Número de historia clínica
Historial de educación	Número de identificación
Historial de empleo	Número de placa médica
Iniciales	Número de registro médico
Posición dentro de la familia (padre, hermano mayor, etc.)	Números de serie de dispositivos identificadores
Nombre del último empleo	Códigos de diagnóstico
Organización en la que trabaja	Códigos de eventos adversos
Profesión	Códigos de tratamiento
Donación fluidos corporales	Fecha de admisión
Donación sustancia corporal	Fecha de alta
Donación de cuerpo	Fecha de deceso
Donación parte corporal	Fecha de recogida de muestras
	Fecha de tratamiento
	Fecha de visita

Tabla 27 Listado de datos médicos protegidos a nivel internacional

2. METODOLOGÍA DE INVESTIGACIÓN

METODOLOGÍA DE INVESTIGACIÓN

2.1 Introducción

En el anterior capítulo, quedaron presentadas normativas nacionales, comunitarias e internacionales vigentes hasta la fecha. Tras su análisis, se destaca la existencia de una disparidad normativa en lo que respecta a los atributos médicos, que deberían ser protegidos tanto en la variedad de su definición como en las medidas de protección que deben ser aplicados. Esto justifica la necesidad de desarrollar un modelo de atributos clínicos a anonimizar para garantizar el cumplimiento de la Ley Orgánica de Protección de Datos de Carácter Personal 15/1999 en España (España, 1999). Con la intención de dar rigor a este trabajo y a su vez obtener un modelo práctico que pueda ser utilizado por médicos e investigadores, se han combinado en él los enfoques teórico y práctico. Por ello, este modelo de atributos clínicos se compone de los siguientes elementos: (a) un enfoque teórico, que se obtiene a través de la normativa y estudios relacionados descritos en el Capítulo 1 y (b), un enfoque empírico que se extrae mediante la consulta a expertos del campo de la medicina. A continuación, se muestran los propósitos que se pretenden lograr a través de este modelo:

1. Elaborar un catálogo con los atributos médicos categorizados como resultado de la revisión interna del listado de atributos médicos obtenidos en el Capítulo 1 para obtener los que son utilizados en el ámbito sanitario de España. Para realizarlo, se utilizará el criterio de expertos en las diferentes áreas de conocimiento involucradas en el proceso de evaluación.

2. Cuantificar la importancia relativa o peso de cada atributo médico en el modelo que permita priorizar los atributos en el momento de la anonimización y apoyar la evaluación de impacto de privacidad propuesta por el Grupo de trabajo del Artículo 29 (Grupo de Trabajo sobre Protección de Datos del Artículo 29, 2014) y recogida como nueva obligación por el RGPD.

La metodología DuMoD (*Domain-oriented qQuality MOdels Development*) (Villalba et al., 2010) fue diseñada para facilitar la definición de criterios de calidad del software representativos y para validarlos. Además, también proporciona la reducción de parámetros de calidad para obtener los más importantes y aporta información sobre el peso de cada criterio y las relaciones existentes entre estos. Dichos criterios o atributos de calidad tienen en común con los atributos médicos que ambos describen una entidad: en el primer caso, un producto software; en el segundo, un paciente. Además, ambos forman una taxonomía clasificada y comparten la complejidad del proceso al que apoyan (evaluación de la calidad del software en el primer caso y anonimización de atributos médicos), lo que conlleva que sean procesos que requieran de una reducción de los atributos a los más importantes para mejorar la eficiencia de dichos procesos. Es, por ello, que esta metodología se ajusta especialmente bien al proceso requerido en este trabajo de tesis. Sin embargo, existen diferencias importantes entre ambos procesos, lo que no permite aplicar el proceso tal cual, siendo necesario su ajuste a las condiciones específicas del proceso de obtención de atributos médicos.

2.2 Metodología para la creación de modelos de datos médicos a anonimizar

La definición de la metodología de creación de un modelo de datos médicos a anonimizar, fruto de la investigación desarrollada en este trabajo de tesis, persigue:

- Facilitar la elección de atributos que necesitan ser anonimizados mediante el uso de un modelo de atributos clínicos que garantice la legislación de protección de datos para la jurisdicción que se desarrolle.
- Aportar un proceso de creación de modelos de atributos que permita desarrollar nuevos modelos de este tipo que cumplan otras normativas de protección de datos aplicables a otros países.

La metodología tiene una fundamentación teórica al basarse en el análisis previo de legislación y trabajos relacionados. Pero, además, está validado por expertos profesionales como usuarios finales de la misma, tanto profesionales sanitarios, como personal relacionado con el campo de la medicina (ingenieros vinculados a través de proyectos tecnológicos y también investigadores de informática aplicada a la salud).

La Figura 2 muestra la metodología DuMoD (Villalba, Fernández-Sanz, Cuadrado-Gallego, & Martínez, 2010; Villalba et al., 2010) y la adaptación de la misma denominada “*metodología adaptada para la creación de modelos de datos médicos a anonimizar*”. El proceso de adaptación consiste, en primer lugar, en sustituir la fase de “*Initial Analysis*” de la

metodología DuMoD, por un listado de datos que actúa como entrada de la Fase 1 de la metodología adaptada. Esta sustitución se realiza ya que en DuMoD se especifican por un lado, la definición y el dominio de aplicación del modelo; y por otro lado, se realiza una adaptación del estándar ISO 9126-1 (*International Organization for Standardization, and International Electrotechnical Commission, 2001*). A diferencia de lo que ocurre en la metodología DuMoD, en este caso los atributos médicos válidos a nivel internacional se han podido obtener del trabajo inicial de investigación del estado de la cuestión por lo que se incluyen como contribución de este trabajo de tesis, y en consecuencia es posible acortar la metodología eliminando dicha fase.

En segundo lugar, la fase "*Definition of preliminary model*" consiste en crear una definición inicial y una revisión interna del modelo inicial mediante la aplicación de la legislación sobre privacidad propia del dominio (país al que se aplica la metodología) y revisiones internas. En la adaptación de la metodología, el listado de datos médicos de la Fase 1 proporciona un listado inicial de atributos a nivel internacional. Sin embargo, es necesario extraer de este listado los atributos médicos de aplicación en el ámbito sanitario español. En esta misma fase del proceso DuMoD, el estándar ISO 9126-1 proporciona los atributos categorizados por grupos o categorías. Por tanto, se conserva este único paso renombrándose como "*Creación del catálogo de atributos médicos*" con el propósito de categorizarlos siguiendo los criterios del personal entrevistado a nivel interno.

En tercer lugar, la Fase 2 de la metodología adaptada se compone de los mismos pasos que la metodología DuMoD, salvo el paso "*Data Collection*", que especifica el acceso restringido a los cuestionarios y el análisis de los datos demográficos. La adaptación consiste en agrupar las tareas relacionadas con el diseño de los cuestionarios (incluido su alojamiento web y su acceso) y la recogida de respuestas agrupándolas bajo "Diseño del instrumento de medida y recogida de datos". Esta adaptación se realizó con el fin de simplificar el proceso. El paso "Análisis de datos" analiza los datos que fueron recogidos en el paso "Diseño del instrumento de medida y recogida de datos", incluidos los datos demográficos y se mantiene sin cambios.

En cuarto lugar, el paso "*Previous Analysis*" de la metodología DuMoD, elimina errores producidos en la recogida de datos y evalúa la posibilidad de realizar el análisis factorial mediante el análisis de aplicabilidad del modelo. En esta adaptación, se ha considerado separarlas en dos pasos dado que se utilizan técnicas estadísticas para propósitos diferentes: el paso "Análisis previo" donde se eliminan errores en la recogida de datos y se excluyen los atributos no utilizados en el ámbito sanitario de España por su baja importancia; y el paso "Análisis de aplicabilidad" donde se evalúa la matriz de correlación y el determinante de la misma con la intención de verificar la aplicación del análisis factorial. Esta adaptación se ha realizado para aportar mayor claridad al proceso.

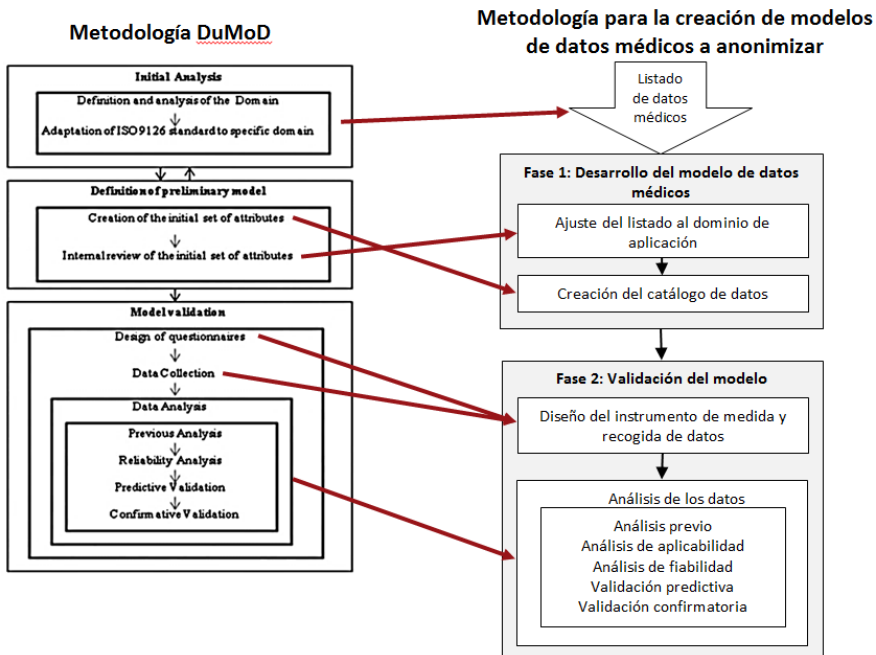


Figura 2 Metodología para la creación de modelos de datos médicos a anonimizar (adaptación de la metodología DuMoD)

Tal y como queda esquematizado en la Figura 2, la metodología para la creación del modelo de datos médicos consta de dos fases. A su vez, estas fases se componen de procesos con documentos de entrada y documentos de salida (Figura 3).

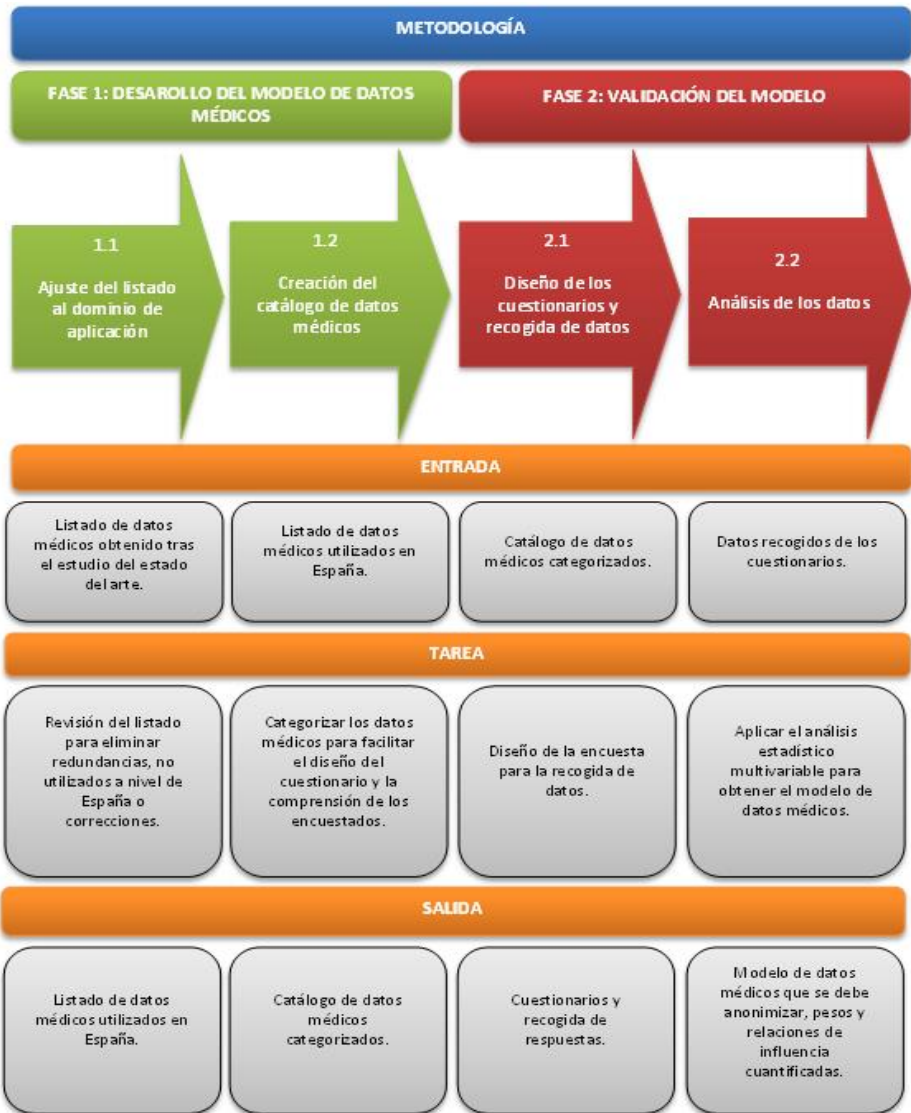


Figura 3 Resumen del proceso llevado a cabo para extraer el modelo de datos médicos

2.2.1 Fase 1. Desarrollo del modelo de datos médicos

En esta fase se define el modelo teórico de datos médicos en el ámbito de España a partir de un listado de datos médicos protegidos a nivel internacional. A continuación, se explican los pasos de forma detallada.

2.2.1.1 Ajuste del listado al dominio de aplicación

El listado de datos médicos obtenidos debe ser reducido y optimizado para conseguir un modelo práctico para el cumplimiento de la normativa europea y nacional aplicable en España. Para ello, se realizan las siguientes acciones:

- Determinar y unificar los datos médicos duplicados.
- Identificar y descartar los datos médicos que no aplican al dominio de estudio; y
- Reescribir los datos médicos de carácter ambiguo.

La revisión se lleva a cabo por personal médico o vinculado al entorno sanitario, con conocimiento sobre los datos a tratar y amplia experiencia tanto en el área de la Medicina como en las TIC. Esta revisión se considera imprescindible para obtener un catálogo de datos médicos que apliquen al dominio definido. No todos los datos médicos recogidos en un listado global, a nivel internacional, poseen la misma importancia para los profesionales (profesionales relacionados en áreas médicas) en España. Para tener otro criterio que aporte información de aplicación al ámbito español se utilizará el Real Decreto 1093/2010, de 3 de septiembre, por el que se aprueba el conjunto mínimo de datos de los informes clínicos en el Sistema Nacional de Salud (España, 2010).

2.2.1.2 Creación del catálogo de datos médicos

Este paso consiste en categorizar los datos médicos revisados en el paso anterior y crear el catálogo que facilite el diseño de la encuesta. El catálogo quedará presentado en tablas que agrupen los datos médicos en categorías según su significado, junto con la fuente de las que fueron obtenidas.

2.2.2 Fase 2. Validación del modelo

La fase de validación se compone de dos pasos principales: el diseño del instrumento de medida y la recogida de datos, y el análisis de los mismos. Para ello, se debe disponer del catálogo de atributos ajustado a la jurisdicción del caso de estudio al inicio de esta fase. Posteriormente y siguiendo los pasos que se describen a continuación, se obtiene un modelo de datos médicos validado.

2.2.2.1 Diseño del instrumento de medida y recogida de datos

Esta fase tiene como finalidad extraer la opinión de expertos en las distintas áreas de conocimiento involucradas en la utilización de datos médicos (investigadores, personal médico y profesionales relacionados con el tratamiento de datos clínicos) para poder reducir y validar el conjunto inicial de datos médicos. Esta tarea se llevará a cabo a través de metodología cuantitativa para poder explorar la opinión de un número grande de sujetos. En concreto, se utilizará el método de la encuesta siguiendo el ciclo de vida definido en (Ameller, Galster, Avgeriou, & Franch, 2015) y representado en la Figura 4.

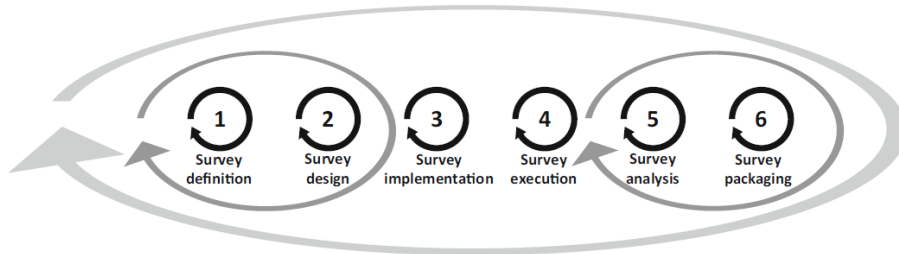


Figura 4: Etapas del método de encuesta (Ameller et al., 2015)

A continuación, se describirán en más detalle las etapas seguidas para llevar a cabo el método de encuesta.

En primer lugar, se define la hipótesis general del cuestionario la cual, será verificada tras el tratamiento estadístico de los datos recogidos. Para ayudar a realizar esta tarea, se plantearon las siguientes preguntas de investigación:

- P1: ¿Qué atributos médicos utiliza el personal médico e investigador en el área de medicina y medicina informática en su trabajo a diario?
- P2: ¿Podría haber otros atributos médicos no obtenidos mediante la investigación teórica previa?
- P3: ¿Sería útil para el personal médico e investigador disponer de un modelo de atributos médicos que asegure la protección de los datos de los pacientes?

La pregunta P1 pretende averiguar qué atributos médicos son utilizados, de una manera habitual, en medicina con el fin de descartar aquellos que no lo son evitando así riesgos innecesarios en la privacidad. La pregunta

P2 permite extraer los atributos médicos que el personal encuestado echa en falta en el cuestionario. De esta forma, se comprueba si los resultados obtenidos en la investigación previa de las leyes relacionadas con privacidad a nivel internacional y los trabajos relacionados publicados contienen todos los atributos necesarios para los usuarios finales de los datos. La pregunta P3 permite extraer la opinión del encuestado sobre la utilidad de disponer de un modelo de atributos que ayude a garantizar la protección de datos de los pacientes.

Una vez definida la hipótesis del cuestionario, el siguiente paso sería el diseño de la encuesta. La encuesta a realizar será de tipo descriptivo y, en concreto, el tipo de constructo a utilizar será un cuestionario descriptivo. Estas encuestas permiten realizar afirmaciones sobre la población y medir lo que ocurrió sin importar el por qué (Ameller et al., 2015). En este caso, se estudiarán aquellos atributos clínicos que son los de mayor uso para poder crear un modelo de los mismos garantizando el cumplimiento de las leyes de protección de datos propias del país en el que se esté aplicando la metodología.

Finalmente, la hipótesis general del cuestionario planteada sería:

H₀: Es posible obtener un conjunto de atributos clínicos consensuado por expertos que contribuya a descubrir relaciones entre atributos y defina la importancia relativa de cada uno de ellos válido en un dominio o país previamente definido

En primer lugar, se debe elegir la población a la que va dirigida, y que en este caso son aquellas personas que desarrollan su trabajo en el campo

de la medicina y que tiene relación directa con la información médica de los pacientes. En esta población elegible, pueden incluirse personal sanitario, investigadores o incluso personal técnico relacionado con los sistemas de información utilizados en medicina.

El segundo paso es elegir el tipo de muestreo. En este caso, se ha elegido un muestreo no probabilístico, más concretamente, el muestreo intencional o a criterio porque solo debe participar en la muestra expertos del sector sanitario, ya sean médicos, investigadores o personal con participación en el uso de datos médicos y, por tanto, la población no es elegida al azar.

Algunas de las recomendaciones que se han seguido para definir el cuestionario son las siguientes (Abascal & Grande Esteban, 2005; Grasso, 2006; Manolov et al., 2015):

- Desarrollar una sección inicial en la que se detalla el estudio y se solicitan los datos demográficos de control para que antes de comenzar el análisis de los datos, se conozca la población muestral y descartar datos que no cumplan el perfil mínimo determinado.
- Utilizar preguntas de tipo Likert porque aportan sencillez en el proceso de diseño del cuestionario y a su vez se consigue ayudar al encuestado a graduar su opinión ante preguntas complejas.
- Se usa una pregunta final abierta para poder recoger nuevos atributos médicos que se consideren importantes en el estudio o comentarios que consideren necesarios.

- Los atributos clínicos se categorizan en base a la clasificación obtenida en la fase previa del proceso con el fin de poder realizar las preguntas del cuestionario por categorías que agrupen pocos elementos de manera que se evite el riesgo de no ser contestadas.
- Las preguntas se deben definir de una manera clara y concisa para poder llegar a una mayor población muestral aprovechando que la población ya está familiarizada con los atributos médicos.
- Previa a la publicación del cuestionario, se debe realizar un piloto del cuestionario para evitar problemas en el cuestionario final.
- Para facilitar la recogida de datos, Internet es un medio a tener en cuenta para realizar la distribución del cuestionario ya que permite la accesibilidad a toda la población muestral y la distribución del mismo mediante correo electrónico.
- Para garantizar que el encuestado está familiarizado con el ámbito del cuestionario, la divulgación se puede realizar en asociaciones especializadas y a través de contactos de profesionales.

2.2.2.2 *Análisis de datos*

El paso inicial en el análisis de datos es el del análisis y descripción de la muestra. Este análisis preliminar consiste en calcular el error muestral y analizar los datos demográficos con el propósito de garantizar que el grupo de usuarios encuestados son válidos y los datos recogidos son fiables antes de llevar a cabo el análisis de los datos. Este paso se hace

necesario debido a que, aunque la encuesta se distribuye a personal relacionado con datos sanitarios, ésta puede ser publicada sin restricciones de acceso y podrían encontrarse datos no relevantes o personal encuestado para el que no fue diseñada. Además, estos datos demográficos no se recaban para realizar un modelo adaptado a un sector o una especialidad médica, ya que no es el objetivo de este trabajo de investigación.

Tras el análisis previo, los datos ya están preparados para ser analizados a través de técnicas de tratamiento estadístico que permitan validar los resultados. Pero antes se debe llevar a cabo un análisis de aplicabilidad para verificar si el catálogo de datos médicos propuesto es de aplicación en el ámbito sanitario español. Para saber si es aplicable se consulta a los expertos añadiendo una pregunta al cuestionario sobre la utilidad de disponer de un modelo de este tipo.

Una vez verificada la viabilidad de los datos recogidos, se lleva a cabo la validación de los mismos que permitirá a su vez construir el modelo. La validez indica el grado en el cual el modelo (en este caso) evalúa el concepto que el investigador pretende medir (Ferrer Guerra, Pascual Miguel, Guerrero Barrón, Ríos Manríquez, & Pérez Carmona, 2011). Veamos a continuación los distintos tipos de validación que se han utilizado en esta tesis y cómo se han llevado a cabo en este trabajo de tesis:

- Validación de contenido. Consiste en evaluar, de manera subjetiva pero sistemática, si las variables comprenden convenientemente la base teórica del modelo que se está midiendo. Para llevar a cabo

esta validación, se lleva a cabo el análisis del estado de la cuestión correspondiente.

- Validación de criterio. Consiste en contrastar la medida o procedimiento por medio de la comparación con otro ya existente y que ha demostrado ser válido. Esta validación no se puede aplicar siempre, ya que para poder ser aplicada debe existir otro procedimiento o medida validado con las mismas especificaciones que el que se pretende validar. Esta validación no se puede aplicar en este caso por la falta de otra metodología de creación de modelos de atributos médicos a anonimizar con el que poder realizar la comparativa.
- Validación de constructo. Mide la relación entre mediciones a partir de las hipótesis teóricas planteadas y que contienen a las variables medidas o constructos incluidos en la teórica planteada. Se divide en tres tipos:
 - Validación convergente: grado en el que dos mediciones sobre el mismo constructo correlacionan entre sí. Para realizar esta validación se estudió el coeficiente α de Cronbach. El α de Cronbach tiene el inconveniente de su alta dependencia lineal con el número de variables usadas para medir un concepto. Por este motivo, se hace necesario la utilización de correlaciones ítem-total (Peter, 1981) que permiten medir la correlación de la variable con el valor de la suma total de todas las variables de cada factor.
 - Validación discriminante: mide el grado en que dos conceptos no se correlacionan utilizando el mismo método de medición. La validación discriminante permite predecir el modelo

subyacente. Para ello, se utilizan técnicas de análisis multivariante (Gerbing & Hamilton, 1996), en concreto análisis factorial exploratorio.

- Validación nomológica: grado de correlación tal y como se ha pronosticado de forma teórica con las medidas de constructos diferentes, pero relacionadas mediante factores. Esta validación permite confirmar el modelo que previamente se ha predicho, utilizando análisis factorial confirmatorio que permite, además, contrastar la hipótesis y demostrar si el ajuste de los datos recogidos concuerda con la estructura esperada (Hair, Tatham, Anderson, & Black, 1998). El análisis de las ecuaciones estructurales y el análisis factorial confirmatorio constituyen el análisis empírico y tienen como fin reducir el error de medida del modelo planteado.

Para mayor claridad, la Tabla 28 muestra un resumen de los tipos de validación realizados, la técnica utilizada y la referencia a la metodología propuesta en este trabajo.

Contexto de validación	Técnica utilizada	Referencia en la metodología propuesta
Validación de contenido	Revisión interna	Ajuste del listado al dominio de aplicación
Validación convergente	Coefficiente α de Cronbach y las correlaciones ítem-total	Análisis de fiabilidad
Validación discriminante	Análisis factorial exploratorio	Validación predictiva
Validación nomológica	Análisis factorial confirmatorio	Validación confirmatoria

Tabla 28 Relación de contextos de validación y técnicas aplicadas

En los siguientes apartados se describe en detalle cada una de las actividades de las que consta esta fase de la metodología (ver Figura 2).

Análisis previo

En el análisis previo de los datos se realiza un análisis preliminar de los datos recogidos utilizando un análisis descriptivo. Este análisis es realizado para cada atributo del modelo de atributos siguiendo los siguientes pasos:

1. En primer lugar y utilizando la herramienta SPSS, se intentan identificar datos atípicos y errores producidos en la recogida de datos mediante los diagramas de caja.
2. Posteriormente, se calculan los siguientes estadísticos descriptivos para cada variable o atributo: (a) media, (b) mediana, (c) moda y (d) desviación típica.
3. Por último, se analizan los histogramas y se verifica la normalidad o no de los datos para evaluar su distribución.

Con el fin de eliminar los atributos médicos con poca importancia del modelo final, se ha establecido que si un atributo médico posee valores de media, mediana y moda menor que 2; y si además la asimetría tiene un carácter negativo resulta que la utilización de ese atributo médico por parte de los encuestados es “nunca” o “casi nunca”, tal y como se ha definido en la escala Likert del cuestionario. Se ha elegido este valor como criterio conservador para evitar eliminar en un estado temprano valores utilizados por parte de los expertos.

Análisis de aplicabilidad

Con el fin de verificar la aplicabilidad del modelo en el ámbito sanitario de España se realizarán dos etapas de consultas sobre el conjunto de atributos. En la fase “Ajuste del listado al dominio de aplicación” de la metodología adaptada de la Figura 2, un grupo reducido de profesionales relacionados con atributos médicos evalúan mediante técnicas cualitativas los atributos médicos que utilizan en el ámbito sanitario español. En la fase “Diseño del instrumento de medida y recogida de datos” de la metodología de la Figura 2 se consulta a profesionales relacionados directamente con el ámbito sanitario, para determinar si un modelo así les resultaría de utilidad en su trabajo diario. De esta forma es posible determinar si los atributos obtenidos de forma teórica son adecuados en el entorno médico de España.

Análisis de fiabilidad

El análisis de fiabilidad minimiza el error de medida para que el resultado sea un modelo fiable. Dicho análisis consiste en tomar los datos médicos agrupados en categorías en las que compartan características, y analizar el coeficiente α de Cronbach y las correlaciones ítem-total para cada variable o dato médico. El valor de aceptación de cada variable para el α de Cronbach debe ser igual o superior a 0,7, siendo considerados los valores por encima de 0,8 como buena fiabilidad, y 0,9 como excelente (Tabachnick & Fidell, 2006). Además, el valor mínimo para la correlación ítem-total debe ser superior a 0,3 (Kline, 2000; Rocha, Pérez, Rodríguez-Sanz, Borrell, & Obiols, 2011). En el caso de que alguna variable no

supere los valores mínimos de ítem-total, ésta será eliminada para no reducir la consistencia de la escala (α de Cronbach) o tiene una relación escasa con el concepto que se quiere medir (ítem-total); mejorando así, la fiabilidad de la escala medida.

Validación predictiva

En esta fase se realiza la validación discriminante, y para ello, se utiliza el análisis factorial exploratorio. De esta manera, se obtiene la estructura compuesta por las correlaciones entre variables (Tabla 28). Para saber si es posible realizar el análisis factorial exploratorio se tomarán los siguientes indicadores de la matriz de correlaciones:

1. Test de esfericidad de Bartlett: Verifica que si se cumple la hipótesis nula de que la matriz de correlaciones tiene determinante 1 por tratarse de la matriz identidad. En este caso, no habría correlación entre variables (Bisquerra Alzina, 1989).
2. Índice Kaiser-Meyer-Olkin (en adelante KMO): Calcula la relación entre los coeficientes de correlación parcial y los coeficientes de correlación de Pearson (Kaiser, 1974):
 - KMO cercano a 0 indica correlación baja o nula y se desaconseja el análisis factorial.
 - KMO mayor que 0,5 debería ser aceptado.
 - KMO mayor o igual de 0,6 es mediocre.
 - KMO mayor que 0,7 se considera un buen valor.
 - KMO mayor o igual que 0,8 es meritorio.
 - KMO mayor o igual que 0,9 es muy bueno.

3. Determinante de la matriz de correlación: Un determinante próximo a 0 indica un alto grado de correlación entre variables (Martín Martín & del Rosario de Paz Santana, Yanira, 2007).

Una vez realizadas las tres comprobaciones anteriores para verificar que se cumplen los mínimos establecidos, se realiza el análisis factorial exploratorio utilizando como método de extracción el método de componentes principales porque ayuda a minimizar el número de los factores originales (Hair et al., 1998). Los criterios que se utilizarán para decidir el número de factores son: (a) los autovalores o “eigenvalues” y (b) el porcentaje de la varianza explicada. Los autovalores permiten incluir el número de factores necesarios que permitan representar, de una forma adecuada, cada variable original (Tabachnick & Fidell, 2006). Para cada categoría se obtiene la matriz de factores compuesta de cargas factoriales sobre cada factor. Las cargas factoriales están ordenadas según su importancia (varianza) sobre dicho factor. Las cargas con valores cercanos a $\pm 0,32$ se encuentran en el nivel mínimo, y las cargas mayores a $\pm 0,50$ son significativas (Tabachnick & Fidell, 2006). De esta forma, se logra suprimir la redundancia manteniendo la información original. Por su parte, el porcentaje de la varianza explicada es un porcentaje acumulado de la varianza total extraída y asegura que los factores derivados explican, al menos, un valor especificado de la varianza. El porcentaje de la varianza explicada se considera satisfactorio cuando es mayor o igual al 60% de la varianza total (Hair et al., 1998).

Por último, es necesario obtener una solución factorial definitiva y dar una interpretación de los factores resultantes. Para facilitar la

interpretación de cada componente, se utiliza el método de rotación. Este método, simplifica el patrón extraído y facilita la interpretación conceptual de los factores extraídos. El método de rotación elegido es el método ortogonal Varimax ya que es el más recomendado en situaciones en las que hay que realizar una reducción del número de variables originales a un conjunto de variables no correlacionadas para su posterior uso en el análisis confirmatorio (Hair et al., 1998). La eliminación de altas correlaciones es importante en el desarrollo de modelos ya que permite evitar redundancias simplificando así los modelos sin perder información significativa y, por ende, aumentando la eficiencia de los procesos. Entre los métodos ortogonales se ha seleccionado Varimax por ser el más robusto (Hair et al., 1998).

Validación confirmatoria

En esta fase se utiliza la técnica de análisis factorial confirmatorio basado en la utilización del modelo de medida de ecuaciones estructurales SEM (*Structural Equation Models*) (Ruiz, Pardo, & San Martín, 2010). Este modelo posee las siguientes características, tal y como se presentan en (Ruiz et al., 2010):

- Permiten estimar el peso y las relaciones entre múltiples variables.
- Son menos restrictivos que los modelos de regresión por permitir incluir errores de medida tanto en las variables dependientes como en las variables independientes.

- Permite plantear el tipo y dirección de las relaciones que se espera encontrar entre las distintas variables contenidas en él, para posteriormente estimar los parámetros que vienen especificados por las relaciones propuestas a nivel teórico.
- Permite controlar el error de medida y realizar un test estadístico de calidad del ajuste para la solución confirmatoria propuesta.

El modelo de ecuaciones estructurales permite representar conceptos no observados y tener en cuenta el error de medida en el proceso de estimación. Además, es considerado como la única técnica que permite analizar, de forma simultánea, una serie de relaciones de dependencias múltiples y cruzadas que constituyen un modelo a gran escala (Hair et al., 1998).

El proceso necesario para llevar a cabo el ajuste de un modelo de análisis factorial confirmatorio se muestra en la Figura 5.

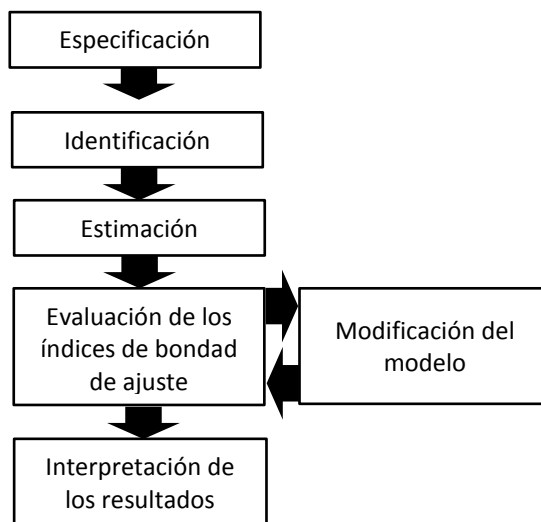


Figura 5 Etapas del análisis confirmatorio

Debido a que la Figura 5 es de elaboración propia para facilitar la comprensión, se explica a continuación cada fase del proceso junto con los valores de corte de las medidas utilizadas (Lévy Mangin & Varela Mallou, 2006):

Especificación: Se definen las ecuaciones estructurales a partir de los resultados obtenidos en el análisis factorial exploratorio para representar el modelo teórico.

Identificación: Se debe asegurar que pueden ser estimados los parámetros del modelo mediante en cálculo de grados de libertad. Un modelo puede estar en tres estados:

- **Identificado:** Se caracteriza por tener cero grados de libertad ($gl=0$). Esto indica un ajuste perfecto del modelo, aunque no se puede generalizar.
- **Sobreidentificado:** El número de grados de libertad es mayor que cero ($gl > 0$). Esto indica que la matriz de datos tiene más información que el número de variables a estimar.
- **Infraestimado:** El número de grados de libertad es menor que cero ($gl < 0$). Esto indica que la estimación intenta explicar más parámetros de los que se encuentran almacenados en la matriz de datos.

Estimación: Mediante programas software (AMOS o SAS) se estiman las variables estandarizadas y no estandarizadas para determinar el error de medición de las variables desconocidas. El método que se utiliza en

nuestro caso es el de Máxima Verosimilitud Robusto ya que es eficiente y no sesgado cuando el modelo cumple con los supuestos de normalidad multivariada (Byrne, 1994).

Evaluación de los índices de bondad de ajuste: A partir de diferentes estudios (Bentler, 1990; Bollen, 1989; Hair et al., 1998; Hu & Bentler, 1999; Schermelleh-Engel, Moosbrugger, & Müller, 2003) en esta fase se analiza si el modelo propuesto se ajusta a los índices de bondad de ajuste. En caso afirmativo, se concluye que el modelo se aproxima a la situación real; en caso contrario, se deberá modificar el modelo para mejorar los índices de ajuste. Estas medidas, tal y como indica el trabajo (Lévy Mangin & Varela Mallou, 2006), son las siguientes:

- Ajuste absoluto: indican el grado de exactitud en que el modelo global pronostica la matriz de covarianzas:
 - Índice de chi-cuadrado (χ^2).
 - Error de aproximación cuadrático medio o RMSEA (*Root Mean Square Error Of Approximation*): se considera un valor bueno de ajuste cuando es menor que 0,05, si el valor está en torno a 0,08 y menor que 0,1 se considera un error razonable de aproximación a la población y valores superiores a 0,1 indican mala aproximación.
 - Índice de ajuste normado GFI (*Goodness of Fit Index*) o índice de Bentler Bonett: Sus valores se encuentran entre 0 y 1. Los valores superiores a 0,9 se consideran aceptados.
 - SRMR (*Standardized Root Mean Square Residual*): se define como la diferencia estandarizada entre la correlación observada y la correlación pronosticada. Es una medida

positivamente sesgada y que el sesgo es mayor para N pequeño y para estudios de baja df . Un valor inferior a 0,08 se considera generalmente un buen ajuste (Hu & Bentler, 1999).

- Ajuste incremental: miden el modelo en comparación al modelo nulo. Se utilizarán los siguientes índices:
 - Índice de ajuste normado NFI (*Normed Fit Index*) o índice de Bentler Bonett: Sus valores se encuentran entre 0 y 1. Los valores superiores a 0,9 se consideran aceptados.
 - Índice de ajuste no normado NNFI (*Non-Normed Fit Index*) o TLI (*Tucker-Lewis index*): Sus valores se encuentran entre 0 y 1. Los valores superiores a 0,9 se consideran aceptados.
 - Índice de ajuste comparativo CFI (*Comparative Fit Index*): Sus valores se encuentran entre 0 y 1. Los valores superiores a 0,9 se consideran aceptados.
- Ajuste de parsimonia: evalúa la parsimonia del modelo propuesto mediante la relación de la bondad de ajuste del modelo con el número de coeficientes estimados necesarios para conseguir ese nivel de ajuste. Los índices utilizados en este caso son:
 - χ^2 (Chi-cuadrado) normalizada o NCS (ratio entre el valor de χ^2 y los grados de libertad) cuyo valor recomendado es entre 1 y 2 para un buen ajuste. Algunos autores indican que los valores máximos recomendables son de 3, y hasta 5 para determinar un ajuste aceptable del modelo (Bollen, 1989).

Modificación del modelo: Esta tarea consta de modificar el modelo propuesto en el caso de que los índices de bondad de ajuste no sean satisfactorios. Por ello, es una tarea opcional que puede no ser necesaria aunque se debe contemplar. Las pruebas y herramientas estadísticas que permiten tomar decisiones sobre el modelo son las siguientes:

- Correlaciones brutas entre variables. Con el propósito de identificar variables indicadoras demasiado correlacionadas entre sí, dado que el modelo asume que hay ausencia de multicolinealidad.
- Prueba de Wald para las covarianzas entre factores latentes. Aunque algunas de estas covarianzas pueden ser no significativas y por lo tanto susceptibles de ser eliminadas del modelo, en análisis confirmatorios es habitual no eliminar estos términos y permitir que todos los factores covaríen.
- Prueba del Multiplicador de Lagrange (ML). Este estadístico mide de forma aproximada el grado de mejoría en el ajuste si se contemplara algún término de covarianza o parámetro inicialmente no considerado. Existen diferentes grupos de parámetros que son evaluados:
 - Entre variables indicadoras. En análisis confirmatorios es habitual asumir que las variables observadas no interactúan más que como producto de las relaciones con los factores latentes. En este sentido, no se considerarán este tipo de relaciones en los modelos propuestos.
 - Para relaciones bidireccionales entre variables indicadoras y factores. El modelo asumido sólo contempla asociaciones entre factores y variables indicadoras, sin bidireccionalidad.

- Para nuevas relaciones entre factores y variables indicadoras. Este análisis sí ofrece la oportunidad de medir relaciones no contempladas inicialmente.
- Para relaciones entre términos de error. Estas relaciones tendrían sentido en el caso de series temporales, pero en el presente modelo no son de aplicación.

Interpretación de los resultados: Como último paso, se debe interpretar y obtener las conclusiones a partir de los resultados alcanzados.

3. EVALUACIÓN DEL PROCESO Y VALIDACIÓN

EVALUACIÓN DEL PROCESO Y VALIDACIÓN

3.1 Introducción

Este capítulo describe el procedimiento llevado a cabo para validar el proceso propuesto a través de su aplicación a un caso de estudio.

La sección 3.2 describe la aplicación de la metodología descrita en el Capítulo 2 al caso de estudio. Esta escenificación permite crear el modelo de datos médicos útil para los profesionales relacionados con los procesos de anonimización de datos médicos en España. Además, se muestra el grado de importancia de los encuestados en disponer de un modelo aplicable.

3.2 Aplicación de la metodología al caso de estudio

A continuación, se describirá la aplicación de las fases definidas en el Capítulo 2 al dominio de atributos clínicos. Así queda como fin, extraer un modelo de atributos clínicos a anonimizar y que cumpla con la LOPD en España.

El propósito del modelo de atributos obtenido, es su utilización en un proceso de anonimización realizado para el cumplimiento de la LOPD en España y además, al no haber otro modelo con la misma finalidad, servir como base para futuras investigaciones o en el desarrollo legislativo en el ámbito para el que fue desarrollado el presente modelo.

3.2.1 Fase 1. Desarrollo del modelo de atributos

A partir del listado de atributos médicos protegidos a nivel internacional y que son utilizados como entrada para esta fase, se obtiene el catálogo de atributos médicos a ser protegidos en España siguiendo los pasos descritos a continuación.

3.2.1.1 Ajuste del listado al dominio de aplicación

El primer paso del proceso consiste en personalizar el conjunto de datos de entrada al dominio de aplicación seleccionado, en este caso, España. Esto permitirá obtener la taxonomía de atributos en el caso particular de España, así como, si el análisis estadístico así lo confirmara, el modelo de atributos final. Para ello, lo primero es eliminar aquellos que no se utilizan en España. Por ejemplo, es el caso del dato médico “Estado” que, al no ser de aplicación en España, se elimina del listado antes de seguir con el análisis posterior. Para hacer esta primera revisión interna, en el caso particular de España, se utilizó el Real Decreto 1093/2010, de 3 de septiembre, por el que se aprobó el conjunto mínimo de datos de los informes clínicos en el Sistema Nacional de Salud (España, 2010).

El grupo de personas para realizar la primera revisión interna es el siguiente:

- Dos investigadores expertos en el área de informática médica.
- Una médica, profesora e investigadora de la Universidad Europea.
- Tres médicos de las especialidades de: urología, medicina forense y psicología clínica, respectivamente.

- Un técnico informático que desempeña su trabajo en el hospital Central de la Defensa Gómez Ulla.

Tras la revisión de estos resultados, se eliminaron los 33 datos médicos mostrados en la Tabla 29 por la falta de utilización en el ámbito sanitario español. Estos 33 atributos clínicos representan un 25% de los atributos clínicos totales existentes al inicio de esta fase, quedando un total de 88 atributos médicos.

Datos médicos eliminados en la revisión interna	
Apodo	Origen religioso
Idioma	Historial criminal
Otros códigos de localización	Símbolos (tatuajes, piercings, escarificaciones, ...)
Matrícula del vehículo	Identidad aborígen
Bastidor del vehículo	Dirección IP (Dirección de red o Internet)
Número de licencia de conducir	Enlaces a páginas de usuarios en Internet
Número de certificado/licencia	Afiliación sindical
Otros identificadores de vehículo	Iniciales
Religión	Transacciones financieras
Creencias	Números de los beneficiarios del plan de salud
Opiniones personales	Datos financieros clínicos
Color	Representante del paciente
Condición de minoría visible	Número del beneficiario del seguro médico
Cultura	Números de cuenta
Opinión política	Datos biométricos de voz
Opiniones, comentarios o preferencias	Otros datos biométricos (iris, firma, ...)
Origen filosófico	

Tabla 29 Datos médicos eliminados en el proceso de revisión interna

Además de esta eliminación y tras realizar la revisión de los atributos que habían resultado, se unificaron los atributos “Edad” y “Elemento indicativo de edad” como un único dato renombrado a “Edad o elemento indicativo de edad”.

En el ANEXO 7 se muestra el resultado final de la revisión interna mediante entrevistas.

3.2.1.2 Creación del catálogo de datos médicos

Dado el alto número de atributos, con el fin de facilitar su manejo y comprensión, los datos médicos se clasifican según su contenido semántico:

- Datos personales de carácter general.
- Datos personales de contacto.
- Datos personales asociados a geolocalización.
- Datos personales asociados a donaciones.
- Datos personales relacionados con la profesión.
- Datos personales relacionados con aseguradoras.
- Datos personales relacionados con creencias, cultura, raza, etnia u opiniones.
- Datos personales relacionados con dispositivos implantados.
- Datos médicos relacionados con datos personales.

La Tabla 30 muestra los atributos médicos dependiendo de su categorización:

Catálogo de datos médicos del ámbito sanitario español	
Datos personales de carácter general	
Nombre (Australia, 1988; Canadá, 1985; El Emam et al., 2006; El Emam, 2014; Meystre et al., 2010)	Estado civil (Canadá, 1985; Institute of Medicine (IOM), 2015; Sudáfrica, 2002)
Primer apellido (Australia, 1988;	Género (El Emam et al., 2006; Sudáfrica,

Catálogo de datos médicos del ámbito sanitario español	
Sweeney, 1996)	2002)
Segundo apellido (Australia, 1988; Sweeney, 1996)	Nacionalidad (Canadá, 1985; Sudáfrica, 2002)
Persona de contacto del paciente (Sweeney, 1996)	Edad o elemento indicativo de edad (Canadá, 1985; Estados Unidos, 1996; Institute of Medicine (IOM), 2015; Meystre et al., 2010; Sudáfrica, 2002; Sweeney, 1996)
Fecha de nacimiento (Australia, 1988; El Emam et al., 2006; Estados Unidos, 1996; Institute of Medicine (IOM), 2015; Sudáfrica, 2002)	Sexo (Australia, 1988; Institute of Medicine (IOM), 2015; Sudáfrica, 2002)
Datos personales de contacto	
Teléfono (El Emam et al., 2006; El Emam, 2014; Estados Unidos, 1996; Meystre et al., 2010; Sweeney, 1996)	Fax (El Emam, 2014; Estados Unidos, 1996; Meystre et al., 2010)
Correo electrónico (El Emam et al., 2006; El Emam, 2014; Estados Unidos, 1996; Meystre et al., 2010)	
Datos personales asociados a geolocalización	
Calle (Meystre et al., 2010; Sweeney, 1996)	Comunidad autónoma (Estados Unidos, 1996)
Ciudad (Estados Unidos, 1996; Meystre et al., 2010; Sweeney, 1996)	Dirección (Australia, 1988; Canadá, 1985; El Emam et al., 2006; El Emam, 2014; Estados Unidos, 1996; Meystre et al., 2010; Sudáfrica, 2002)
Municipio (Estados Unidos, 1996)	País de nacimiento (Institute of Medicine (IOM), 2015; Sweeney, 1996)
Distrito (Estados Unidos, 1996)	Código postal (Estados Unidos, 1996; Institute of Medicine (IOM), 2015; Meystre et al., 2010; Sweeney, 1996)
Datos personales asociados a donaciones	
Donación de sustancia corporal (Australia, 1988; Canadá, 2000)	Donación de parte corporal (Australia, 1988; Canadá, 2000)
Datos personales relacionados con la profesión	
Años de escolaridad (Institute of Medicine (IOM), 2015)	Nombre del último empleo (Australia, 1988)
Historial de educación (Canadá, 1985)	Organización en la que trabaja (Sweeney, 1996)
Historial de empleo (Canadá, 1985)	Profesión (Institute of Medicine (IOM), 2015)

Catálogo de datos médicos del ámbito sanitario español	
Posición dentro de la familia (padre, hermano mayor, etc.) (Meystre et al., 2010)	
Datos personales relacionados con aseguradoras	
Ingresos totales (Institute of Medicine (IOM), 2015)	Número de la Seguridad Social (El Emam et al., 2006; El Emam, 2014; Estados Unidos, 1996; Meystre et al., 2010; Sweeney, 1996)
Proveedor de atención médica (Meystre et al., 2010)	Número de tarjeta sanitaria (El Emam et al., 2006; El Emam, 2014)
Datos personales relacionados con creencias, cultura, raza, etnia u opiniones	
Origen étnico (Canadá, 1985; España, 1999; Institute of Medicine (IOM), 2015; Marruecos, 2009b; Sudáfrica, 2002)	Origen social (Sudáfrica, 2002)
Necesidades personales (Australia, 1988)	Pluralidad de nacimiento (Institute of Medicine (IOM), 2015)
Orientación sexual (España, 1999; España, 2007; Sudáfrica, 2002; Unión Europea, 1995)	
Datos personales relacionados con dispositivos implantados	
Uso de aparatos auditivos (Nuevo León, 2008)	Prótesis (Nuevo León, 2008)
Uso de aparatos oftalmológicos (Nuevo León, 2008)	Número identificativo de los dispositivos implantados (El Emam, 2014)
Uso de aparatos ortopédicos (Nuevo León, 2008)	
Datos médicos relacionados con datos personales	
Tipo sanguíneo (Canadá, 1985; Sudáfrica, 2002; Trinidad y Tobago, 2011)	Salud física (otra información física no recogida en anteriores campos)(Canadá, 2000; España, 2002; España, 2007; Mauricio, 2004; Sudáfrica, 2002)
Salud mental (otra información mental no recogida en anteriores campos) (Canadá, 2000; España, 2007; Mauricio, 2004; Sudáfrica, 2002)	Servicio prestado (Australia, 1988)
Ácido desoxirribonucleico (Trinidad y Tobago, 2011)	Tratamiento (España, 2002; Lituania, 1996)
Alergias (Nuevo León, 2008)	Centros de salud (Meystre et al., 2010)
Diagnos (España, 2002; Lituania, 1996)	Consumo de sustancias tóxicas (Nuevo León, 2008)
Términos médicos que identifican	Enfermedades (Nuevo León, 2008)

Catálogo de datos médicos del ámbito sanitario español	
patologías no habituales (Sweeney, 1996)	
Estado del embarazo (Sudáfrica, 2002)	Estado de enfermedad (Finlandia, 1999)
Historial clínico (Canadá, 1985; Nuevo León, 2008; Sudáfrica, 2002)	Información psicológica (Nuevo León, 2008)
Medidas corporales (talla, peso) (Sweeney, 1996)	Prognosis (Lituania, 1996)
Estado de discapacidad (Australia, 1988; España, 2007; Finlandia, 1999; Sudáfrica, 2002)	Residencias de ancianos (Meystre et al., 2010)
Imágenes fotográficas identificables (El Emam, 2014; Estados Unidos, 1996)	Vacunas (Nuevo León, 2008)
Incapacidades médicas (Nuevo León, 2008)	Genética del paciente (Australia, 1988; España, 2007; Marruecos, 2009b)
Genética familiar cercano (Australia, 1988; España, 2007; Marruecos, 2009b)	Número de identificación (Finlandia, 1999)
Número de historia clínica (Meystre et al., 2010)	Número de placa médica (El Emam, 2014)
Códigos de diagnóstico (Institute of Medicine (IOM), 2015)	Número de registro médico (El Emam, 2014)
Códigos de eventos adversos (Institute of Medicine (IOM), 2015)	Fecha de recogida de muestras (Institute of Medicine (IOM), 2015)
Fecha de admisión (Estados Unidos, 1996; Institute of Medicine (IOM), 2015)	Fecha de tratamiento (Institute of Medicine (IOM), 2015)
Fecha de alta (Estados Unidos, 1996; Institute of Medicine (IOM), 2015)	Fecha de visita (Institute of Medicine (IOM), 2015)
Fecha de deceso (Estados Unidos, 1996; Institute of Medicine (IOM), 2015)	Números de serie de dispositivos identificadores (Estados Unidos, 1996)
Códigos de tratamiento (Institute of Medicine (IOM), 2015)	Información recolectada durante un servicio sanitario (Canadá, 2000)
Información psiquiátrica (Nuevo León, 2008)	Laboratorios que analizaron las pruebas médicas (Meystre et al., 2010)
Intervenciones quirúrgicas (Nuevo León, 2008)	Números de registros de ensayos clínicos (El Emam, 2014)

Tabla 30 Catálogo de datos médicos del ámbito sanitario de España

3.2.2 Fase 2. Validación del modelo

Esta fase provee de los pasos necesarios para diseñar el instrumento de medida, recoger los datos de las encuestas y analizar los datos para

finalmente obtener un modelo de datos médicos validado para el ámbito sanitario de España.

3.2.2.1 *Diseño del instrumento de medida y recogida de datos*

El número de atributos obtenido en la fase anterior sigue siendo significativamente grande para que pueda ser utilizado, de forma eficiente, en distintos procesos de anonimización o en el diseño de ficheros o bases de datos (*privacy by design*). Además, se debe validar que el conjunto de atributos médicos obtenidos de la normativa y los estudios científicos analizados en el Capítulo 1, presentan utilidad para los profesionales relacionados con el área de conocimiento. Debido a esto y aunque los atributos médicos se han obtenido de fuentes fiables como legislación y estudios sobre protección de datos médicos, aun así, se recoge el criterio de profesionales vinculados con el dominio de la medicina como criterio de validación.

La hipótesis del cuestionario que se utilizó para el desarrollo de esta fase fue:

H₀: Es posible obtener un conjunto de atributos clínicos consensuado por expertos que contribuya a descubrir relaciones entre atributos y defina la importancia relativa de cada uno de ellos válido en España.

Durante esta fase, se realiza una revisión externa de todos los atributos clínicos obtenidos en la fase anterior. Esta revisión, se realiza mediante cuestionarios en el que participaron 169 profesionales tales como:

personal médico, personal investigador, así como, personal vinculado con el tratamiento de datos médicos.

Los cuestionarios se pusieron disponibles de forma online para ser accedidos a través de Internet. El cuestionario se estructuró en tres secciones principales, tal y como puede verse en el cuestionario completo en el ANEXO 4:

- Primera parte: En esta primera parte del cuestionario denominada “Antes de empezar” se informa de la finalidad que se pretende conseguir con la realización del mismo y las cuestiones relacionadas con el anonimato. Además, se incluyen ocho preguntas cerradas, de tipo sugerido, que pretenden recabar datos demográficos del participante (edad, nacionalidad, cargo que desempeña, profesión, años de experiencia, etc.).
- Segunda parte: La segunda parte titulada “Cuestionario sobre el grado de importancia”, está formada por una única pregunta en la que se intentan extraer los atributos más utilizados por el personal encuestado. Para ello, se muestran tablas que recogen los atributos organizados en las categorías mostradas en la anterior Fase 1 de la metodología presentada en el Capítulo 2 de esta tesis, para poder evaluarlo mediante una escala Likert de 5 niveles:
 - Nunca o casi nunca (nivel: 1).
 - A veces (nivel: 2).
 - Normalmente (nivel: 3).
 - Con frecuencia (nivel: 4).
 - Siempre (nivel: 5).

- Tercera parte: En esta última parte denominada “Para terminar”, se realizan dos preguntas. La primera pregunta es abierta, en la que se intenta recabar del encuestado algún atributo que no se haya incluido anteriormente y que considera importante. En la segunda pregunta, se utiliza también una escala Likert de 5 niveles para evaluar la utilidad del modelo de datos médicos que garantice el cumplimiento de la LOPD después de realizar un proceso de anonimización.

El cuestionario se creó con preguntas cortas y concisas, diseñadas especialmente para los dos grupos de profesionales para los que se han diseñado los cuestionarios evitando así, el uso de negaciones que suscitarán ambigüedades al encuestado.

Los cuestionarios fueron revisados y validados por 2 investigadores de la Universidad Europea expertos en análisis de datos, un médico de la Universidad Europea, un técnico informático de la Asociación de Técnicos de Informática (en adelante ATI) y un experto en informática médica del departamento de informática del hospital Central de la Defensa Gómez Ulla. Tras dicha revisión se depuraron las preguntas del cuestionario según los comentarios aportados por los expertos. La mayoría de los comentarios se referían a mejorar la comprensión. A este respecto, tanto por la parte médica como por la informática, se señalaron dificultades con la comprensión del cuestionario. Por ello, se decidió crear dos cuestionarios con el mismo contenido, pero cada uno de ellos dirigido al área específica usando la terminología propia de cada experto.

Posteriormente se reordenaron los atributos clínicos en las diferentes tablas. Así, por un lado, los atributos que tienen una definición más amplia de información se colocaron al final de cada una de las tablas. Y, por otro lado, los datos más específicos se situaron en la parte superior de cada tabla. El propósito fue evitar la fatiga del participante cuando intenta pensar sobre qué atributos podría abarcar una amplia definición.

Para establecer el reclutamiento y selección de los participantes, se desarrolló un listado de los Colegios Oficiales de Médicos, asociaciones médicas, fundaciones médicas, hospitales recogidos en Catálogo Nacional de Hospitales proporcionado por el Ministerio de Sanidad y Servicios Sociales e Igualdad del Gobierno de España. Por último, se pidió la participación de agentes relacionados con el entorno de medicina y la protección de datos como la Asociación Profesional Española de Privacidad (APEP) y de ATI, así como, al personal médico interesado en participar. La distribución pública del acceso al cuestionario se incluyó en un correo diseñado para tal fin. Este correo incluyó una breve descripción sobre el marco y objetivos de la investigación y el enlace al propio cuestionario.

3.2.2.2 Análisis de los datos

El proceso de recogida de datos recopiló un total de 169 cuestionarios, pero de estos, solo 120 eran cuestionarios en los que el encuestado había realizado el cuestionario de forma completa y, por tanto, podían ser utilizados, ya que el resto no tenían respuestas en las primeras preguntas de atributos o se había detectado que se respondía siempre la

primera respuesta de cada pregunta sin llegar a completar categorías que pudieran ser utilizados en el análisis. Por ello, el margen de error fue del 7,5% con un intervalo de confianza del 90%.

La mayoría de los encuestados tenían entre 36 y 49 años (42,37%) o de 50 a 65 años (33,89%) (Figura 6) y en un 96,60% provenían de España (Figura 7).

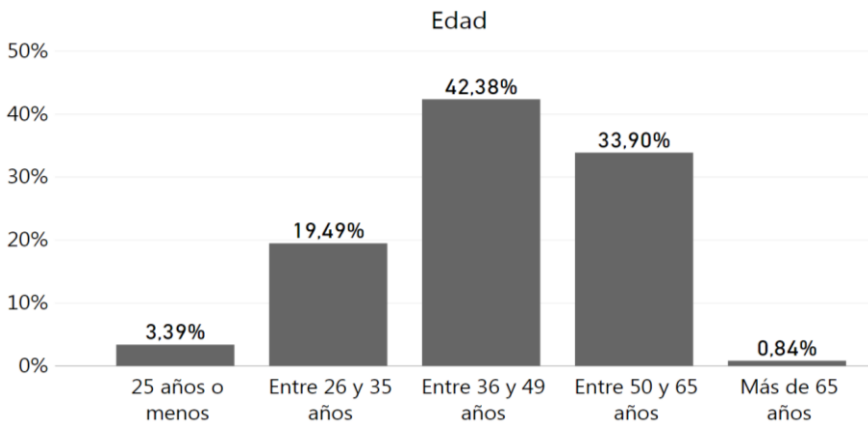


Figura 6 Datos demográficos relacionados con la edad de los encuestados

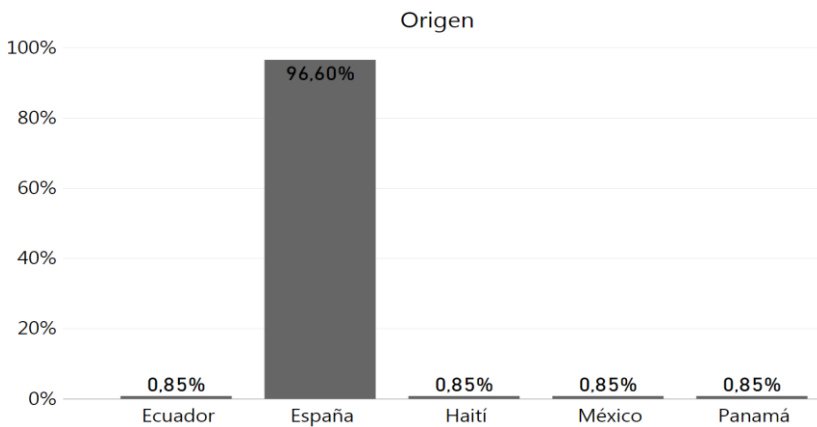


Figura 7 Datos demográficos relacionados con el país de origen de los encuestados

El sector más representado por los encuestados es el de salud pública (55,93%), seguido del sector de educación y formación (12,71%). La muestra estaba homogéneamente balanceada en relación a la especialidad médica de los profesionales de la salud. Además, los profesionales contaban con alta experiencia en su trabajo (41,52%). Por último, aproximadamente la mitad de los profesionales encuestados se dedicaba al cuidado de pacientes (49,15%), seguidos por personal dedicado a la investigación en el área de la salud (14,41%) (ver Tabla 31).

Datos demográficos relacionados con la profesión			
SECTOR		ESPECIALIDAD	
Informática y Telecomunicaciones	1,69%	Otros	9,32%
Administraciones públicas	0,85%	Dermatología	1,69%
Acción humanitaria	0,85%	Alergología	1,69%
Educación y Formación	12,71%	Neurocirugía	0,85%
Organismos oficiales	0,85%	Obstetricia y ginecología	16,95%
I+D	3,39%	Medicina del trabajo	2,54%
Salud pública	55,93%	Cirugía ortopédica y traumatología	4,23%
Salud privada	9,32%	Epidemiología	4,23%
Salud militar	2,54%	Ingeniería biomédica	1,69%
Gestión hospitalaria	5,08%	Medicina interna	14,41%
Servicios y consultoría informática	5,93%	Gestión general de aplicaciones y sistemas	2,54%
Entorno laboral	0,85%	Psiquiatría	6,78%
EXPERIENCIA		Medicina intensiva	1,69%
Ninguna	0,85%	Microbiología y parasitología	5,08%
Menos de 5 años	19,49%	Educación y formación	5,08%
Entre 5 y 9 años	22,03%	Medicina familiar y comunitaria	2,54%
Entre 10 y 15 años	16,10%	Infectología	2,54%
Más de 15 años	41,52%	Farmacología clínica	2,54%
TAREA QUE DESARROLLA		Anestesiología y reanimación	1,69%
MIR	5,93%	Formador, profesor, investigador	1,69%
Cuidado de pacientes	49,15%	Diseñador, analista orgánico	0,85%
Investigación	14,41%	Técnico sanitario	0,85%
Gestión hospitalaria	12,71%	Técnico especialista, integrador, etc.	4,24%
Gestión TIC	5,08%	Jefe de proyecto	1,69%
Sin respuesta	12,71%	Programador	1,69%
		Consultor, auditor	0,85%

Tabla 31 Resultados de datos demográficos relacionados con la profesión de los encuestados

En relación a la protección de datos, la mayoría de los encuestados tenía un nivel de familiaridad de usuario/paciente (64,40%) (Figura 8); disponiendo el 42,37% de formación a nivel básico en esta materia (Figura 9) y el 27,12% nivel alto o de responsable de ficheros de datos (Figura 8).

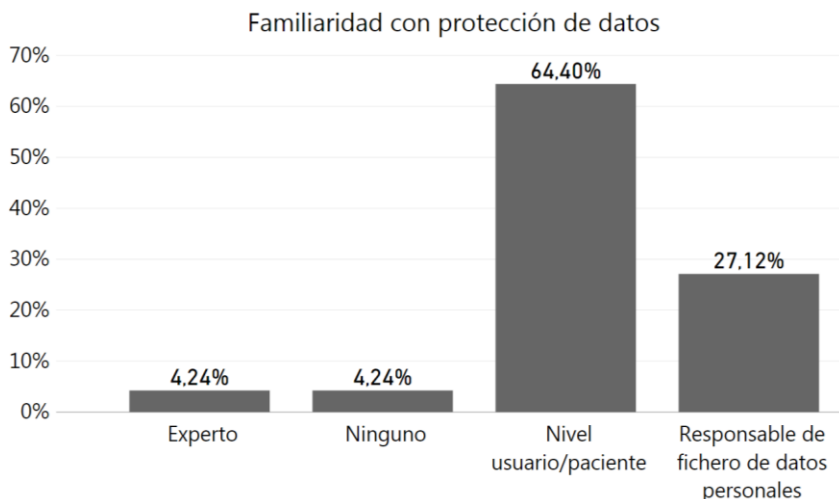


Figura 8 Datos demográficos relacionados con la familiaridad con protección de datos

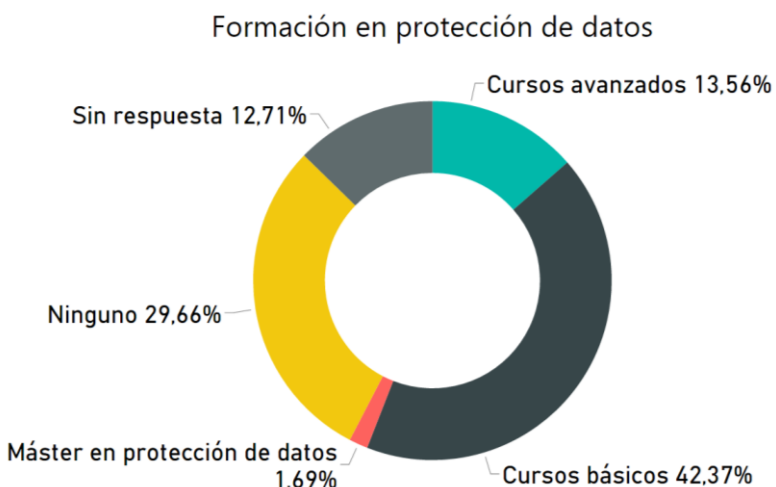


Figura 9 Datos demográficos relacionados con la formación en protección de datos

Los programas informáticos utilizados en esta fase de análisis de datos son los siguientes: *IBM SPSS Statistics Version 22* (George & Mallery, 2016) para realizar el análisis descriptivo y el análisis factorial exploratorio; e *IBM SPSS AMOS 24* (Meyers, Gamst, & Guarino, 2012) y *SAS University Edition* (O'Rourke & Hatcher, 2013) para realizar el análisis factorial confirmatorio.

Análisis previo

En primer lugar, como medida de control, se ha llevado a cabo un análisis exploratorio con el fin de detectar posibles errores durante la recogida de datos, así como para comprobar la viabilidad del análisis factorial. Posteriormente se han analizado los datos a través de un análisis descriptivo de todas las variables o atributos médicos del estudio (medias, desviación estándar, mediana, mínimo y máximo y las frecuencias absolutas y relativas). Además, se realizó una revisión de los diagramas de caja para determinar los errores de entrada de datos y el coeficiente de variación para comprobar la homogeneidad de los datos.

Con el propósito de eliminar los datos con poca importancia para los encuestados, y tal y como se describió en el Capítulo 2, se siguieron los siguientes criterios de eliminación de atributos denominados “poco importantes”, por tener todos ellos una baja importancia según los expertos (media inferior a 2, mediana menor o igual a 2, moda de 1 y asimetría negativa).

A continuación, se muestra de forma detallada el análisis exploratorio llevado a cabo por las categorías inicialmente definidas. Con la finalidad de mejorar la comprensión lectora de esta parte del análisis de datos, se han mostrado los histogramas vinculados a los atributos médicos que cumplen los criterios de baja importancia y que han sido eliminados.

(1) Datos personales de carácter general

En esta categoría los atributos médicos que hay que resaltar son los atributos DPCG4 (Persona de contacto del paciente) y DPCG5 (Estado civil). En ambos casos las medias, las medianas y las modas se encuentran por debajo de 2 (ver atributos resaltados en la Tabla 32) lo que indica que para los encuestados tienen poca importancia.

Análisis descriptivo: Datos personales de carácter general							
Variable estadística	Atributo clínico	Media	Mediana	Moda	Desviación estándar	Kolmogorov-Smirnova	
						Estadístico de prueba	Signatura asintótica ^a
DPCG1	Nombre	2,69	2,00	1	1,654	0,229	0,000
DPCG2	Primer apellido	2,68	2,00	1	1,615	0,220	0,000
DPCG3	Segundo apellido	2,40	2,00	1	1,542	0,252	0,000
DPCG4	Persona de contacto del paciente	1,93	1,50	1	1,235	0,278	0,000
DPCG5	Estado civil	1,70	1,00	1	1,026	0,328	0,000
DPCG6	Género	3,68	5,00	5	1,561	0,309	0,000
DPCG7	Nacionalidad	2,80	3,00	1	1,435	0,173	0,000
DPCG8	Fecha de nacimiento	3,33	4,00	5	1,508	0,204	0,000
DPCG9	Edad o elemento indicativo de edad	3,89	5,00	5	1,333	0,314	0,000
DPCG10	Sexo	4,05	5,00	5	1,437	0,396	0,000
	TOTAL DE MUESTRA	120					

a. Corrección de Lilliefors

Tabla 32 Análisis descriptivo para las variables englobadas dentro de datos personales de carácter general

Además, los histogramas de ambos atributos médicos (DPCG4 y DPCG5) poseen asimetría negativa (ver Figura 10 y Figura 11). Por todo ello, se ha decidido eliminar estos atributos médicos.

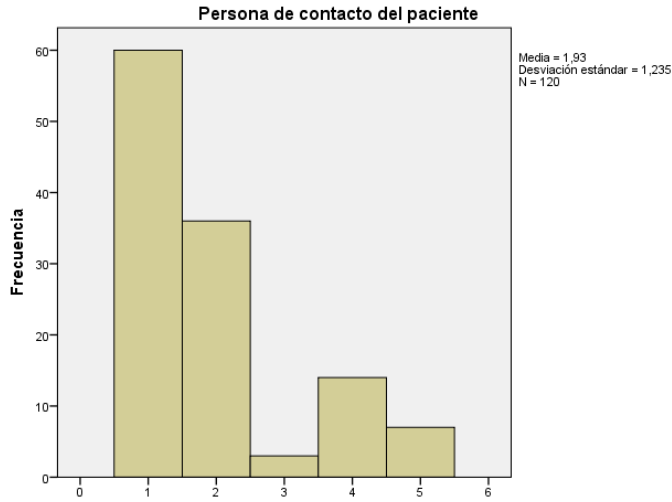


Figura 10 Histograma del atributo DPCG4 (Persona de contacto del paciente) con asimetría negativa

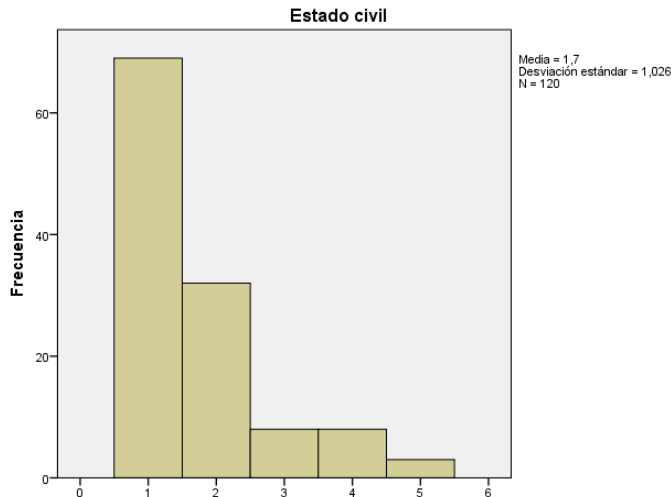


Figura 11 Histograma del atributo DPCG5 (Estado civil) con asimetría negativa

Los histogramas de los atributos médicos que no se han eliminado, se encuentran recogidos en la categoría “(1) Datos personales de carácter general” del ANEXO 5.

(2) Datos personales de contacto

En la categoría “Datos personales de contacto” los atributos médicos que hay que resaltar son los atributos DPC2 (Fax) y DPC3 (Correo electrónico). En ambos casos las medias, las medianas y las modas se encuentran por debajo de 2 (ver atributos resaltados Tabla 33), indicando poca importancia.

Análisis descriptivo: Datos personales de contacto							
Variable estadística	Atributo clínico	Media	Mediana	Moda	Desviación estándar	Kolmogorov-Smirnova	
						Estadístico de prueba	Signatura asintótica ^a
DPC1	Teléfono	2,21	2,00	1	1,347	0,286	0,000
DPC2	Fax	1,16	1,00	1	0,594	0,513	0,000
DPC3	Correo electrónico	1,30	1,00	1	0,717	0,462	0,000
	TOTAL DE MUESTRA	120					

a. Corrección de Lilliefors

Tabla 33 Análisis descriptivo para las variables englobadas dentro de datos personales de contacto

Tras analizar los histogramas (Figura 12 y Figura 13) de estas variables estadísticas con poca importancia, se observa que poseen asimetría negativa. Por todo ello, se ha decidido eliminar estos atributos médicos.

El histograma del atributo médico que no se ha eliminado, se encuentran recogidos en la categoría “(2) Datos personales de contacto” del ANEXO 5.

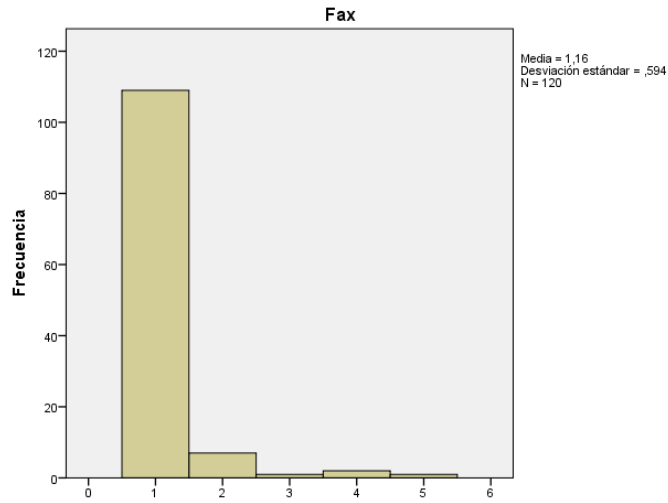


Figura 12 Histograma del atributo DPC2 (Fax) con asimetría negativa

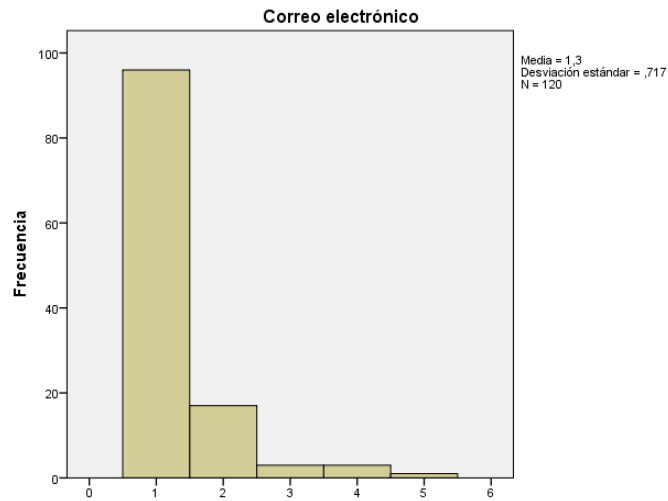


Figura 13 Histograma del atributo DPC3 (Correo electrónico) con asimetría negativa

(3) Datos personales asociados a geolocalización

En esta categoría los atributos médicos que poseen valores estadísticos con valores por debajo de los criterios establecidos de importancia (ver

atributos resaltados Tabla 34) son los atributos médicos: DPAG1 (Calle), DPAG4 (Distrito), DPAG5 (Dirección) y DPAG8 (Código postal).

Análisis descriptivo: Datos personales asociados a geolocalización							
						Kolmogorov-Smirnova	
Variable estadística	Atributo clínico	Media	Mediana	Moda	Desviación estándar	Estadístico de prueba	Signatura asintótica ^a
DPAG1	Calle	1,88	1,00	1	1,382	0,370	0,000
DPAG2	Ciudad	2,19	2,00	1	1,398	0,278	0,000
DPAG3	Municipio	2,13	2,00	1	1,345	0,274	0,000
DPAG4	Distrito	1,95	1,00	1	1,222	0,307	0,000
DPAG5	Dirección	1,92	1,00	1	1,394	0,370	0,000
DPAG6	Comunidad autónoma	2,31	2,00	1	1,389	0,235	0,000
DPAG7	País de nacimiento	2,39	2,00	1	1,428	0,235	0,000
DPAG8	Código postal	1,76	1,00	1	1,108	0,353	0,000
	TOTAL DE MUESTRA	120					

a. Corrección de Lilliefors

Tabla 34 Análisis descriptivo para las variables englobadas dentro de datos personales asociados a geolocalización

A continuación se muestran, los histogramas de los atributos médicos DPAG1, DPAG4, DPAG5 y DPAG8 poseen asimetría negativa (Figura 14, Figura 15, Figura 16 y Figura 17). Por todo ello, se ha decidido eliminar estos atributos médicos.

Los histogramas de los atributos médicos que no se han eliminado, se encuentran recogidos en la categoría “(3) Datos personales asociados a geolocalización” del ANEXO 5.

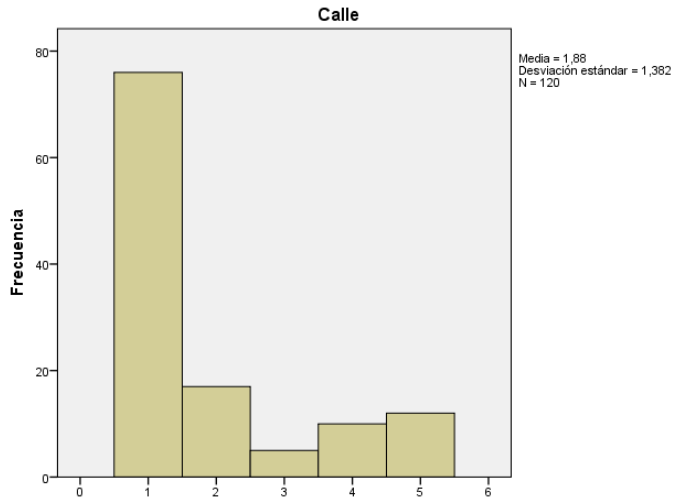


Figura 14 Histograma del atributo DPAG1 (Calle) con asimetría negativa

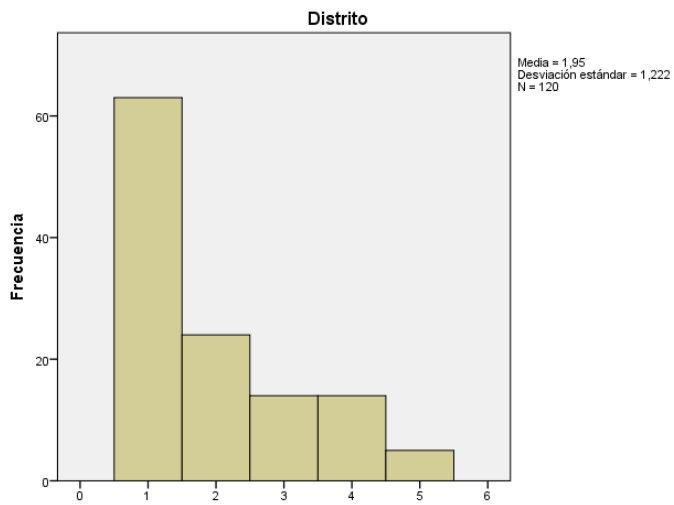


Figura 15 Histograma del atributo DPAG4 (Distrito) con asimetría negativa

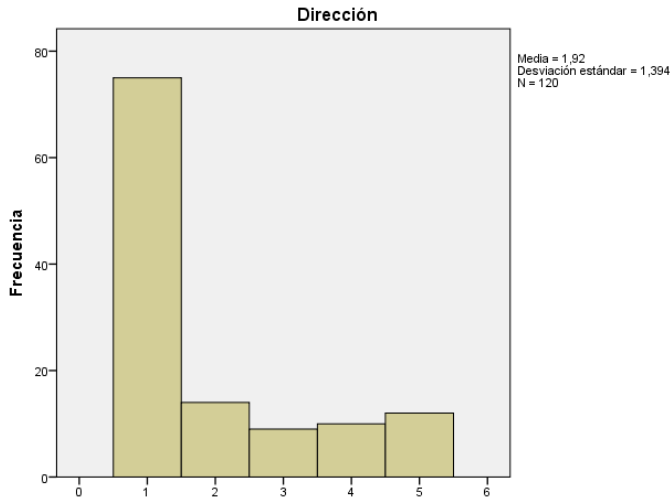


Figura 16 Histograma del atributo DPAG5 (Dirección) con asimetría negativa

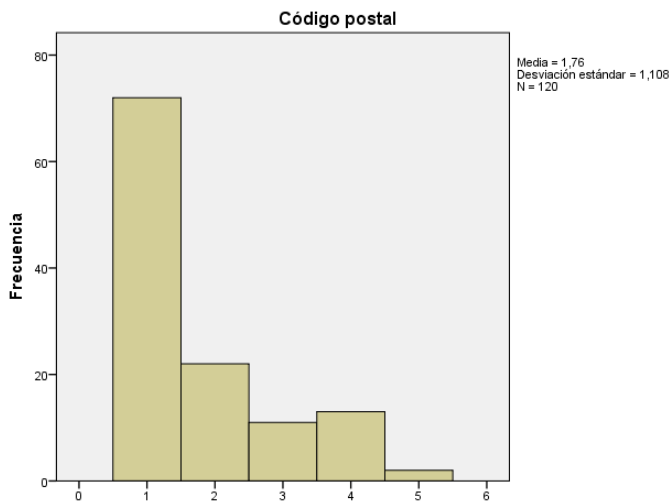


Figura 17 Histograma del atributo DPAG8 (Código postal) con asimetría negativa

(4) Datos personales relacionados con creencias, cultura, raza, etnia u opiniones

En esta categoría los atributos médicos que cumplen los criterios para ser eliminados del análisis factorial posterior son: DPRCCREO2 (Necesidades

personales), DPRCCREO3 (Orientación sexual), DPRCCREO4 (Origen social) y DPRCCREO5 (Pluralidad de nacimiento). En todas las variables indicadas anteriormente los valores de las medias, las medianas y las modas se encuentran por debajo de 2 (ver atributos resaltados Tabla 35) lo que indica que para los encuestados tienen poca importancia.

Análisis descriptivo: Datos personales relacionados con creencias, cultura, raza, etnia u opiniones							
Variable estadística	Atributo clínico	Media	Mediana	Moda	Desviación estándar	Kolmogorov-Smirnova	
						Estadístico de prueba	Signatura asintótica ^a
DPRCCREO1	Origen étnico	2,20	2,00	1	1,294	0,248	0,000
DPRCCREO2	Necesidades personales	1,78	1,00	1	1,096	0,302	0,000
DPRCCREO3	Orientación sexual	1,68	1,00	1	1,195	0,383	0,000
DPRCCREO4	Origen social	1,59	1,00	1	0,921	0,348	0,000
DPRCCREO5	Pluralidad de nacimiento	1,46	1,00	1	0,869	0,409	0,000
	TOTAL DE MUESTRA	120					

a. Corrección de Lilliefors

Tabla 35 Análisis descriptivo para las variables englobadas dentro de datos personales relacionados con creencias, cultura, raza, etnia u opiniones

Además, los histogramas de los atributos médicos DPRCCREO2, DPRCCREO3, DPRCCREO4, DPRCCREO5 poseen asimetría negativa (Figura 18, Figura 19, Figura 20 y Figura 21). Por todo ello, se ha decidido eliminar estos atributos médicos.

El histograma del atributo médico que no se ha eliminado, se encuentra recogido en la categoría “(4) Datos personales relacionados con creencias, cultura, raza, etnia u opiniones” del ANEXO 5.

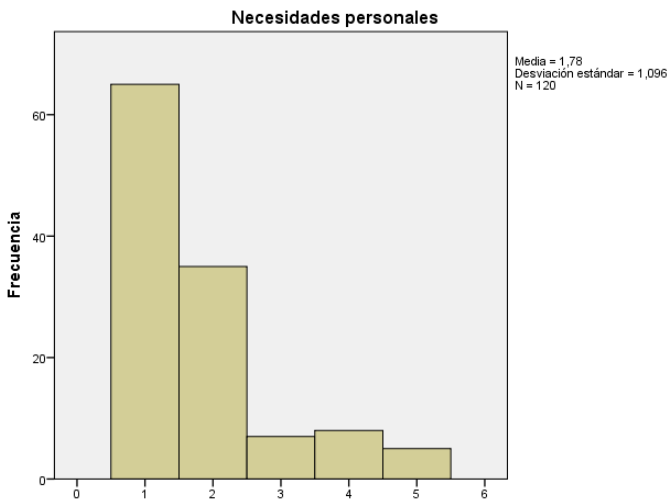


Figura 18 Histograma del atributo DPRCCREO2 (Necesidades personales) con asimetría negativa

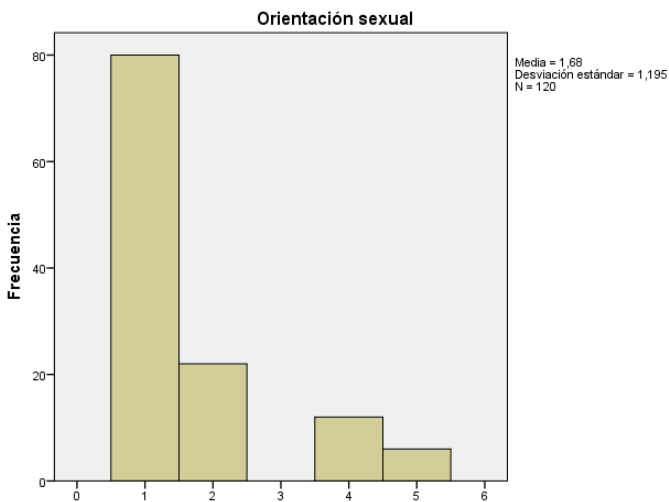


Figura 19 Histograma del atributo DPRCCREO3 (Orientación sexual) con asimetría negativa

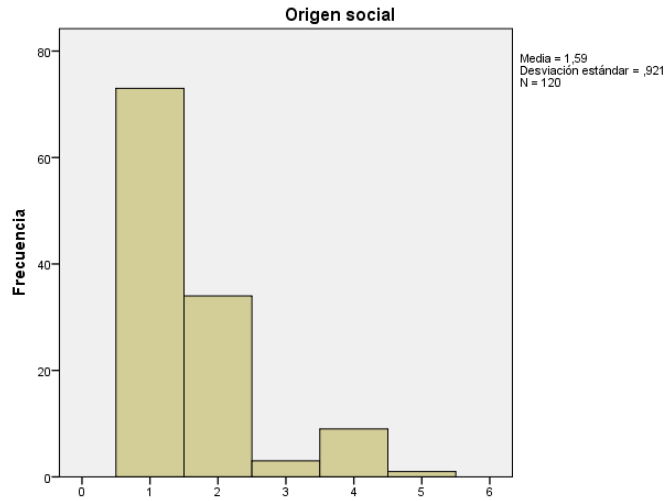


Figura 20 Histograma del atributo DPRCCREO4 (Origen social) con asimetría negativa

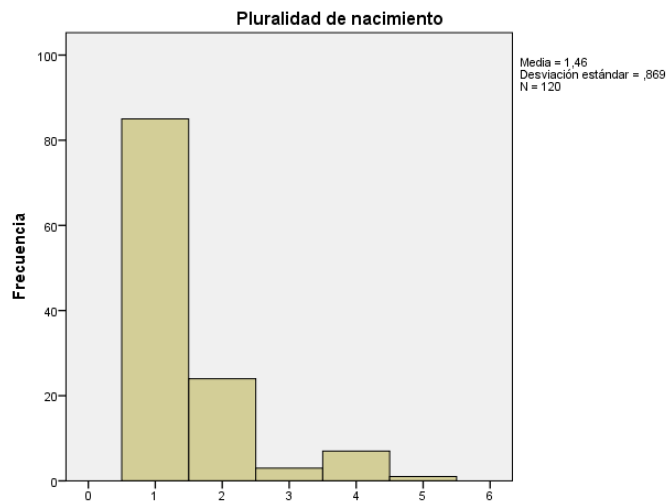


Figura 21 Histograma del atributo DPRCCREO5 (Pluralidad de nacimiento) con asimetría negativa

(5) Datos personales relacionados con la profesión

En esta categoría de atributos médicos relacionados con la profesión del paciente, los atributos que cumplen con los criterios para ser eliminados

por su baja importancia son: DPRP1 (Años de escolaridad), DPRP2 (Historial de educación), DPRP3 (Historial de empleo), DPRP4 (Posición dentro de la familia), DPRP5 (Nombre del último empleo) y DPRP6 (Organización en la que trabaja), ya que los valores de la media, la mediana y la moda se encuentran por debajo o igual de 2 (ver atributos resaltados Tabla 36).

Análisis descriptivo: Datos personales relacionados con la profesión							
Variable estadística	Atributo clínico	Media	Mediana	Moda	Desviación estándar	Kolmogorov-Smirnova	
						Estadístico de prueba	Signatura asintótica ^a
DPRP1	Años de escolaridad	1,64	1,00	1	1,002	0,347	0,000
DPRP2	Historial de educación	1,80	1,00	1	1,097	0,284	0,000
DPRP3	Historial de empleo	1,93	2,00	1	1,158	0,282	0,000
DPRP4	Posición dentro de la familia (padre, hermano mayor, etc)	1,54	1,00	1	0,943	0,367	0,000
DPRP5	Nombre del último empleo	1,68	1,00	1	1,161	0,361	0,000
DPRP6	Organización en la que trabaja	1,76	1,00	1	1,243	0,354	0,000
DPRP7	Profesión	2,53	2,00	2	1,359	0,244	0,000
	TOTAL DE MUESTRA	120					

a. Corrección de Lilliefors

Tabla 36 Análisis descriptivo para las variables englobadas dentro de datos personales relacionados con la profesión

Además, los histogramas de los atributos médicos DPRP1, DPRP2, DPRP3, DPRP4, DPRP5 y DPRP6 poseen asimetría negativa (Figura 22, Figura 23,

Figura 24, Figura 25, Figura 26 y Figura 27). Por todo ello, se ha decidido eliminar estos atributos médicos.

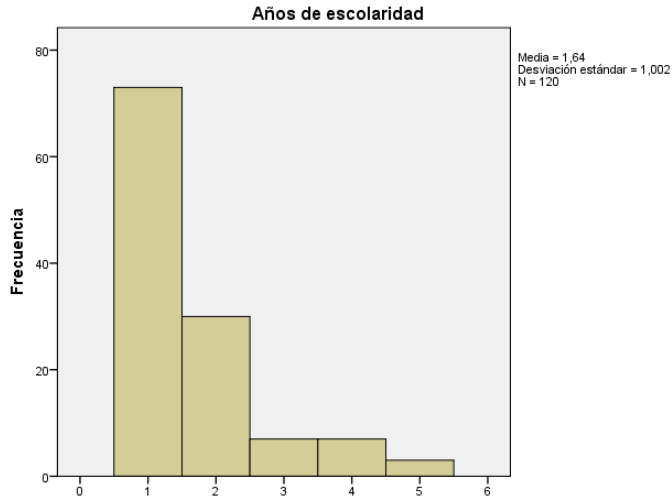


Figura 22 Histograma del atributo DPRP1 (Años de escolaridad) con asimetría negativa

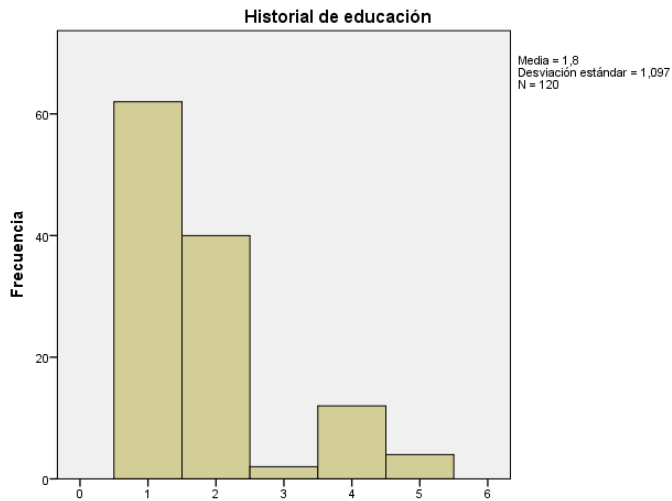


Figura 23 Histograma del atributo DPRP2 (Historial de educación) con asimetría negativa

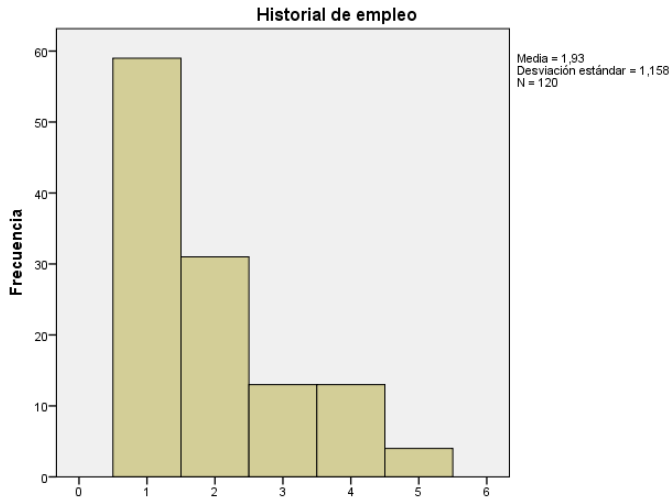


Figura 24 Histograma del atributo DPRP3 (Historial de empleo) con asimetría negativa

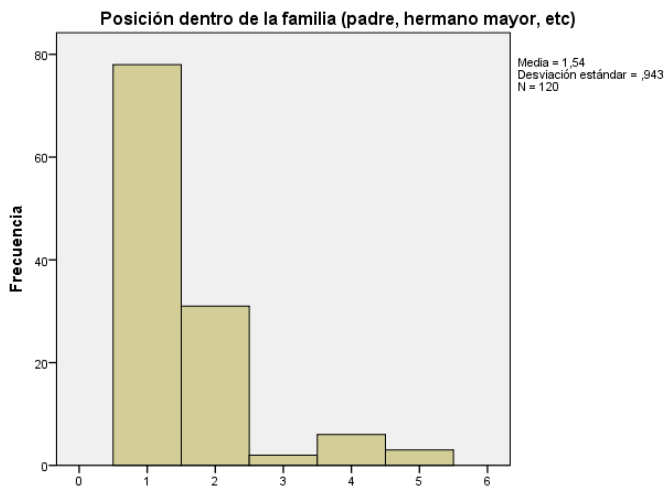


Figura 25 Histograma del atributo DPRP4 (Posición dentro de la familia) con asimetría negativa

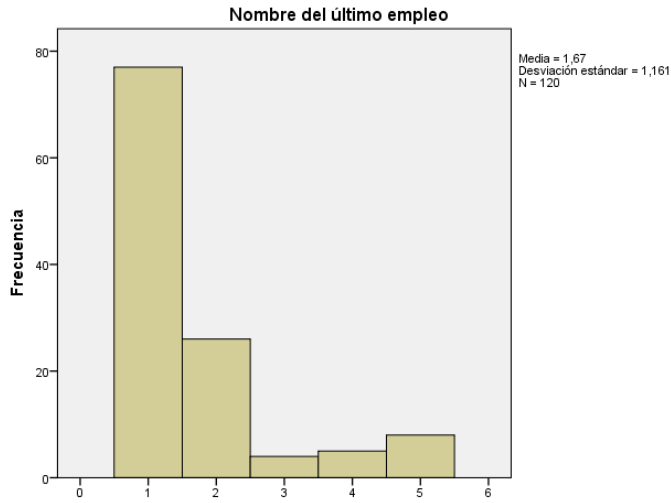


Figura 26 Histograma del atributo DPRP5 (Nombre del último empleo) con asimetría negativa

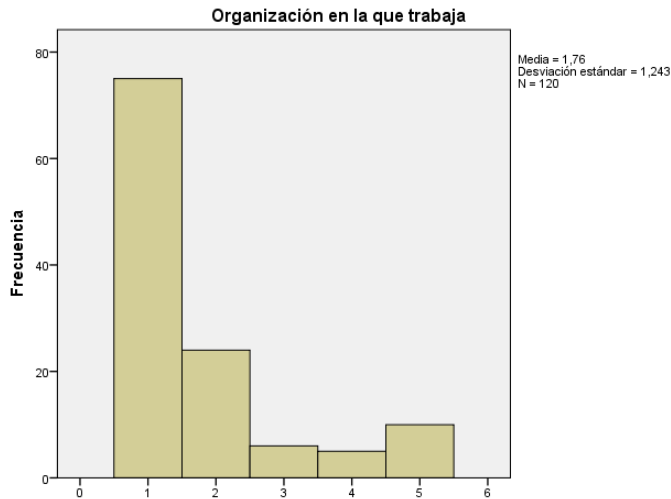


Figura 27 Histograma del atributo DPRP6 (Origen en la que trabaja) con asimetría negativa

El histograma del atributo médico que no se ha eliminado, se encuentra recogido en la categoría “(5) Datos personales relacionados con la profesión” del ANEXO 5.

(6) Datos personales asociados a donaciones

En esta categoría ambos atributos médicos: DPAD1 (Donación de sustancia corporal) y DPAD2 (Donación de parte corporal) cumplen los criterios para ser eliminados (ver atributos resaltados Tabla 37), lo que indica que para los encuestados tienen poca importancia.

Análisis descriptivo: Datos personales asociados a donaciones							
						Kolmogorov-Smirnova	
Variable estadística	Atributo clínico	Media	Mediana	Moda	Desviación estándar	Estadístico de prueba	Signatura asintótica ^a
DPAD1	Donación de sustancia corporal	1,28	1,00	1	0,698	0,478	0,000
DPAD2	Donación de parte corporal	1,18	1,00	1	0,534	0,509	0,000
	TOTAL DE MUESTRA	120					

a. Corrección de Lilliefors

Tabla 37 Análisis descriptivo para las variables englobadas dentro de datos personales asociados a donaciones

Además, los histogramas de los atributos médicos DPAD1 y DPAD2 poseen asimetría negativa (Figura 28 y Figura 29). Por todo ello, se ha decidido eliminar estos atributos médicos.

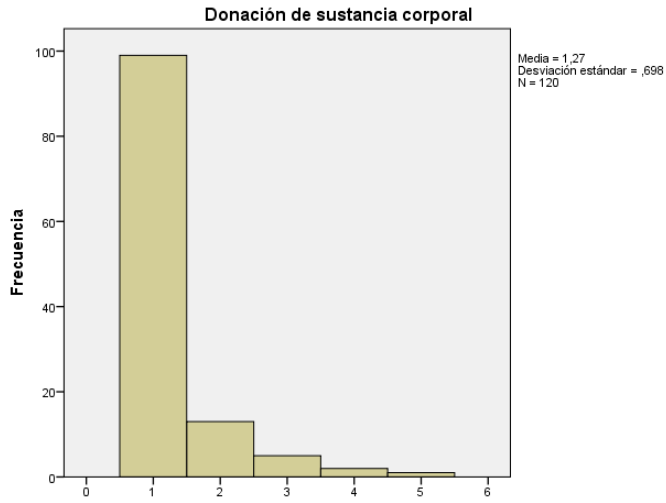


Figura 28 Histograma del atributo DPAD1 (Donación de sustancia corporal) con asimetría negativa

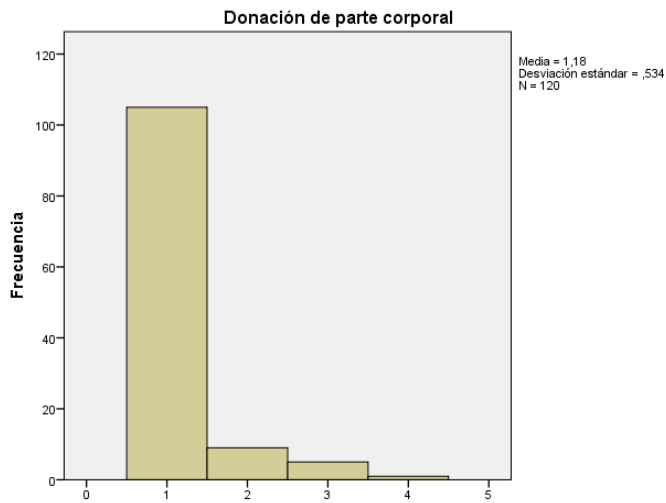


Figura 29 Histograma del atributo DPAD2 (Donación de parte corporal) con asimetría negativa

(7) Datos personales asociados a aseguradoras

En esta categoría todos los atributos médicos recogidos cumplen los criterios para ser eliminados (ver atributos resaltados Tabla 38) por su baja importancia para los encuestados.

Análisis descriptivo: Datos personales asociados a aseguradoras							
Variable estadística	Atributo clínico	Media	Mediana	Moda	Desviación estándar	Kolmogorov-Smirnova	
						Estadístico de prueba	Signatura asintótica ^a
DPAA1	Ingresos totales	1,38	1,00	1	0,881	0,452	0,000
DPAA2	Proveedor de atención médica	1,78	1,00	1	1,298	0,394	0,000
DPAA3	Número de la Seguridad Social	1,98	1,00	1	1,396	0,351	0,000
DPAA4	Número de tarjeta sanitaria	1,92	1,00	1	1,351	0,351	0,000
	TOTAL DE MUESTRA	120					

a. Corrección de Lilliefors

Tabla 38 Análisis descriptivo para las variables englobadas dentro de datos personales asociados a aseguradoras

Además, los histogramas de los atributos médicos DPAA1, DPAA2, DPAA3 y DPAA4 poseen asimetría negativa (Figura 30, Figura 31, Figura 32 y Figura 33). Por todo ello, se ha decidido eliminar estos atributos médicos.

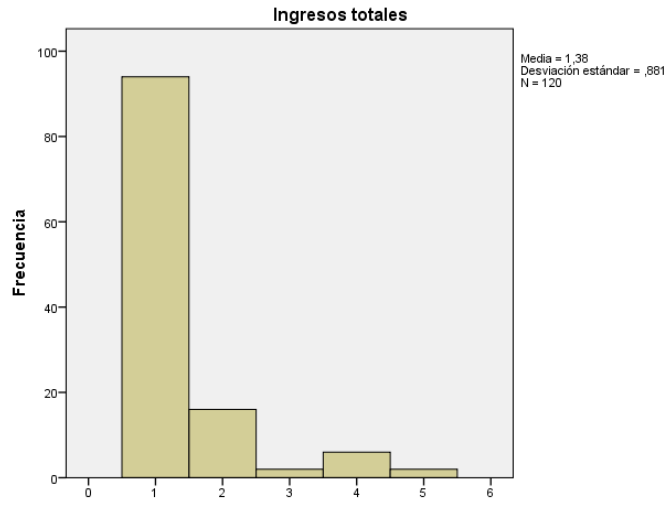


Figura 30 Histograma del atributo DPAA1 (Ingresos totales) con asimetría negativa

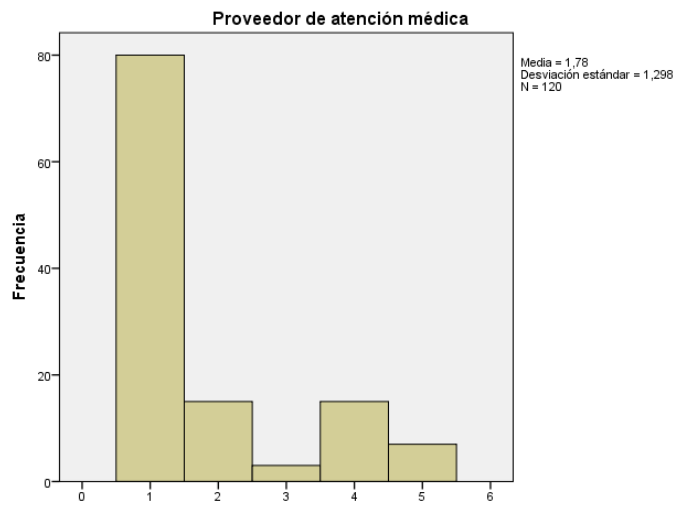


Figura 31 Histograma del atributo DPAA2 (Proveedor de atención médica) con asimetría negativa

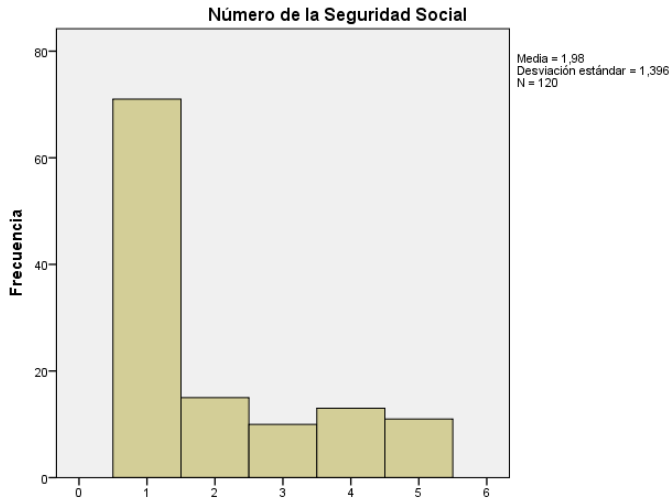


Figura 32 Histograma del atributo DPAA3 (Número de la Seguridad Social) con asimetría negativa

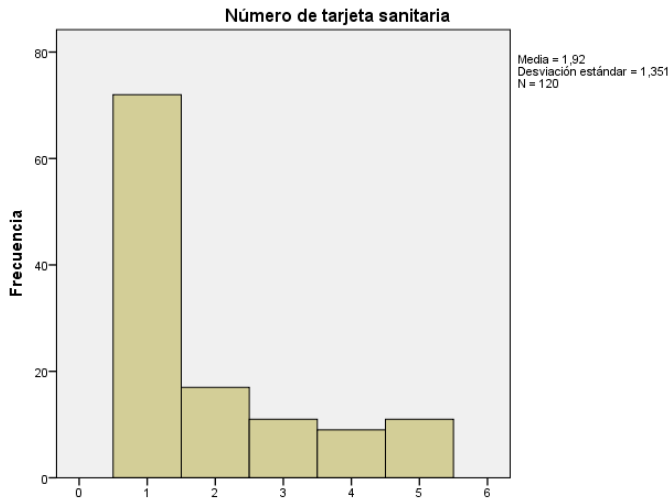


Figura 33 Histograma del atributo DPAA4 (Número de tarjeta sanitaria) con asimetría negativa

(8) Datos de dispositivos implantados

En esta categoría de atributos médicos relacionados con dispositivos implantados en paciente, los atributos médicos que cumplen con los

criterios de eliminación por su baja importancia son: DDI1 (Uso de aparatos auditivos), DDI2 (Uso de aparatos oftalmológicos), DDI3 (Uso de aparatos ortopédicos), DDI5 (Número identificativo de los dispositivos implantados); porque los valores de la media, la mediana y la moda se encuentran por debajo o igual de 2 (ver atributos resaltados Tabla 39).

Análisis descriptivo: Datos de dispositivos implantados							
Variable estadística	Atributo clínico	Media	Mediana	Moda	Desviación estándar	Kolmogorov-Smirnova	
						Estadístico de prueba	Signatura asintótica ^a
DDI1	Uso de aparatos auditivos	1,61	1,00	1	0,990	0,372	0,000
DDI2	Uso de aparatos oftalmológico	1,67	1,00	1	1,015	0,353	0,000
DDI3	Uso de aparatos ortopédicos	1,79	1,00	1	1,076	0,319	0,000
DDI4	Prótesis	2,04	2,00	1	1,191	0,276	0,000
DDI5	Número identificativo de los dispositivos implantados	1,45	1,00	1	0,858	0,425	0,000
	TOTAL DE MUESTRA	120					

a. Corrección de Lilliefors

Tabla 39 Análisis descriptivo para las variables englobadas dentro de datos de dispositivos implantados

Además, los histogramas de los atributos médicos DDI1, DDI2, DDI3 Y DDI5 poseen asimetría negativa (Figura 34, Figura 35, Figura 36 y Figura 37). Por todo ello, se ha decidido eliminar estos atributos médicos.

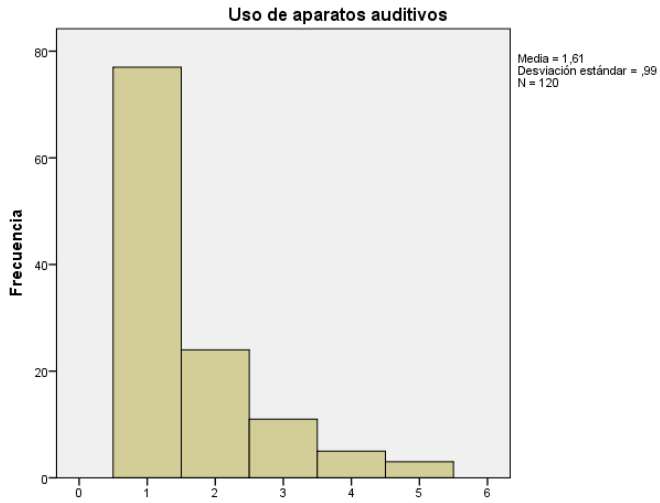


Figura 34 Histograma del atributo DDI1 con asimetría negativa

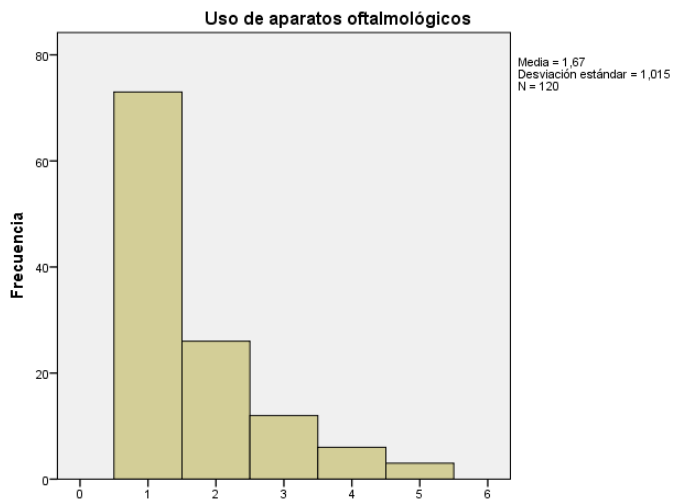


Figura 35 Histograma del atributo DDI2 con asimetría negativa

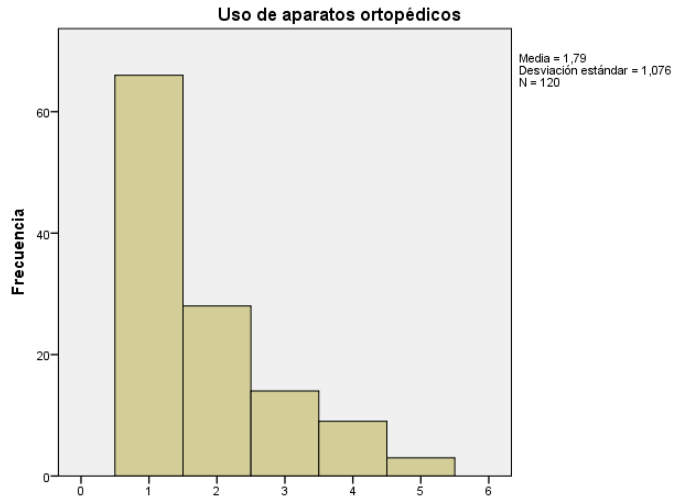


Figura 36 Histograma del atributo DDI3 con asimetría negativa

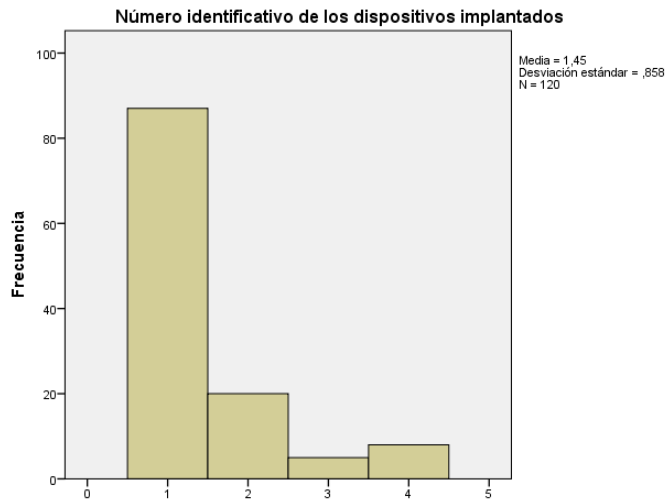


Figura 37 Histograma del atributo DDI5 con asimetría negativa

El histograma del atributo médico que no se ha eliminado, se encuentra recogido en la categoría “(6) Datos de dispositivos implantados” del ANEXO 5.

(9) Datos médicos

En esta categoría de datos médicos los siguientes atributos cumplen los criterios para ser eliminados (ver atributos resaltados Tabla 40): DM1 (Tipo sanguíneo), DM4 (Ácido desoxirribonucleico), DM23 (Imágenes fotográficas identificables), DM25 (Genética de familiar cercano), DM26 (Genética propia), DM29 (Número de placa médica), DM30 (Número de registro médico), DM31 (Número de serie de dispositivos identificadores), DM33 (Códigos de eventos adversos), lo que indica que para los encuestados tienen poca importancia.

Análisis descriptivo: Datos médicos							
Variable estadística	Atributo clínico	Media	Mediana	Moda	Desviación estándar	Kolmogorov-Smirnova	
						Estadístico de prueba	Signatura asintótica ^a
DM1	Tipo sanguíneo	1,93	1,00	1	1,320	0,327	0,000
DM2	Servicio prestado	2,25	2,00	1	1,445	0,281	0,000
DM3	Tratamiento	3,68	4,00	5	1,444	0,244	0,000
DM4	Ácido desoxirribonucleico	1,43	1,00	1	0,923	0,452	0,000
DM5	Alergias	3,36	4,00	5	1,719	0,255	0,000
DM6	Centro de salud	2,13	2,00	1	1,429	0,268	0,000
DM7	Consumo de sustancias tóxicas	3,18	3,00	5	1,581	0,191	0,000
DM8	Diagnos	4,07	5,00	5	1,333	0,325	0,000
DM9	Enfermedades	4,27	5,00	5	1,179	0,383	0,000
DM10	Estado de embarazo	3,45	4,00	5	1,544	0,226	0,000
DM11	Estado de enfermedad	3,86	4,00	5	1,398	0,276	0,000
DM12	Historial clínico	4,07	5,00	5	1,262	0,337	0,000
DM13	Información psicológica	2,64	2,00	1	1,437	0,197	0,000
DM14	Información psiquiátrica	2,71	3,00	1	1,469	0,169	0,000
DM15	Información recolectada durante un servicio sanitario	3,41	4,00	5	1,452	0,188	0,000
DM16	Intervenciones quirúrgicas	3,86	4,00	5	1,272	0,249	0,000
DM17	Laboratorios que	2,31	2,00	1	1,500	0,256	0,000

Análisis descriptivo: Datos médicos							
Variable estadística	Atributo clínico	Media	Mediana	Moda	Desviación estándar	Kolmogorov-Smirnova	
						Estadístico de prueba	Signatura asintótica ^a
	analizaron las pruebas médicas						
DM18	Medidas corporales (talla, peso)	3,25	3,00	3	1,343	0,178	0,000
DM19	Prognosis	2,67	3,00	1	1,410	0,223	0,000
DM20	Residencias de ancianos	2,00	1,00	1	1,270	0,301	0,000
DM21	Vacunas	2,39	2,00	1	1,463	0,254	0,000
DM22	Incapacidades médicas	2,58	2,50	1	1,532	0,240	0,000
DM23	Imágenes fotográficas identificables	1,61	1,00	1	1,056	0,376	0,000
DM24	Estado de discapacidad	2,45	2,00	1	1,472	0,220	0,000
DM25	Genética de familiar cercano	1,64	1,00	1	1,060	0,369	0,000
DM26	Genética propia	1,78	1,00	1	1,124	0,324	0,000
DM27	Número de historia clínica	3,62	4,50	5	1,589	0,308	0,000
DM28	Número de identificación	2,85	2,50	1	1,795	0,274	0,000
DM29	Número de placa médica	1,49	1,00	1	1,012	0,428	0,000
DM30	Número de registro médico	1,71	1,00	1	1,212	0,404	0,000
DM31	Número de serie de dispositivos identificadores	1,31	1,00	1	0,776	0,463	0,000
DM32	Códigos de diagnosis	2,26	2,00	1	1,447	0,271	0,000
DM33	Códigos de eventos adversos	1,88	1,00	1	1,245	0,303	0,000
DM34	Códigos de tratamiento	2,03	1,00	1	1,399	0,285	0,000
DM35	Fecha de admisión	3,22	3,00	5	1,479	0,178	0,000
DM36	Fecha de alta	3,52	4,00	5	1,438	0,216	0,000
DM37	Fecha de deceso	2,89	3,00	1	1,576	0,198	0,000
DM38	Fecha de recogida de muestras	2,93	3,00	1	1,565	0,200	0,000
DM39	Fecha de tratamiento	3,48	4,00	5	1,561	0,227	0,000
DM40	Fecha de visita	3,25	4,00	5	1,557	0,210	0,000
DM41	Número de registros de ensayos clínicos	2,33	2,00	1	1,445	0,231	0,000
DM42	Salud física (otra información no	2,81	3,00	1	1,519	0,186	0,000

Análisis descriptivo: Datos médicos							
Variable estadística	Atributo clínico	Media	Mediana	Moda	Desviación estándar	Kolmogorov-Smirnova	
						Estadístico de prueba	Signatura asintótica ^a
	recogida en anteriores campos)						
DM43	Salud mental (otra información mental no recogida en anteriores campos)	2,58	2,00	1	1,504	0,218	0,000
DM44	Términos médicos que identifican patologías no habituales	2,78	2,00	1	1,551	0,210	0,000
	TOTAL DE MUESTRA	120					

a. Corrección de Lilliefors

Tabla 40 Análisis descriptivo para las variables englobadas dentro de datos médicos

Además, los histogramas de los atributos médicos DM1, DM4, DM23, DM25, DM26, DM29, DM30, DM31 y DM33 poseen asimetría negativa (Figura 38, Figura 39, Figura 40, Figura 41, Figura 42, Figura 43, Figura 44, Figura 45 y Figura 46). Por todo ello, se ha decidido eliminar estos atributos.

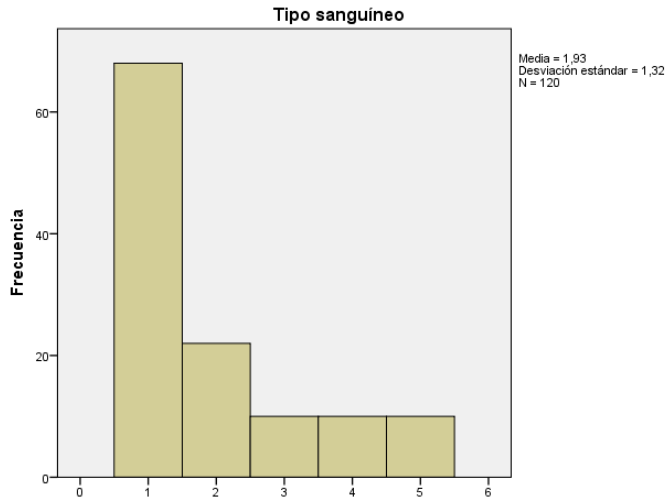


Figura 38 Histograma del atributo DM1 (Tipo sanguíneo) con asimetría negativa

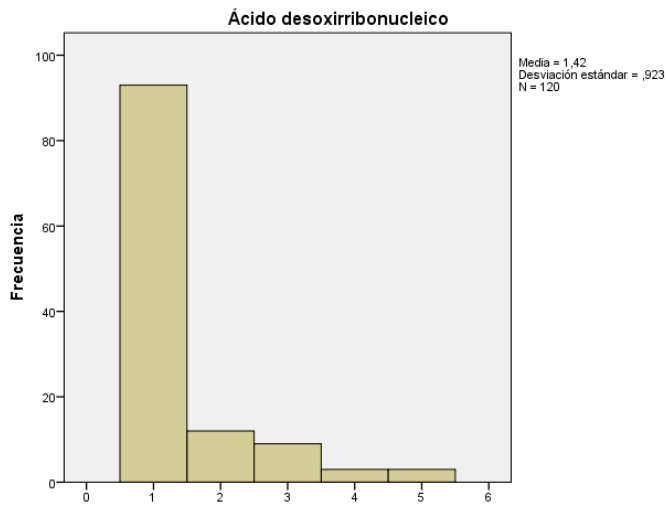


Figura 39 Histograma del atributo DM4 (Ácido desoxirribonucleico) con asimetría negativa

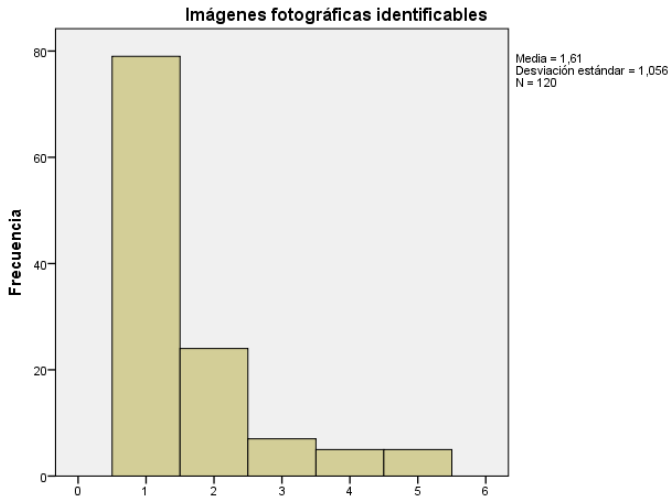


Figura 40 Histograma del atributo DM23 (Imágenes fotográficas identificables) con asimetría negativa

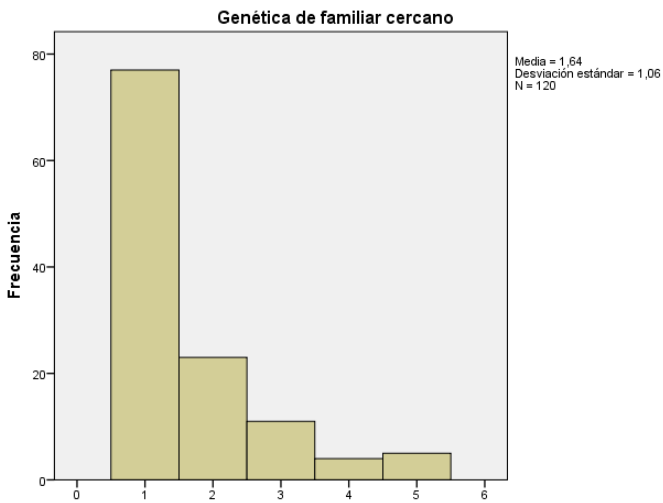


Figura 41 Histograma del atributo DM25 (Genética de familiar cercano) con asimetría negativa

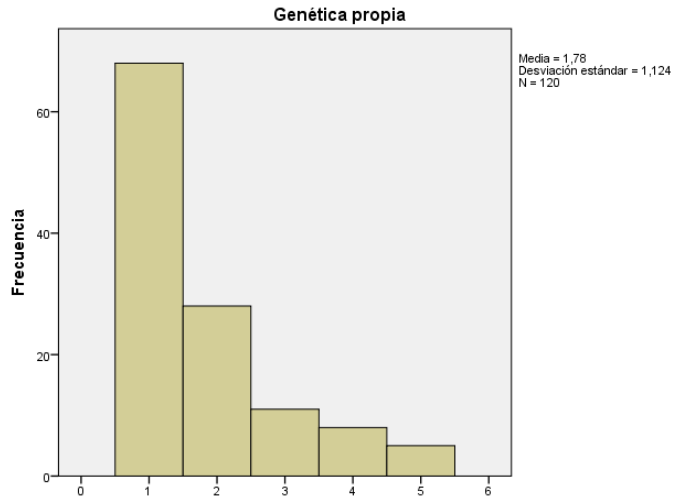


Figura 42 Histograma del atributo DM26 (Genética propia) con asimetría negativa

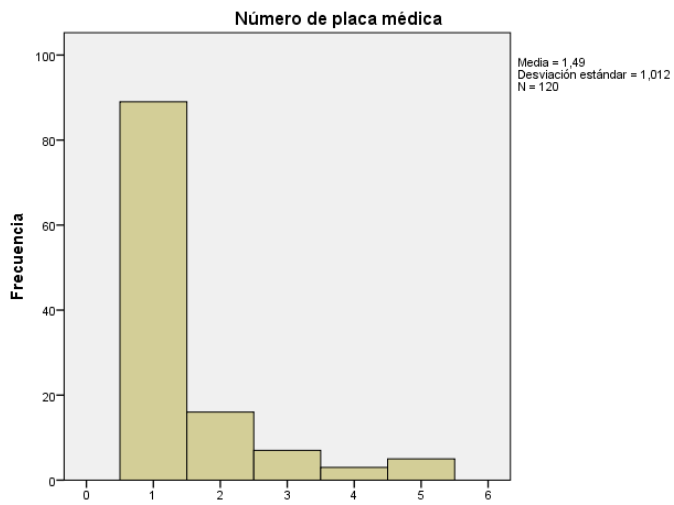


Figura 43 Histograma del atributo DM29 (Número de placa médica) con asimetría negativa

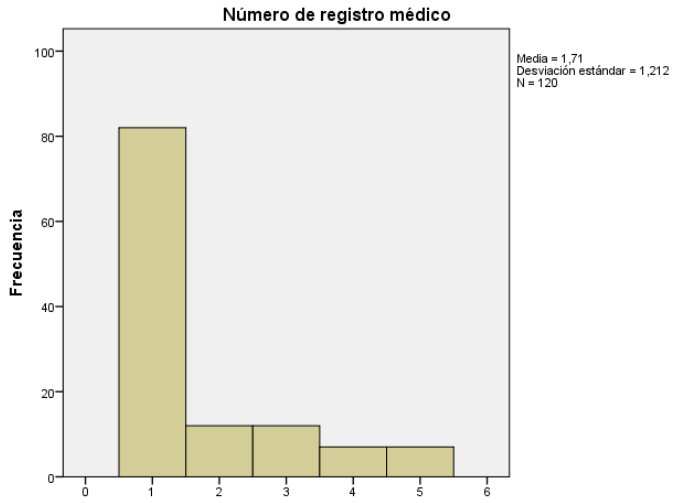


Figura 44 Histograma del atributo DM30 (Número de registro médico) con asimetría negativa

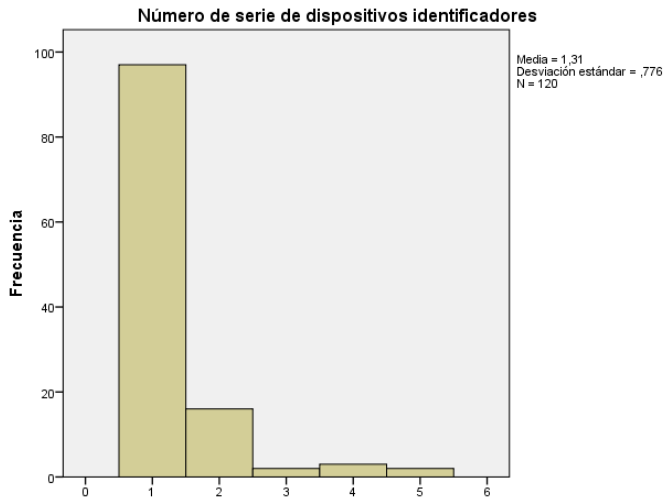


Figura 45 Histograma del atributo DM31 (Número de serie de dispositivos identificadores) con asimetría negativa

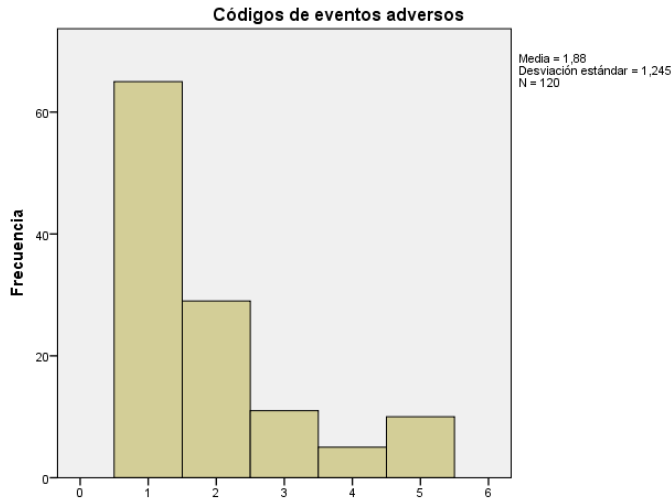


Figura 46 Histograma del atributo DM33 (Códigos de eventos adversos) con asimetría negativa

Los histogramas de los atributos médicos que no se han eliminado, se encuentran recogidos en la categoría “(7) Datos médicos” del ANEXO 5.

Conclusiones del análisis previo de datos

Se eliminaron 37 atributos médicos por tener, todos ellos, una baja importancia según los expertos (media inferior a 2, mediana menor o igual a 2, moda de 1 y asimetría negativa). Tras realizar esta eliminación, resultaron un total de 51 atributos médicos para seguir realizando el análisis estadístico. Debido a que hay categorías que han quedado sin atributos médicos, se ha realizado una nueva distribución de atributos médicos en la que se ha utilizado el criterio de significado del atributo.

Los atributos médicos relacionados con el ámbito personal han sido recogidos en la categoría “Datos personales” (ver Tabla 41). Los atributos

médicos relacionados directamente con el ámbito de salud se han agrupado en la categoría “Datos médicos” (ver Tabla 42).

Datos personales	
DPCG1	Nombre
DPCG2	Primer apellido
DPCG3	Segundo apellido
DPCG6	Género
DPCG7	Nacionalidad
DPCG8	Fecha de nacimiento
DPCG9	Edad o elemento indicativo de edad
DPCG10	Sexo
DPC1	Teléfono
DPAG2	Ciudad
DPAG3	Municipio
DPAG6	Comunidad autónoma
DPAG8	País de nacimiento
DPRCCREO1	Origen étnico
DPRP7	Profesión

Tabla 41 Categoría de datos personales como resultado del análisis previo

Datos médicos	
DDI4	Prótesis
DM2	Servicio prestado
DM3	Tratamiento
DM5	Alergias
DM6	Centros de salud
DM7	Consumo de sustancias tóxicas
DM8	Diagnos
DM9	Enfermedades
DM10	Estado de embarazo
DM11	Estado de enfermedad
DM12	Historial clínico
DM13	Información psicológica
DM14	Información psiquiátrica
DM15	Información recolectada durante un servicio sanitario
DM16	Intervenciones quirúrgicas
DM17	Laboratorios que analizaron las pruebas médicas
DM18	Medidas corporales (talla, peso)
DM19	Prognosis
DM20	Residencias de ancianos
DM21	Vacunas
DM22	Incapacidades médicas

Datos médicos	
DM24	Estado de discapacidad
DM27	Número de historia clínica
DM28	Número de identificación
DM32	Códigos de diagnóstico
DM34	Códigos de tratamiento
DM35	Fecha de admisión
DM36	Fecha de alta
DM37	Fecha de deceso
DM38	Fecha de recogida de muestras
DM39	Fecha de tratamiento
DM40	Fecha de visita
DM41	Número de registros de ensayos clínicos
DM42	Salud física (otra información física no recogida en anteriores campos)
DM43	Salud mental (otra información mental no recogida en anteriores campos)
DM44	Términos médicos que identifican patologías no habituales

Tabla 42 Categoría de datos médicos como resultado del análisis previo

Posteriormente, se llevó a cabo una exploración preliminar de las relaciones entre los atributos a través del análisis de la matriz de correlación usando el coeficiente de Pearson. Se considera que existe correlación entre las variables cuando los valores del coeficiente de Pearson superiores a 0,30 y los determinantes de matrices cercanos a 0 (Hair et al., 1998). En el caso de este trabajo, el análisis de la matriz de correlación resultó en variables correlacionadas al arrojar valores superiores a 0,30 del coeficiente de Pearson y determinantes con valores positivos y cercanos a 0 ($4,558E-8$ para la categoría de datos personales y $2,526E-21$ para la categoría de datos médicos) por lo que se rechaza la hipótesis nula y se concluye que hay relación lineal entre las variables. Por lo tanto, se confirma la aplicabilidad del análisis factorial.

Por último, se recogen los datos médicos que los encuestados aportaron como respuesta a la pregunta “Si cree que hay algún atributo clínico que sería importante incluir y que no aparece aquí, por favor, indíquelo a continuación” con los siguientes atributos médicos que consideraban que sería de utilidad:

- Información de autopsias.
- Uso de mallas.
- Número de hijos vivos y sanos.
- Dependiente para las actividades básicas de la vida diaria o dependiente.
- Historia de maltrato.
- Historia de abuso sexual.
- Viajes recientes.

Estos atributos médicos, en la normativa analizada, no se considera que deban ser protegidos por lo que no han sido incorporados al modelo de ámbito nacional al no tratarse de datos personales.

Análisis de aplicabilidad

Se ha realizado una primera revisión del conjunto de atributos en la fase “Ajuste del listado al dominio de aplicación” para seleccionar los atributos que son de aplicación en España. Además, en el cuestionario y para realizar el análisis sobre la aplicabilidad del modelo de atributos clínicos presentado a los encuestados tras la realización de la segunda parte del cuestionario, se incluyó una tercera parte en el mismo con la

siguiente pregunta: ¿Cree que es importante para su trabajo de investigador y/o docente utilizar el sistema que mejor asegure la confidencialidad de los datos de sus pacientes? En la Figura 47 se observa que un 83% de los encuestados han valorado como importante o muy importante la utilización de un modelo de este tipo que garantice la confidencialidad de los datos médicos en su trabajo diario (Media=4.34, Mediana=5, Moda=5, Desviación típica=0.80) y solo un 1% (corresponde a 1 encuestado) ha considerado que es poco o nada importante el desarrollo de un sistema que ayude a proteger la confidencialidad de los datos médicos.

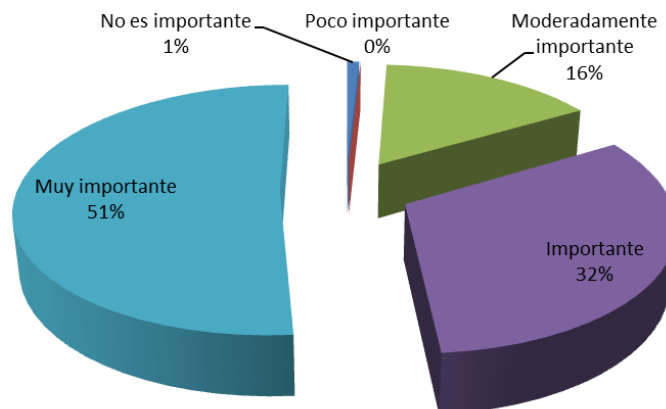


Figura 47 Distribución de las respuestas a la pregunta sobre la importancia de un modelo aplicable de protección de datos médicos

Análisis de fiabilidad

Se ha realizado el análisis de fiabilidad con el propósito de minimizar el error de medida y poder obtener como resultado la validación de un modelo fiable. Este análisis consiste en agrupar las variables (atributos

médicos) para posteriormente analizar el coeficiente α de Cronbach y las correlaciones ítem-total para cada variable. Por tanto, el propósito de este análisis consiste en maximizar la fiabilidad minimizando el error de medida para obtener como resultado un modelo fiable.

Convencionalmente, el mínimo utilizado para el coeficiente α de Cronbach como aceptable está establecido como 0,7, siendo considerados los valores por encima de 0,8 como buena fiabilidad, y 0,9 como excelente (Tabachnick & Fidell, 2006), y que el valor mínimo para la correlación ítem-total sea superior a 0,30 (Kline, 2000; Rocha et al., 2011). En el caso de que alguna variable no supere los valores mínimos, se suprimirán porque disminuyen la consistencia de la escala (α de Cronbach) o tienen una relación escasa con el concepto que se quiere medir (ítem-total); mejorando la fiabilidad de la escala medida.

El análisis de fiabilidad se ha realizado de forma particular para cada categoría presentada en la Tabla 41 (datos personales) y en la Tabla 42 (datos médicos) mediante un proceso iterativo en el que se han eliminado aquellas variables que no cumplen con los valores mínimos de los coeficientes previo análisis de los mismos por considerarse poco relacionados con el resto. En la Tabla 43 se muestra el resultado del análisis de fiabilidad realizado.

Categoría	Atributo		Correlación Ítem-Total	α Cronbach
Datos Personales	DPCG1	Nombre	0,539	0,880
	DPCG2	Primer apellido	0,536	
	DPCG3	Segundo apellido	0,527	
	DPCG6	Género	0,434	

Categoría	Atributo		Correlación Ítem-Total	α Cronbach
	DPCG7	Nacionalidad	0,567	
	DPCG8	Fecha de nacimiento	0,532	
	DPCG9	Edad o elemento indicativo de edad	0,483	
	DPCG10	Sexo	0,362	
	DPC1	Teléfono	0,485	
	DPAG2	Ciudad	0,686	
	DPAG3	Municipio	0,681	
	DPAG6	Comunidad autónoma	0,724	
	DPAG8	País de nacimiento	0,577	
	DPRCCREO1	Origen étnico	0,488	
	DPRP7	Profesión	0,461	
Datos Médicos	DDI4	Prótesis	0,610	0,964
	DM2	Servicio prestado	0,371	
	DM3	Tratamiento	0,733	
	DM5	Alergias	0,691	
	DM6	Centros de salud	0,512	
	DM7	Consumo de sustancias tóxicas	0,745	
	DM8	Diagnos	0,708	
	DM9	Enfermedades	0,641	
	DM10	Estado de embarazo	0,752	
	DM11	Estado de enfermedad	0,830	
	DM12	Historial clínico	0,804	
	DM13	Información psicológica	0,599	
	DM14	Información psiquiátrica	0,666	
	DM15	Información recolectada durante un servicio sanitario	0,689	
	DM16	Intervenciones quirúrgicas	0,778	
	DM17	Laboratorios que analizaron las pruebas médicas	0,414	
	DM18	Medidas corporales (talla, peso)	0,584	
	DM19	Prognosis	0,688	
	DM20	Residencias de ancianos	0,514	
	DM21	Vacunas	0,623	
	DM22	Incapacidades médicas	0,683	
	DM24	Estado de discapacidad	0,655	
	DM27	Número de historia clínica	0,553	
	DM28	Número de identificación	0,562	
	DM32	Códigos de diagnóstico	0,524	
	DM34	Códigos de tratamiento	0,517	
	DM35	Fecha de admisión	0,670	
	DM36	Fecha de alta	0,744	
DM37	Fecha de deceso	0,578		
DM38	Fecha de recogida de muestras	0,727		

Categoría	Atributo		Correlación Ítem-Total	α Cronbach
	DM39	Fecha de tratamiento	0,744	
	DM40	Fecha de visita	0,683	
	DM41	Número de registros de ensayos clínicos	0,533	
	DM42	Salud física (otra información física no recogida en anteriores campos)	0,695	
	DM43	Salud mental (otra información mental no recogida en anteriores campos)	0,690	
	DM44	Términos médicos que identifican patologías no habituales	0,715	
	TOTAL			0,964

Tabla 43 Análisis de fiabilidad

En la Tabla 43, se observa que para todas las variables los coeficientes alfa están comprendidos entre 0,880 y 0,964, demostrando así, una alta consistencia y fiabilidad. Además, para todas las variables la correlación ítem-total se encuentra por encima de 0,30, concluyendo que no hay que eliminar ninguna variable porque se encuentran dentro del rango de aceptación.

Validación predictiva

La validez del constructo es el grado en el que el instrumento (cuestionario) mide aquello para lo que fue diseñado. Para medir dicha validez se utilizó análisis factorial exploratorio que además, permite identificar la estructura subyacente de las relaciones entre las variables, obteniendo así una validación predictiva del modelo. Antes de aplicar dicho análisis es necesario verificar que se cumplen las condiciones para ello. El índice Kaiser-Meyer-Oklin, o KMO, y el test de esfericidad de Bartlett son las medidas utilizadas en este caso. El índice KMO mide la relación entre las variables. El valor mínimo comúnmente aceptado para

dicho índice es 0,5, considerando buenos valores a aquellos entre 0,7 y 0,8, y meritorios los superiores a 0,8 (Punter, Solingen, & Trienekens, 1997). En relación al test de esfericidad de Bartlett el modelo es significativo para p-valores menores de 0,05. Como se observa en la Tabla 44, los valores de KMO se encuentran entre 0,807 y 0,826 considerados como meritorios. En lo que al test de esfericidad de Bartlett se refiere, en todos los casos fue inferior al valor mínimo de referencia. En ese caso, se concluye por tanto que puede aplicarse el análisis multifactorial.

Categoría	KMO	Test de Bartlett		
		Sig	gl	Chi-square
Datos personales	0,807	0,000	105	1962,729
Datos médicos	0,826	0,000	630	5035,253
Total	0,712	0,000	1275	8329,498

Tabla 44 Pruebas de validez para realizar el análisis factorial exploratorio

A continuación, se realiza el análisis factorial exploratorio para cada categoría de atributos médicos que permite predecir las relaciones de influencia entre las variables que pertenecen a diferentes factores.

(a) Análisis factorial exploratorio para los datos personales

La Tabla 45 muestra el resultado final tras aplicar el análisis factorial exploratorio a los atributos recogidos en la categoría “datos personales”. Como la idea era reducir los atributos eliminando duplicidades, se aplicó rotación Varimax con normalización de Kaiser y el criterio de los eigenvalues (Hair et al., 1998) para decidir qué factores o atributos mantener. El valor de corte utilizado para mantener un factor fue 0,32

(Tabachnick & Fidell, 2006). Además, para mejorar la lectura de la tabla, las cargas inferiores al valor de corte se han eliminado. El método de rotación convergió en 8 iteraciones.

Además, en la Tabla 45 se muestran las variables estadísticas junto con el atributo al que representan y las cargas factoriales para cada uno de los 4 factores obtenidos. Estos factores han sido etiquetados atendiendo a la relación semántica de los atributos en cada uno de ellos contenidos según el valor de corte establecido 0,32 (Tabachnick & Fidell, 2006), del siguiente modo:

- Contacto: Factor que agrupa los atributos relacionados con la comunicación entre el paciente y el personal sanitario.
- Envejecimiento: Factor que recoge los atributos relacionados con las características que influyen en el envejecimiento del paciente.
- Ubicación: Factor que concentra los atributos relacionados con la dirección física del paciente.
- Procedencia: Este factor reúne los atributos que se encuentran relacionados con el origen del paciente.

En la Tabla 45 se observa que hay variables con cargas factoriales por encima de valor de corte 0,32 repartidos en varios factores. Por ejemplo, la variable DPCG7 (Nacionalidad) tiene influencia en los factores de “envejecimiento” y “procedencia”. La variable DPC1 (Teléfono) tiene influencia en los factores “contacto” y “ubicación”. La variable DPAG2 (Ciudad) y DPAG3 (Municipio) poseen influencia en “contacto” y “ubicación”. Las variables DPRCCREO1 (Origen étnico) y DPAG8 (País de nacimiento) tienen influencia en los factores “envejecimiento” y

“procedencia”, aunque además DPAG8 también tiene influencia en el factor “ubicación”. Además, el valor de la varianza total explicada con 4 factores es de 80,30%.

Variable estadística	Atributo	Factor			
		Contacto	Envejecimiento	Ubicación	Procedencia
DPCG1	Nombre	0,942			
DPCG2	Primer apellido	0,949			
DPCG3	Segundo apellido	0,887			
DPCG6	Género		0,759		
DPCG7	Nacionalidad		0,618		0,566
DPCG8	Fecha de nacimiento		0,726		
DPCG9	Edad o elemento indicativo de edad		0,850		
DPCG10	Sexo		0,900		
DPC1	Teléfono	0,673		0,424	
DPAG2	Ciudad	0,383		0,883	
DPAG3	Municipio	0,436		0,869	
DPAG6	Comunidad autónoma			0,889	
DPAG8	País de nacimiento		0,439	0,414	0,621
DPRCCREO1	Origen étnico		0,474		0,724
DPRP7	Profesión				0,816

Tabla 45 Matriz de componente rotado para datos personales

(b) Análisis factorial exploratorio para los datos médicos

La Tabla 46 muestra el resultado final tras aplicar el análisis factorial exploratorio a los atributos recogidos en la categoría “datos médicos”. Como la finalidad era reducir los atributos eliminando duplicidades, se

aplicó rotación Varimax con normalización de Kaiser y el criterio de los eigenvalues (Hair et al., 1998) para decidir qué factores o atributos mantener. El valor de corte utilizado para mantener un factor fue 0,32 (Tabachnick & Fidell, 2006) al igual que se realizó en el análisis factorial exploratorio para los datos personales. Además, para mejorar la lectura de la tabla, las cargas inferiores al valor de corte se han eliminado. El método de rotación convergió en 11 iteraciones.

Además, la Tabla 46 muestran las variables estadísticas junto con el atributo al que representan y las cargas factoriales para cada uno de los 6 factores obtenidos. Los factores se han nombrado, de acuerdo al valor de corte establecido 0,32 (Tabachnick & Fidell, 2006), de la siguiente forma:

- Atención hospitalaria: Factor que agrupa los atributos médicos utilizados en una atención hospitalaria.
- Estado mental: Factor que recoge los atributos médicos relacionados con los estados mentales del paciente.
- Códigos: Factor que concentra los atributos relacionados con los códigos identificadores utilizados en el ámbito sanitario.
- Ensayos clínicos: Factor que agrupa los atributos médicos utilizados en investigación médica.
- Analíticas: Factor que representa el grupo de atributos médicos relacionados con pruebas analíticas.
- Otra información médica: Este factor reúne los atributos médicos que no han quedado agrupados en otros factores.

En la Tabla 46 se puede observar que los factores “ensayos clínicos” y “analíticas” comprenden cada uno, tres cargas factoriales máximas (mayor de 0,7). Esto garantiza la existencia de ambos factores porque se cumple el mínimo número de variables que deben representar un factor (Costello & Osborne, 2005; Kim & Mueller, 1994). Además, el valor de la varianza total explicada con 6 factores es de 75,31%.

Variable estadística	Atributo	Factor					
		Atención hospitalaria	Estado mental	Códigos	Otra información médica	Ensayos clínicos	Analíticas
DDI4	Prótesis		0,446	0,479			
DM2	Servicio prestado		0,371				0,598
DM3	Tratamiento	0,744					
DM5	Alergias	0,520	0,454			0,444	0,373
DM6	Centros de salud		0,678				0,423
DM7	Consumo de sustancias tóxicas	0,458	0,585			0,321	
DM8	Diagnosís	0,806					
DM9	Enfermedades	0,811					
DM10	Estado de embarazo	0,702	0,370				
DM11	Estado de enfermedad	0,830					
DM12	Historial clínico	0,832					
DM13	Información psicológica		0,826				
DM14	Información psiquiátrica		,791				
DM15	Información recolectada durante un servicio sanitario	0,605	0,324		0,524		

Variable estadística	Atributo	Factor					
		Atención hospitalaria	Estado mental	Códigos	Otra información médica	Ensayos clínicos	Análíticas
DM16	Intervenciones quirúrgicas	0,811					
DM17	Laboratorios que analizaron las pruebas médicas						0,836
DM18	Medidas corporales (talla, peso)	0,413	0,402		0,495		
DM19	Prognosis		0,353		0,669		
DM20	Residencias de ancianos				0,677		
DM21	Vacunas		0,352			0,759	
DM22	Incapacidades médicas		0,762	0,463			
DM24	Estado de discapacidad		0,798	0,431			
DM27	Número de historia clínica	0,680					0,381
DM28	Número de identificación	0,426		0,462		0,514	
DM32	Códigos de diagnosis			0,803			
DM34	Códigos de tratamiento			0,847			
DM35	Fecha de admisión	0,482		0,597	0,430		
DM36	Fecha de alta	0,697		0,437			
DM37	Fecha de deceso	0,341		0,418	0,634		
DM38	Fecha de recogida de muestras	0,358		0,407			0,451
DM39	Fecha de tratamiento	0,635		0,474			
DM40	Fecha de visita	0,599		0,485		0,378	
DM41	Número de registros de ensayos clínicos			0,328	0,347	0,544	

Variable estadística	Atributo	Factor					
		Atención hospitalaria	Estado mental	Códigos	Otra información médica	Ensayos clínicos	Análíticas
DM42	Salud física (otra información física no recogida en anteriores campos)	0,440	0,458		0,477		
DM43	Salud mental (otra información mental no recogida en anteriores campos)		0,724				
DM44	Términos médicos que identifican patologías no habituales	0,342		0,372	0,551	0,324	

Tabla 46 Matriz de componente rotado para datos médicos

Conclusiones del análisis factorial exploratorio

Como resultado, las categorías iniciales fueron subdivididas a su vez en otras categorías o componentes que a su vez fueron renombradas según los atributos que contenían, dando lugar a la taxonomía de atributos médicos personales para el caso de España que se muestra en la Tabla 47. Así, la taxonomía queda definida por 2 dominios, datos personales y datos médicos, y cada uno de ellos, por 4 y 6 factores o dimensiones respectivamente.

Dominios	Factores
Datos personales	Datos de ubicación
	Datos de envejecimiento
	Datos de contacto
	Datos de procedencia
Datos médicos	Atención hospitalaria
	Estado mental
	Códigos
	Ensayos clínicos
	Análíticas
	Otra información médica

Tabla 47 Taxonomía de datos médicos de carácter personal

Validación confirmatoria

En esta fase de la metodología se confirma el modelo de atributos médicos obtenido. Para ello, se hace uso del modelo de ecuaciones estructurales (SEM), que mediante un conjunto de ecuaciones de regresión múltiple distintas e interrelacionadas, permiten una transición entre el modelo factorial exploratorio y el modelo factorial confirmatorio.

Tal y como se indicó en la sección 2.2.2 del Capítulo 2, las ecuaciones estructurales facilitan la validación nomológica del modelo, ya que permiten controlar el error de medida y desarrollar un test estadístico de bondad de ajuste. Además, SEM es la única técnica que permite explorar simultáneamente una serie de relaciones de dependencia múltiple y cruzada que constituyen un modelo a gran escala, capaz de representar a partir de variables observadas (indicadores) conceptos no observados (factores latentes) y tener en cuenta el error de medida en el proceso de estimación (Hair et al., 1998).

Para realizar este análisis, se han implementado una serie de fases tal y como se describen a continuación.

Especificación

Las ecuaciones estructurales especifican, de manera formal, el modelo teórico presentado. A continuación, se presentan las ecuaciones estructurales para ambas categorías:

(1) Datos personales:

$$\text{Contacto} = \alpha_{1,1}\text{DPCG1} + \alpha_{1,2}\text{DPCG2} + \alpha_{1,3}\text{DPCG3} + \alpha_{1,4}\text{DPC1} + \omega_1$$

$$\text{Envejecimiento} = \alpha_{2,1}\text{DPCG6} + \alpha_{2,2}\text{DPCG7} + \alpha_{2,3}\text{DPCG8} + \alpha_{2,4}\text{DPCG9} + \alpha_{2,5}\text{DPCG10} + \omega_2$$

$$\text{Ubicación} = \alpha_{3,1}\text{DPAG2} + \alpha_{3,2}\text{DPAG3} + \alpha_{3,3}\text{DPAG6} + \omega_3$$

$$\text{Procedencia} = \alpha_{4,1}\text{DPAG8} + \alpha_{4,2}\text{DPRCCREO1} + \alpha_{4,3}\text{DPRP7} + \omega_4$$

Siendo:

- $\alpha_{i,j}$ los coeficientes de regresión que relacionan los factores con las variables observadas
- ω_i los errores de medida para las variables observadas.

(2) Datos médicos:

$$\text{Aten_hosp} = \beta_{1,1}\text{DM3} + \beta_{1,2}\text{DM5} + \beta_{1,3}\text{DM8} + \beta_{1,4}\text{DM9} + \beta_{1,5}\text{DM10} + \beta_{1,6}\text{DM11} + \beta_{1,7}\text{DM12} + \beta_{1,8}\text{DM15} + \beta_{1,9}\text{DM16} + \beta_{1,10}\text{DM27} + \beta_{1,11}\text{DM36} + \beta_{1,12}\text{DM39} + \beta_{1,13}\text{DM40} + \psi_1$$

$$\text{Est_mental} = \beta_{2,1}\text{DM6} + \beta_{2,2}\text{DM7} + \beta_{2,3}\text{DM13} + \beta_{2,4}\text{DM14} + \beta_{2,5}\text{DM22} + \beta_{2,6}\text{DM24} + \beta_{2,7}\text{DM43} + \psi_2$$

$$\text{Códigos} = \beta_{3,1}\text{DDI4} + \beta_{3,2}\text{DM32} + \beta_{3,3}\text{DM34} + \beta_{3,4}\text{DM35} + \psi_3$$

$$\text{Otra_info} = \beta_{4,1}\text{DM18} + \beta_{4,2}\text{DM19} + \beta_{4,3}\text{DM20} + \beta_{4,4}\text{DM37} + \beta_{4,5}\text{DM42} + \beta_{4,6}\text{DM44} + \psi_4$$

$$\text{Ens_clínicos} = \beta_{5,1}\text{DM21} + \beta_{5,2}\text{DM28} + \beta_{5,3}\text{DM41} + \psi_5$$

$$\text{Analíticas} = \beta_{6,1}\text{DM2} + \beta_{6,2}\text{DM17} + \beta_{6,3}\text{DM38} + \psi_6$$

Siendo:

- β_{ij} los coeficientes de regresión que relacionan los factores con las variables observadas.
- ψ_i los errores de medida para las variables observadas.

Identificación

En esta fase, el resultado del cálculo de los grados de libertad queda recogido en la Tabla 44. Dado que los valores de grados de libertad (gl) son mayores que cero, el modelo está sobreidentificado.

Estimación

Para determinar si el modelo teórico inicial (ecuaciones estructurales iniciales) satisface los requisitos de potencia estadística, se evalúan las ecuaciones estructurales iniciales, obteniéndose unos resultados que no se ajustan a los valores de aceptación del modelo.

Evaluación de los índices de bondad de ajuste del modelo

(1) Datos personales:

A continuación, se muestran los valores de bondad de ajuste proporcionado por el programa SAS al ajustar el modelo inicial de ecuaciones estructurales (Tabla 48). Tras el análisis de los valores que se muestran en la Tabla 48, se concluye que el ajuste es pobre debido a que los valores que obtienen los índices NCS, RMSEA, SRMR y GFI; se

encuentran muy alejados de los valores de aceptación. En este sentido, podría suceder que el modelo esté contemplando relaciones poco relevantes, esté sobreparametrizado o haya excluido alguna relación que debería contemplarse. Por esta razón, el modelo debe ser modificado para obtener una nueva versión que ofrezca un ajuste aceptable. Además, la potencia estadística se sitúa en un 69%, muy por debajo del 80% requerido (O'Rourke & Hatcher, 2013).

Índices de bondad de ajuste	Valores sugeridos de corte	Valores de bondad de ajuste del modelo inicial
χ^2	-	340,0756
NCS	$1 > NCS < 5$	4,0485
RMSEA	Próximos a 0,08	0,1601
SRMR	$< 0,08$	0,1038
GFI	$> 0,9$	0,7377
CFI	$> 0,9$	0,8657
NFI	$> 0,9$	0,8309
NNFI	$> 0,9$	0,8321

Tabla 48 Análisis de bondad de ajuste para las ecuaciones estructurales iniciales para la categoría “datos personales”

(2) Datos médicos:

El proceso para obtener el modelo de ecuaciones estructurales ha sido el mismo que en el caso de “datos personales”.

Los indicadores de bondad de ajuste para el modelo inicial proporcionado por el programa SAS al introducir las ecuaciones estructurales se muestran en la Tabla 49. Tras el análisis de los valores que se muestran en la Tabla 49, se concluye que el ajuste es pobre debido a que los valores que obtienen los índices NCS, RMSEA, SRMR, GFI, CFI, NFI y NNFI; se encuentran muy alejados de los valores de

aceptación. Por ello, podría ocurrir que el modelo esté contemplando relaciones poco relevantes, esté sobreparametrizado o haya excluido alguna relación que debería contemplarse. Por esta razón, el modelo debe ser modificado para obtener una nueva versión que ofrezca un ajuste aceptable (O'Rourke & Hatcher, 2013).

Índices de bondad de ajuste	Valores sugeridos de corte	Valores de bondad de ajuste del modelo inicial
χ^2	-	2682,3944
NCS	$1 > NCS < 5$	4,6328
RMSEA	Próximos a 0,08	0,1747
SRMR	$< 0,08$	0,1179
GFI	$> 0,9$	0,4919
CFI	$> 0,9$	0,5805
NFI	$> 0,9$	0,5247
NNFI	$> 0,9$	0,5435

Tabla 49 Análisis de bondad de ajuste para las ecuaciones estructurales iniciales para la categoría “datos médicos”

Modificación del modelo

(1) Datos personales:

Para ayudar en la toma de decisiones, con respecto a la modificación del modelo, se evaluó la prueba del Multiplicador de Lagrange entre factores y variables indicadoras. Los cambios propuestos por estas pruebas no siempre derivan en modificaciones del modelo, ya que es necesario que la decisión esté basada en la comprensión de los atributos y en la investigación existente sobre el tema (Byrne, 1998).

A partir de la información obtenida en la prueba del Multiplicador de Lagrange (Tabla 50) se han ido tomando decisiones para ajustar los índices de bondad de ajuste al valor de aceptación mediante la adición de relaciones entre factores e indicadores. Además, la toma de

decisiones en la modificación del modelo se complementa con un análisis de la matriz de correlación (ANEXO 6) para eliminar las variables que estén muy correlacionadas y así proponer modificaciones que permiten ajustar el modelo de bondad de ajuste siempre sin perder información que haya sido considerada importante por los expertos.

Rank Order of the 10 Largest LM Stat for Paths from Exogenous Variables				
To	From	LM Stat	Pr > ChiSq	Parm Change
DPAG6	Procedencia	28.29080	<.0001	0.30477
DPCG7	Procedencia	20.89909	<.0001	1.13585
DPC1	Ubicacion	20.37240	<.0001	0.52907
DPAG6	Envejecimiento	17.22745	<.0001	0.22326
DPCG3	Ubicacion	14.69157	0.0001	0.26154
DPRP7	Contacto	12.32547	0.0004	0.39181
DPAG6	Contacto	12.29553	0.0005	-0.19417
DPRP7	Envejecimiento	10.01245	0.0016	-0.91545
DPRCCREO1	Ubicacion	8.84506	0.0029	-0.31723
DPAG3	Contacto	7.18289	0.0074	0.11632

Tabla 50 Prueba del Multiplicador de Lagrange para la inclusión de nuevas relaciones entre factores e indicadores para la categoría “Datos personales”

Se han realizado las modificaciones de las ecuaciones estructurales utilizando para la toma de decisiones la prueba del Multiplicador de Lagrange, la matriz de correlación entre variables indicadoras y el significado conceptual de las variables involucradas. Los cambios realizados son los siguientes:

- DPCG1 (nombre) se renombra como “Nombre completo” porque según la Tabla 26, es un atributo compuesto de, al menos, DPCG2 (primer apellido) y DPCG3 (segundo apellido). Además, la matriz de correlación indica una alta correlación entre las variables.
- DPAG2 (ciudad) se renombra como “Localización” porque según la Tabla 26, podría denominarse como un atributo compuesto

que englobe a DPAG3 (municipio) y DPAG6 (comunidad autónoma). Además, la matriz de correlación indica una alta correlación (valores cercanos a 1) entre las variables.

Tras las modificaciones adoptadas, las ecuaciones estructurales modificadas son las siguientes:

$$\text{Contacto} = \alpha_{1,1}\text{DPCG1} + \alpha_{1,2}\text{DPC1} + \alpha_{1,3}\text{DPRP7} + \omega_1$$

$$\text{Envejecimiento} = \alpha_{2,1}\text{DPCG6} + \alpha_{2,2}\text{DPCG7} + \alpha_{2,3}\text{DPCG8} + \alpha_{2,4}\text{DPCG9} + \alpha_{2,5}\text{DPCG10} + \alpha_{2,6}\text{DPRP7} + \omega_2$$

$$\text{Ubicación} = \alpha_{3,1}\text{DPCG10} + \alpha_{3,2}\text{DPAG2} + \alpha_{3,3}\text{DPAG8} + \omega_3$$

$$\text{Procedencia} = \alpha_{4,1}\text{DPCG7} + \alpha_{4,2}\text{DPAG8} + \alpha_{4,3}\text{DPRCCREO1} + \alpha_{4,4}\text{DPRP7} + \omega_4$$

Siendo:

- $\alpha_{i,j}$ los coeficientes de regresión que relacionan los factores con las variables observadas
- ω_i los errores de medida para las variables observadas.

(2) Datos médicos:

Para ayudar en la toma de decisiones con respecto a la modificación del modelo, se evaluó la prueba del Multiplicador de Lagrange entre factores y variables indicadoras. Los cambios propuestos por estas pruebas no siempre derivan en modificaciones del modelo, ya que es necesario que la decisión esté basada en la comprensión de los atributos y en la investigación existente sobre el tema (Byrne, 1998).

A partir de la información obtenida en la prueba del Multiplicador de Lagrange (Tabla 51) se han ido tomando decisiones para ajustar los

índices de bondad de ajuste al valor de aceptación mediante la adición de relaciones entre factores e indicadores. Además, la toma de decisiones en la modificación del modelo se complementa con un análisis de la matriz de correlación (ANEXO 6) para eliminar las variables que estén muy correlacionadas y así proponer modificaciones que permiten ajustar el modelo de bondad de ajuste siempre sin perder información que haya sido considerada importante por los expertos.

Rank Order of the 10 Largest LM Stat for Paths from Exogenous Variables				
To	From	LM Stat	Pr > ChiSq	Parm Change
DM35	Aten_hosp	39.00880	<.0001	0.79757
DM13	Codigos	28.55745	<.0001	-0.56139
DM35	Otra_info	27.57750	<.0001	0.75901
DM7	Analiticas	27.21563	<.0001	0.80835
DM7	Aten_hosp	27.06026	<.0001	0.70035
DM38	Ens_clinicos	21.46936	<.0001	1.49430
DM39	Analiticas	20.85805	<.0001	0.69042
DM43	Otra_info	20.43302	<.0001	0.62663
DDI4	Ens_clinicos	19.55816	<.0001	0.56361
DM40	Ens_clinicos	19.12479	<.0001	0.68347

Tabla 51 Prueba del Multiplicador de Lagrange para la inclusión de nuevas relaciones entre factores e indicadores para la categoría "Datos médicos"

Se han realizado las modificaciones de las ecuaciones estructurales utilizando para la toma de decisiones la prueba del Multiplicador de Lagrange entre factores y variables indicadoras, la matriz de correlación entre variables indicadoras, el significado conceptual de las variables involucradas, el Real Decreto 1093/2010 (para comprobar los datos médicos que podría recoger al menos el historial médico) y la frecuencia de utilización de los datos evaluada en el análisis estadístico; quedan recogidos en la Tabla 52. En este caso no se produce pérdida de información en la eliminación ya que en el proceso se ha tenido en cuenta el significado de los atributos médicos y porque en la historia

clínica de un paciente pueden quedar incluidos como parte de ella como así recoge el Real Decreto 1093/2010.

Variables eliminadas	Motivo		
	Matriz de correlación ¹	Real Decreto 1093/2010 ²	Análisis Descriptivo ³
DM5	X (DM16)	X (DM16)	
DM10	X (DM16)		
DM39	X (DM40)		
DM40	X (DM36)		
DM9	X (DM8)	X (DM8)	
DM27	X (DM12)	X (DM12)	
DM16	X (DM12)	X (DM12)	
DM15	X (DM12)		
DM11	X (DM12)		
DM7	X (DM24)	X (DM24)	
DM24	X (DM22)		X (DM22)
DM22	X (DM14)		X (DM14)
DM43	X (DM14)		X (DM14)
DM6	X (DM13)	X (DM13)	X (DM13)
DM32	X (DM34)		X (DM34)
DM35	X (DM34)		
DM18	X (DM19)		
DM20	X (DM19)		X (DM19)
DM19	X (DM44)		X (DM44)
DM44	X (DM42)		X (DM42)
DM28	X (DM21)		X (DM21)
DM17	X (DM2)	X (DM2)	X (DM2)
¹ Matriz de correlación ² Real Decreto 1093/20103 ³ Variables del análisis descriptivo con valores de media cercanos al valor "poco importante".			

Tabla 52 Modificaciones realizadas para mejorar la bondad de ajuste del modelo

Tras las modificaciones adoptadas, las ecuaciones estructurales modificadas son las siguientes:

$$\text{Aten_hosp} = \beta_{1,1}\text{DM3} + \beta_{1,2}\text{DM8} + \beta_{1,3}\text{DM12} + \beta_{1,4}\text{DM36} + \psi_1$$

$$\text{Est_mental} = \beta_{2,1}\text{DM13} + \beta_{2,2}\text{DM14} + \psi_2$$

$$\text{Códigos} = \beta_{3,1}\text{DDI4} + \beta_{3,2}\text{DM34} + \psi_3$$

$$\text{Otra_info} = \beta_{4,1}\text{DM20} + \beta_{4,2}\text{DM42} + \psi_4$$

$$\text{Ens_clínicos} = \beta_{5,1}\text{DM21} + \beta_{5,2}\text{DM41} + \psi_5$$

$$\text{Analíticas} = \beta_{6,1}\text{DM2} + \beta_{6,2}\text{DM38} + \psi_6$$

Siendo:

- $\beta_{i,j}$ los coeficientes de regresión que relacionan los factores con las variables observadas.
- ψ_i los errores de medida para las variables observadas.

Evaluación de los índices de bondad de ajuste del modelo

(1) Datos personales:

Los indicadores necesarios para realizar el análisis de bondad de ajuste, proporcionados tanto por el programa SAS como por el programa AMOS. La Tabla 53, obtenida con el programa SAS, aglutina todos los indicadores de bondad de ajuste.

Índices de bondad de ajuste	Valores sugeridos de corte	Valores de bondad de ajuste del modelo final
χ^2	-	65,6521
NCS	$1 > \text{NCS} < 5$	1,9908
RMSEA	Próximos a 0,08	0,0912
SRMR	$< 0,08$	0,0481
GFI	$> 0,9$	0,9129
CFI	$> 0,9$	0,9470
NFI	$> 0,9$	0,9021
NNFI	$> 0,9$	0,9116

Tabla 53 Bondad de ajuste del ajuste final del modelo para la categoría "datos personales"

En la Tabla 54, se puede observar que la modificación del modelo ha permitido ajustar los índices de bondad de ajuste a valores de corte

aceptados. Por lo tanto, el modelo final para la categoría “datos personales” queda validado.

Índices de bondad de ajuste	Valores sugeridos de corte	Valores de bondad de ajuste del modelo inicial	Valores de bondad de ajuste del modelo final
χ^2	-	340,0756	65,6521
NCS	$1 > NCS < 5$	4,0485	1,9908
RMSEA	Próximos a 0,08	0,1601	0,0912
SRMR	$< 0,08$	0,1038	0,0481
GFI	$> 0,9$	0,7377	0,9129
CFI	$> 0,9$	0,8657	0,9470
NFI	$> 0,9$	0,8309	0,9021
NNFI	$> 0,9$	0,8321	0,9116

Tabla 54 Comparación de índices de bondad de ajuste para la categoría "datos personales"

(2) Datos médicos:

Los índices necesarios para realizar el análisis de bondad de ajuste, proporcionados tanto por el programa SAS como por el programa AMOS. La Tabla 55, obtenida con el programa SAS, aglutina todos los índices de bondad de ajuste.

Índices de bondad de ajuste	Valores sugeridos de corte	Valores de bondad de ajuste del modelo final
χ^2	-	142,9941
NCS	$1 > NCS < 5$	2,3063
RMSEA	Próximos a 0,08	0,1048
SRMR	$< 0,08$	0,0569
GFI	$> 0,9$	0,8668
CFI	$> 0,9$	0,9103
NFI	$> 0,9$	0,8562
NNFI	$> 0,9$	0,8684

Tabla 55 Bondad de ajuste del modelo final para la categoría "datos médicos"

En la Tabla 56, se puede observar que la modificación del modelo ha permitido ajustar los índices de bondad de ajuste a valores de corte próximos a los de aceptación (GFI, NFI y NNFI que se encuentran próximos a 0,9; por lo que el ajuste es adecuado). Además, el valor del índice del residuo cuadrático medio (SRMR) con valor 0,0569 (menor de 0,08) indica un valor de aceptación del modelo (Hu & Bentler, 1999).

En este caso, el valor del índice RMSEA supera el límite de aceptación, que se encuentra por encima de 0,1. Con el fin de justificar la aceptación del modelo, se puede observar que el límite inferior del intervalo de confianza de RMSEA para el 90% de los valores de la muestra es de 0,0823. Dicho valor se encuentra muy próximo al valor de aceptación del índice.

Índices de bondad de ajuste	Valores sugeridos de corte	Valores de bondad de ajuste del modelo inicial	Valores de bondad de ajuste del modelo final
χ^2	-	2682,3944	142,9941
NCS	$1 > NCS < 5$	4,6328	2,3063
RMSEA	Próximos a 0,08	0,1747	0,1048
SRMR	$< 0,08$	0,1179	0,0569
GFI	$> 0,9$	0,4919	0,8668
CFI	$> 0,9$	0,5805	0,9103
NFI	$> 0,9$	0,5247	0,8562
NNFI	$> 0,9$	0,5435	0,8684

Tabla 56 Comparación de índices de bondad de ajuste para la categoría "datos médicos"

Interpretación de los resultados

Se utiliza el método de máxima verosimilitud robusto con el fin de solucionar los problemas de no normalidad de los datos (Byrne, 1994). Para la estimación y la representación gráfica del modelo en notación

LISREL, se ha utilizado el programa AMOS, obteniendo así, la Figura 48 y la Figura 49 para ambas categorías de datos. Ambas figuras representan la última iteración de ajuste del modelo. En esta iteración se alcanzan unos valores aceptables para los índices de bondad de ajuste. Los cuadrados se asocian a los indicadores correspondientes a los atributos médicos. Los círculos pequeños identifican los errores de muestreo de cada una de las variables discretas. Los círculos grandes representan los factores latentes que actúan como relaciones de cada uno de los dominios de los atributos médicos. En último lugar, los arcos dirigidos tienen asociado un valor de correlación estandarizado (índice de correlación) ayudando a visualizar el grado de actuación entre las diferentes variables presentes en la figura.

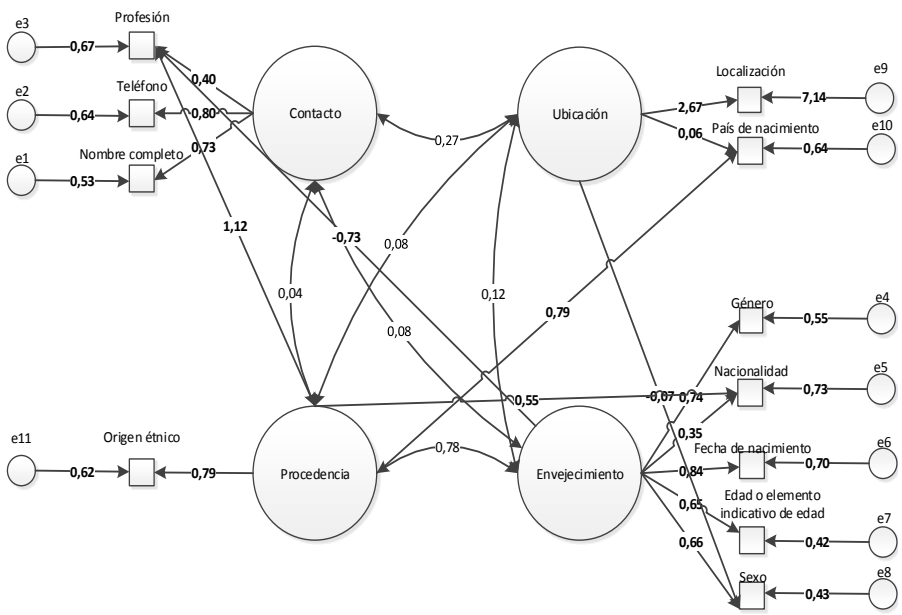


Figura 48 Representación gráfica del modelo obtenido para la categoría "datos personales" en notación LISREL

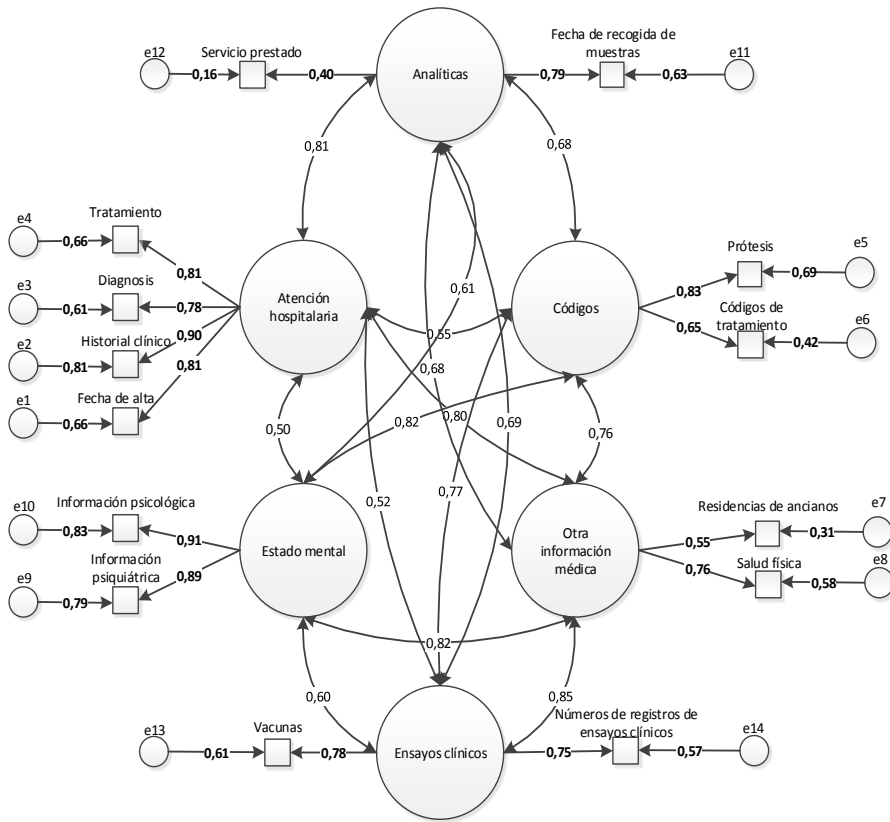


Figura 49 Representación gráfica del modelo obtenido para la categoría "datos médicos" en notación LISREL

Tras la realización de la metodología propuesta en el Capítulo 2 se ha obtenido un modelo final de atributos médicos validado. Este modelo ha tenido en cuenta la legislación internacional para extraer el estado de la cuestión y la legislación nacional para aplicar la metodología al caso de España. Durante el proceso de creación del modelo se ha logrado reducir el listado desde 9 dimensiones y 122 atributos médicos, a 2 dimensiones y 25 atributos médicos (83,61% de reducción del número de atributos

médicos). Finalmente, se ha evaluado la aplicabilidad, validez y fiabilidad obteniendo resultados positivos en los tres análisis.

Para aplicar este modelo a un proceso específico de anonimización, lo primero que se deberá hacer es eliminar aquellos atributos que no se encuentren en los datos o no sean necesarios, ya que la eliminación es siempre lo más seguro. Los datos que se encuentren en el fichero, pero no en el modelo, deberán ser cuidadosamente estudiados. Lo más probable es que se encuentren con otro nombre o un significado similar. El modelo servirá para saber qué atributos son los considerados más importantes y, al haber eliminado redundancias, supone un conjunto de atributos con mayores garantías de protección de los datos personales. Normalmente las exportaciones de datos médicos son llevadas a cabo por el personal del departamento de las TIC, por lo que esta información es de gran utilidad para saber qué datos es importante mantener, y cuáles deben ser anonimizados en la extracción de información médica. El disponer de las relaciones entre atributos y los pesos de cada uno, permite, además que el conjunto de atributos sea reducido aún más según las necesidades del proceso de anonimización particular a llevar a cabo, permitiendo así disminuir el riesgo de re-identificación en las EIPD.

Otra utilidad aún más importante del modelo es su uso como modelo de datos médicos en los desarrollos de aplicaciones en el ámbito de la salud y el bienestar. Al definir el modelo de datos a utilizar, el modelo aquí obtenido puede utilizarse para conocer cuál es la información de carácter personal considerada más importante para el cuidado de la salud en los ámbitos de cuidado de pacientes, investigación y docencia. Los datos personales en este modelo contenidos han sido obtenidos utilizando

como referencia el marco teórico internacional, pero además han sido validados empíricamente por expertos y de esta información se ha obtenido, tras el tratamiento estadístico de los datos, la clasificación de los mismos o taxonomía, y también el nivel de importancia y las relaciones entre ellos. Por tanto, permitirá crear un modelo de datos basado en estándares y normativa internacional y adaptado al marco español a través de la consulta a expertos de diferentes ámbitos, tratándose entonces de un modelo de validez global, en lugar de local. Por tanto, esto supone un importante marco para la definición del modelo de datos en el área, contribuyendo así a la privacidad desde el diseño.

4. CONCLUSIONES Y CONTRIBUCIONES

CONCLUSIONES Y CONTRIBUCIONES

4.1 Conclusiones

Este trabajo forma parte del Doctorado en Tecnologías de la Información aplicadas y, como tal, se trata de un trabajo de seguridad Informática, más específicamente de privacidad que hace uso del Derecho TIC y la Ingeniería del Software, y que se ha aplicado al área de la salud informática. Se propone en este trabajo un proceso sistemático utilizado para la obtención de un modelo de datos personales médicos, utilizando para ello técnicas propias de la Ingeniería del Software. Por último, para validar el modelo se utilizó estadística multivariable.

Los modelos cualitativos, en el ámbito de la protección de datos médicos, son instrumentos de gran importancia en la protección de datos de carácter personal. La caracterización, diseño, adaptación y desarrollo de un modelo cualitativo (modelo que simula el comportamiento de un proceso de anonimización de atributos clínicos) es un arduo trabajo dada la gran cantidad de factores involucrados en el mismo, tales como, el número de legislaciones existentes, zonas geográficas, tipos de pacientes, tecnologías utilizadas o umbrales biométricos. Para ello, se realizó un análisis exhaustivo de la normativa nacional, comunitaria e internacional sobre protección de datos tras lo que se concluye que existen países que incluyen en su ley de protección de datos aquellos datos personales relativos al campo de la salud como, por ejemplo, EE.UU. a través de la HIPAA que lleva ya aplicándose 21 años con éxito. Este modelo presentado por HIPAA no permite ser aplicado en otros

países porque se desconoce la metodología aplicada para replicarla en otras legislaciones y, esto conlleva a que no haya modelos validados en otras jurisdicciones. En España y en Europa, la legislación vigente no llega a este nivel de detalle, y tampoco es posible aplicar la metodología que se utilizó para desarrollar HIPAA, por lo que no es posible realizar una propuesta de protección de datos médicos validada para España. En España, ahora que la LOPD está siendo revisada para adaptarse al RGPD, se debería revisar en este punto e incluir el modelo validado en este trabajo como ampliación de la definición de datos personales relacionados con la salud. Y a nivel internacional, este trabajo aporta la metodología a ser aplicada para obtener un modelo validado adaptado a la normativa en protección de datos de cualquier país del mundo para ayudar a proteger los datos relacionados con la salud. Además, tal y como ha quedado indicado, la protección de los datos relacionada con la salud es una necesidad de demanda creciente con el desarrollo de las nuevas tecnologías, en concreto, del IoT. Por otro lado, como se ha puesto de manifiesto en este trabajo a través de la encuesta realizada queda patente que los profesionales implicados en el uso de estos datos consideran que es necesario disponer de un modelo así.

En la Tabla 57 se recogen los atributos recogidos en la ley HIPAA y los atributos del modelo validado sin categorías para facilitar la comparación. Se puede observar que en el caso de España las fechas no quedan agrupadas en un atributo compuesto. En cambio, los atributos relacionados con la localización física se agrupan en Estados Unidos al igual que en España bajo un mismo atributo compuesto de localización. También llama la atención la evolución de los sistemas de comunicación

ya que el atributo médico “fax” de la ley HIPAA no se encuentra como un atributo médico a anonimizar en España. Y por último, es reseñable que el modelo para España incluye atributos médicos que aunque por definición semántica no debería identificar a un paciente, si se producen casos puntuales de enfermedades sí que facilita la identificación del usuario. A modo de ejemplo, un tratamiento con el medicamento ZMapp en España podría identificar a las personas con infección de ébola dado el número limitado de los pacientes infectados.

HIPAA	MODELO VALIDADO
Nombres	Nombre completo: nombre, primer apellido, segundo apellido
Dirección, ciudad, condado, distrito, código postal	Localización
Fechas: fecha de nacimiento, fecha de admisión, fecha de alta, fecha de muerte, edad>89, elemento indicativo de edad	Fecha de nacimiento
Teléfono	Teléfono
Fax	Origen étnico
Correo electrónico	País de nacimiento
Número seguridad social	Género
Número de registros médicos	Nacionalidad
Número de los beneficiarios del plan de salud	Profesión
Números de cuenta	Edad o elemento indicativo de edad
Número certificado/licencia	Sexo
Identificadores de vehículo, número de serie, números de matrícula	Servicio prestado
Dispositivos identificadores y números de serie	Tratamiento
URL	Diagnóstico
IP	Historial clínico
Biométricos (huellas dactilares y voz)	Fecha de alta
Imágenes fotográficas de rostro completo y cualquier imagen comparable	Información psicológica
Cualquier número de identificación, característica o código	Información psiquiátrica
	Vacunas
	Fecha de recogida de muestras
	Prótesis

HIPAA	MODELO VALIDADO
	Códigos de tratamiento
	Residencias de ancianos
	Salud física
	Números de registros de ensayos clínicos

Tabla 57 Comparativa del modelo proporcionado por HIPAA y el modelo validado

Otra importante contribución de este trabajo es al campo de la privacidad desde el diseño más conocida como *privacy by design*. Puesto que se definen los datos personales de ámbito sanitario considerados como más importantes por los profesionales que los van a usar con una mínima redundancia de datos, éstos pueden ser utilizados en la definición de la estructura de los ficheros en general, como pueden ser las bases de datos médicas, así como, en el desarrollo de aplicaciones, software y, en general, sistemas de información que hacen uso de las mismas.

Por último, como también se mencionó al inicio de este trabajo, éste supone también una importante contribución a la evaluación y análisis de riesgos porque mediante el proceso de generación de modelos se pueden desarrollar modelos de datos médicos para cada jurisdicción, de tal manera que ayude a reducir el proceso de análisis del conjunto completo de atributos en las EIPD, que tengan que llevarse a cabo para cumplir las exigencias de la legislación de protección de datos.

Antes de responder a la pregunta de investigación definida en el punto 0.2.2, se verifica el cumplimiento de los objetivos específicos definidos junto con el resultado obtenido.

El primer objetivo específico consistía en extraer los datos médicos protegidos a través del estudio de la normativa disponible en protección de datos. El estudio se inició por la normativa nacional y comunitaria, pero debido a la falta de definición de datos médicos, se decidió ampliar el estudio a la normativa aplicable en otros países. Además, también se analizaron trabajos relacionados con la protección de datos, que han ayudado a contemplar los conceptos de cuasi-identificador y de atributo compuesto ayudando a identificar como atributos cuasi-identificables la agrupación de atributos no identificables.

El segundo objetivo específico consistía en la adaptación de una metodología de creación de modelos validada, al dominio de la protección de datos médicos desarrollando para ello un nuevo diseño (adaptable y aplicable a otras áreas de conocimiento) de la metodología más adecuado a las necesidades de este trabajo de investigación. Este objetivo fue satisfecho, a partir de la adaptación de la metodología DuMoD definiendo un nuevo proceso basado en éste utilizando para ello técnicas de Ingeniería del Software.

El tercer objetivo específico consistía en definir una taxonomía de datos médicos con las relaciones y pesos entre ellos a partir del análisis factorial exploratorio y de aplicación en España. Se debe destacar el valor de la taxonomía propuesta de atributos clínicos como pilar base para la creación de futuros modelos más robustos y amplios no sólo para la protección de datos clínicos, sino para cualquier otra área relacionada con las tecnologías de uso masivo de datos. Este objetivo se cumplió a través de la implementación práctica de la metodología en el Capítulo 3.

El cuarto objetivo específico que consistía en la validación de la metodología propuesta a través de la validación empírica del modelo de datos médicos aplicado al ámbito de España. Esto queda descrito en el Capítulo 3 mediante la aplicación de diferentes técnicas multivariantes descritas en el Capítulo 2 de metodología.

Se puede concluir que todos los objetivos específicos descritos en el Capítulo 0, con el propósito de cumplir el objetivo definido en el punto 0.2.2, se han desarrollado de manera satisfactoria, ya que se han obtenido unos resultados favorables. Por ello, queda validada la viabilidad de desarrollar un modelo de atributos clínicos a anonimizar que permita cumplir las exigencias de la normativa sobre protección de datos.

Por otro lado, se reseña que el cuestionario no sólo ha sido respondido por personal nacional sino que por ejemplo, había encuestados y estudiantes de MIR que provenían de países centroamericanos y que en su respuesta indicaban la alta importancia que tenía el poder disponer de un modelo de atributos que garantice la correcta anonimización de datos y por ende la privacidad del paciente. Esto creó la necesidad de seguir aplicando el proceso en otros países para ayudar en la protección de datos de los pacientes.

En relación al modelo final confirmado se debe indicar que hay relaciones entre factores y atributos que pueden llamar la atención y que *a priori* no se habían detectado por su significado o la falta de especificación del atributo médico. Por ejemplo, el atributo médico

“Vacunas” ha resultado estar relacionado con el factor “Ensayos clínicos” por lo que el alcance del atributo incluye las vacunas experimentales. Con respecto al atributo “Servicio prestado”, en un primer momento de la investigación se creía que estaba relacionado con el historial médico. En cambio, se ha confirmado que está relacionado con los servicios de “Analíticas”. Otro atributo ha sido “Profesión” al encontrarse relacionado con el factor “Envejecimiento”, con lo que se puede inferir una relación entre el trabajo de un paciente y su envejecimiento. Esto ocurre también con el atributo “Nacionalidad” y “Envejecimiento”.

Finalmente, el desarrollo de esta investigación ha aportado una metodología que permite desarrollar un modelo de atributos de cada legislación de protección de datos a aplicar. Y por ello, se ha respondido a la cuestión abierta planteada y no resuelta en (Avancha et al., 2012): *“How do we design a system to easily accommodate these regional differences?”*, ya que se ha desarrollado un proceso de generación de modelos de datos médicos, validado mediante la aplicación al caso práctico en el ámbito sanitario de España.

4.2 Contribuciones

Tras el estudio, análisis y desarrollo de la tesis se han obtenido las siguientes contribuciones:

PRIMERA

Análisis de la normativa nacional e internacional relacionada con protección de datos médicos, obteniendo así un catálogo para el desarrollo del modelo de datos médicos que se publicó en una revista en enero del año pasado:

- Lozoya de Diego, Abel; Villalba de Benito, María Teresa; Arias Pou, María 2016. “Análisis sobre la heterogeneidad en la legislación de protección de datos personales de carácter médico”. *Diario La Ley*, ISSN 1989-6913, v. 8688, 2016 (Lozoya de Diego, Villalba de Benito, & Arias Pou, 2016a).

SEGUNDA

Descripción de un proceso para la obtención de datos médicos. Dicha contribución fue publicada en las actas de un congreso en julio de 2016:

- Lozoya de Diego, Abel; Villalba de Benito, María Teresa; Arias Pou, María 2016. “Seudonimización de datos en el entorno sanitario”. *Building a European Digital Space, 12th International Conference on Internet, Law & Politics*. HUYGENS Editorial, ISBN: 978-84-608-9670-8, Barcelona (Lozoya de Diego, Villalba de Benito, & Arias Pou, 2016b).

TERCERA

Modelo de datos médicos propuesto para el ámbito español validado empíricamente utilizando la opinión de expertos en el área y realizando el tratamiento estadístico de los datos recogidos. Este modelo ha sido publicado en marzo de 2017 en una revista indexada JCR impact factor

(2016): 1.063 (Q3, Information Science). SJR (2016): Q1 (Information Sciences):

- Lozoya de Diego, Abel; Villalba de Benito, María Teresa; Arias Pou, María 2017. "Taxonomía de información personal de salud para garantizar la privacidad de los individuos". *El profesional de la información*, marzo de 2017, vol. 26, nº 2. ISSN: 1699-2407. DOI: <https://doi.org/10.3145/epi.2017.mar.16> (Lozoya de Diego, Villalba de Benito, & Arias Pou, 2017).

CUARTA

Este trabajo aporta los datos médicos más importantes a anonimizar, por lo que se puede aplicar en la privacidad desde el diseño y en las evaluaciones de impacto en protección de datos. Para ello, entendemos que este trabajo es un ejemplo de cómo se debe aplicar el principio de privacidad desde el diseño en el proceso de anonimización desde la etapa inicial y durante todo el ciclo de vida del tratamiento de la información que vaya a realizarse. Así este principio deberá tenerse en cuenta en el diseño de la estructura de los historiales médicos y de las bases de datos con el propósito de eliminar aquellos datos médicos que los expertos no han considerado como importantes y aumentar de esta manera la seguridad.

QUINTA

La metodología propuesta y validada permite ser reproducible en procesos similares con el propósito de desarrollar modelos de atributos médicos en otros países para cumplir los requisitos impuestos en las diferentes normativas de protección de datos.

SEXTA

La Tabla 29 aporta los atributos médicos eliminados que pueden ser utilizados como un criterio de protección de datos a la hora de realizar las EIPD, los procesos de anonimización o los análisis de riesgos de reidentificación. Este criterio de protección ha sido validado porque se ha extraído de la legislación de protección de datos analizada.

5. LÍNEAS FUTURAS DE INVESTIGACIÓN

LÍNEAS FUTURAS DE INVESTIGACIÓN

En este capítulo se describen los aspectos de investigación que serían deseables abordar. Estas cuestiones y líneas definen el trabajo futuro de la siguiente forma:

- Aumentar el número de aplicaciones del proceso. Sería de gran utilidad aplicar el proceso desarrollado a un amplio número de procesos de anonimización nacionales en el ámbito sanitario. Así incorporaríamos nuevos *benchmarks* (casos de prueba) a nuestro modelo lo que podría ayudar en la mejora del modelo.
- Aplicar el modelo con una muestra más amplia y segregando por perfiles. Como paso previo a la incorporación de estos atributos como parte de la legislación, sería necesario ampliar la muestra de expertos. Además, con una muestra lo suficientemente amplia, sería posible segregar los datos según el perfil, lo que permitiría comparar las muestras para ver si existen diferencias que permitan reducir aún más los datos según el perfil del usuario experto que los usará. Dado que se aporta como contribución de este trabajo el proceso a seguir para ello, así como, el catálogo inicial de datos, llevar a cabo este trabajo sería sencillo.
- Desarrollo de herramientas de apoyo a la gestión del proceso. Sería de gran utilidad el uso de herramientas que permitieran ganar tiempo en algunas fases del proceso.
- Aplicación del modelo a legislaciones de otros países. El 96,61% de los encuestados recoge la opinión de profesionales que desempeñan su trabajo en España. Por ello, la aplicación del

modelo a otros países permitiría mejorar el proceso de generación de modelos aquí planteado.

- Uso del modelo en áreas específicas donde se manipulan cantidades masivas de datos. La constante evolución de tecnologías como *Big Data* e IoT (*Internet of Things*) hacen que el entorno en el que los datos personales son manipulados sea cada vez menos restringido, lo que conlleva a un peligro de privacidad para el usuario de datos catalogados como de nivel alto de protección. La Unión Europea está tomando medidas al respecto a través del Grupo de Trabajo 29 enviando notificaciones formales los proveedores para que se ajusten a las leyes de protección de datos, como es el caso de Microsoft Windows 10 en Francia (Francia, 2016). Sin embargo, no existen mecanismos que ofrezcan aquellos atributos que deban ser anonimizados bajo estas condiciones, donde que esta cantidad de datos requiera de máquinas muy concretas para poder procesar esta información.
- Cuantificar el tiempo de mejora en los procesos de EIPD y en las evaluaciones del riesgo de un proceso de reidentificación. Los procesos de evaluación de riesgos relacionados con las EIPD y la reidentificación de los usuarios podrían aplicar el modelo a procesos ya finalizados para poder cuantificar el tiempo de mejora.

6. BIBLIOGRAFÍA

BIBLIOGRAFÍA

Abascal, E., & Grande Esteban, I. (2005). *Análisis de encuestas*. Madrid: Editor ESIC Editorial.

Agencia Española de Protección de Datos. (2014). Guía para una Evaluación de Impacto en la de Protección Datos Personales. Recuperado el 7 de abril de 2017, de https://www.agpd.es/portalwebAGPD/canaldocumentacion/publicaciones/common/Guias/Guia_EIPD.pdf

Agencia Española de Protección de Datos. (2016). Orientaciones y garantías en los procedimientos de ANONIMIZACIÓN de datos personales. Recuperado el 10 de abril de 2017, de http://www.agpd.es/portalwebAGPD/canaldocumentacion/publicaciones/common/Guias/2016/Orientaciones_y_garantias_Anonimizacion.pdf

Agencia Española de Protección de Datos, & ISMS Forum Spain. (2017). Código de buenas prácticas en proyectos de protección de datos para proyectos Big Data. Recuperado el 29 de junio de 2017, de https://www.agpd.es/portalwebAGPD/canaldocumentacion/publicaciones/common/Guias/2017/Guia_Big_Data_AEPD-ISMS_Forum.pdf

Aguascalientes. (2006). Ley de Transparencia y Acceso a la Información Pública del Estado de Aguascalientes. *Periódico Oficial del Estado [Decreto No. 169]*, tomo LXIX, 22, de 27 de julio de 2015. Recuperado el 30 de marzo de 2017, de <http://www.poderjudicialags.gob.mx/servicios/legislacion/Ley%20de%20Transparencia.pdf>

Alemania. (1969). Sentencia del Tribunal de Justicia de 12 de noviembre de 1969 (Caso Erich Stauder contra Stadt Ulm - Sozialamt). Recuperado el 10 de abril de 2017, de <http://eur-lex.europa.eu/legal-content/ES/TXT/PDF/?uri=CELEX:61969CJ0029&from=ES>

Alemania. (1970). Sentencia del Tribunal de Justicia de 17 de diciembre de 1970 (Caso Internationale Handelsgesellschaft mbH contra

Einfuhr- und Vorratsstelle für Getreide und Futtermittel). Recuperado el 10 de abril de 2017, de <http://eur-lex.europa.eu/legal-content/ES/TXT/PDF/?uri=CELEX:61970CJ0011&from=ES>

Alemania. (2003). Federal Data Protection Act. *Federal Law Gazette p. 66*, 1, de 14 de enero de 2003. Recuperado el 28 de marzo de 2017, de https://www.gesetze-im-internet.de/englisch_bdsgr/federal_data_protection_act.pdf

Ameller, D., Galster, M., Avgeriou, P., & Franch, X. (2015). A survey on quality attributes in service-based systems. *Software Qual Journal*, 24, 271-299. doi:10.1007/s11219-015-9268-4

Amin, A. N., Keshishian, A., Trocio, J., & Vo, L. (2017). Effectiveness and safety of apixaban, dabigatran, and rivaroxaban compared to warfarin among non-valvular atrial fibrillation patients in the us medicare population. *Journal of the American College of Cardiology*, 69(11), 316. doi:10.1016/S0735-1097(17)33705-1

Andorra. (2004). Llei 15/2003, del 18 de desembre, qualificada de protecció de dades personals. *Butlletí Oficial*, 3, de 24 de enero de 2004. Recuperado el 28 de marzo de 2017, de <https://www.bopa.ad/bopa/documents/016003.pdf>

Angola. (2011). Law nº. 22/11 Data Protection Law. *Diário da República*, série I, 114, de 17 de junio de 2011. Recuperado el 2 de abril de 2017, de http://files.mwe.com/info/pubs/Law_22_11_Data_Privacy_Law.pdf

Arenas Ramiro, M. (2008). La protección de datos personales en los países de la Unión Europea. *Revista Jurídica De Castilla Y León*, 16, 113-168.

Argentina. (2000). Ley 25326/2000 de Protección de los Datos Personales. *Boletín Oficial*, de 2 de noviembre de 2000. Recuperado el 30 de marzo de 2017, de <https://www.boletinoficial.gob.ar/#!DetalleNormativa/239279/null>

- Asociación Española de Usuarios de Telecomunicaciones y de la Sociedad de la Información. (2014). Resumen ejecutivo estudio AUTELSI "Privacy by design". Recuperado el 3 de abril de 2017, de http://www.autelsi.es/cms/index.php?option=com_autelsi&pagina=showpdf.php&pdf=/cms/autel/desayunos_tecnologicos/privacybydesign/Resumen_Ejecutivo_Estudio_PbD_AUTELSI.pdf
- Australia. (1988). Privacy Act Nº 119 of 1988. *Gazette*, S399, de 18 de julio de 2016. Recuperado el 8 de abril de 2017, de <https://www.legislation.gov.au/Details/C2016C00838>
- Austria. (1999). Federal Act concerning the Protection of Personal Data. *Federal Law Gazette*, parte I, 165/1999, de 17 de agosto de 1999. Recuperado el 28 de marzo de 2017, de <http://archiv.dsb.gv.at/site/6274/default.aspx>
- Avancha, S., Baxi, A., & Kotz, D. (2012). Privacy in Mobile Technology for Personal Healthcare. *ACM Computing Surveys, Artículo 3*, de noviembre de 2012, 45(1), 1-54. doi:10.1145/2379776.2379779
- Bailey, S., Ukoumunne, O. C., Shephard, E. A., & Hamilton, W. (2017). Clinical relevance of thrombocytosis in primary care: a prospective cohort study of cancer incidence using English electronic medical records and cancer registry data. *Royal College of General Practitioners*, 67(659), 405-413. doi:10.3399/bjgp17X691109
- Baja California. (2010). Ley de Transparencia y Acceso a la Información Pública para el Estado de Baja California. *Periódico Oficial*, tomo CXVII, 42, de 1 de octubre de 2010. Recuperado el 30 de marzo de 2017, de <http://www.ordenjuridico.gob.mx/Documentos/Estatal/Baja%20California/wo54372.pdf>
- Baja California Sur. (2010). Ley de Transparencia y Acceso a la Información Pública para el Estado de Baja California Sur. *Boletín Oficial [Decreto No. 1838]*, tomo XXXVI, 10, de 12 de marzo de 2010. Recuperado el 30 de marzo de 2017, de http://www.ctainl.org.mx/descargas/leg_estatal/B.C.S.pdf

- Barker, R. (1994). *El modelo entidad-relación CASE*METHOD*. España: Editorial Díaz de Santos, S.A.
- Bélgica. (1992). Act of 8 December 1992 on the protection of privacy in relation to the processing of personal data. *Belgian Official Journal*, de 18 de marzo de 1993. Versión consolidada de 7 de abril de 2014. Recuperado el 30 de marzo de 2017, de https://www.privacycommission.be/sites/privacycommission/files/documents/Privacy_Act_1992.pdf
- Bélgica. (2001). Royal Decree implementing the Act of 8 December 1992 on the protection of privacy in relation to the processing of personal data. *Belgian Official Gazette*, de 13 de marzo de 2001. Recuperado el 29 de marzo de 2017, de http://www.legislationline.org/download/action/download/id/2639/file/Belgium_Royal_Decree_Protection_Privacy_Processing_Personal_Data_2001.pdf
- Benin. (2009). Loi n° 2009-09 portant protection des données à caractère personnel en République du Bénin. *Republique du Benin, Assemblée Nationale*, de 29 de abril de 2009. Recuperado el 28 de marzo de 2017, de <http://www.afapdp.org/wp-content/uploads/2012/01/B%C3%A9nin-LOI-SUR-PROTECTION-DES-DONNEES-A-CARACTERE-PERSONNEL-2009.pdf>
- Bentler, P. M. (1990). Comparative fit indexes in structural models. *Psychological Bulletin*, 107(2), 238-246. doi:10.1037/0033-2909.107.2.238
- Bielorrusia. (2008). Law on Information, Informatization and Protection of Information N° 455-Z. *Law of the Republic of Belarus*, 102-Z, de 1 de abril de 2014. Recuperado el 10 de abril de 2017, de http://www.right2info.org/resources/publications/laws-1/laws_belarus-foi-law
- Billings, J., Dixon, J., Mijanovich, T., & Wennberg, D. (2006). Case finding for patients at risk of readmission to hospital: development of algorithm to identify high risk patients. *BMJ: British Medical Journal*, 333(7563), 327. doi:10.1136/bmj.38870.657917.AE

- Bisquerra Alzina, R. (1989). *Introducción conceptual al análisis multivariable : un enfoque informático con los paquetes SPSS-X, BMDP, LISREL y SPAD*. España: PPU S.A.
- Bollen, K. A. (1989). *Structural Equation with Latent Variables*. John Wiley & Sons, Inc
- Bosnia and Herzegovina. (2006). Law on the Protection of Personal Data. *Official Gazette of Bosnia and Herzegovina [49/06]*, de 4 de julio de 2006. Recuperado el 28 de marzo de 2017, de http://www.azlp.gov.ba/images/PropisiEN/Law_on_Protection_of_Personal_Data.doc
- Bradley, J., Barbier, J., & Handler, D. (2013). Embracing the Internet of Everything to capture your share of \$14.4 trillion. *Cisco Systems, Inc.*, Recuperado el 14 de junio de 2017, de http://www.cisco.com/c/dam/en_us/about/ac79/docs/innov/loE_Economy.pdf
- Buenos Aires. (2006). Ley Nº 1985 de Protección de Datos Personales. *Boletín Oficial de la Ciudad de Buenos Aires*, 2494, de 3 de agosto de 2006. Recuperado el 30 de marzo de 2017, de http://boletinoficial.buenosaires.gob.ar/areas/leg_tecnica/boletinOficial/?c=Boletin&a=descargarBoletin&numero=2494
- Bulgaria. (2002). Law for Protection of Personal Data. *State Gazette.*, de enero de 2002. Recuperado el 28 de marzo de 2017, de https://www.cpdp.bg/download.php?part=rubric_element&aid=1407
- Burkina Faso. (2004). Loi nº 010-2004/AN Portant Protection des Donnees a Caractere Personnel. *Assemblée Nationale*, de 20 de abril de 2004. Recuperado el 7 de abril de 2017, de <http://www.afapdp.org/wp-content/uploads/2012/01/Burkina-Faso-Loi-portant-protection-des-donn%C3%A9es-%C3%A0-caract%C3%A8re-personnel-20042.pdf>

- Byrne, B. M. (1994). *Structural Equation Modeling with EQS and EQS/WINDOWS: Basic Concepts, Applications, and Programming*. California: SAGE Publications, Inc.
- Byrne, B. M. (1998). *Structural Equation Modeling with LISREL, PRELIS, and SIMPLIS: Basic concepts, Applications, and Programming*. New Jersey: Lawrence Erlbaum Associates Inc.
- Cabo Verde. (2001). Law 133/V/2001 Data Protection Act. *Comissão Nacional de Protecção de Dados*, de 22 de enero de 2001. Recuperado el 7 de abril de 2017, de <http://www.cnpd.cv/leis/DATA%20PROTECTION%20Law%20133.pdf>
- Campeche. (2012). Ley de Protección de Datos Personales del Estado de Campeche y sus Municipios. *Periódico Oficial del Estado [Decreto No. 231]*, 5034, de 9 de julio de 2012. Recuperado el 30 de marzo de 2017, de <http://inicio.ifai.org.mx/LeyesEstados/campdp.pdf>
- Canadá. (1985). Privacy Act. *Ministerio de Justicia [R.S.C., 1985, c. P-21]*, de versión de 12 de abril de 2016. Recuperado el 29 de marzo de 2017, de <http://laws-lois.justice.gc.ca/PDF/P-21.pdf>
- Canadá. (2000). Personal Information Protection and Electronic Documents Act. *Ministerio de Justicia*, de 28 de marzo de 2016. Recuperado el 29 de marzo de 2017, de <http://laws-lois.justice.gc.ca/PDF/P-8.6.pdf>
- Castilla y León. (2014). Aprobada la contratación del programa informático para la historia de salud única por 1,7 millones de euros. Recuperado el 5 de abril de 2017, de http://www.comunicacion.jcyl.es/web/jcyl/Comunicacion/es/Plantilla100Detalle/1281372057192/_/1284367769416/Comunicacion
- Chihuahua. (2005). Ley de Transparencia y Acceso a la Información Pública del Estado de Chihuahua. *Periódico Oficial del Estado [Decreto No. 278-05]*, 83, de 15 de octubre de 2005. Recuperado el 30 de marzo de 2017, de

<http://www.congresochoihuahua.gob.mx/biblioteca/leyes/archivosLeyes/115.pdf>

- Chile. (1999). Ley Nº 19628 sobre Protección de la Vida Privada. *Diario Oficial de la República de Chile*, 36451, de 17 de febrero de 2012. Recuperado el 30 de marzo de 2017, de <http://www.leychile.cl/Navegar?idNorma=141599>
- China. (2012). National Standard Guiding Technical Document of the People's Republic of China. *Chinese Standard [GB/Z 28828-2012]*, de 5 de noviembre de 2012. Recuperado el 29 de marzo de 2017, de <http://www.chinesestandard.net/PDF-Excerpt/ShowPDFexcerpt.aspx?ExcerptID=GB/Z%2028828-2012>
- Chipre. (2001). The Processing of Personal Data (Protection of Individuals) Law 138 (i) 2001. *Cyprus Government Gazette*, de 23 de noviembre de 2001. Recuperado el 6 de abril de 2017, de [http://www.dataprotection.gov.cy/dataprotection/dataprotection.nsf/697e70c0046f7759c2256e8c004a0a49/f8e24ef90a27f34fc2256eb4002854e7/\\$FILE/138\(I\)-2001_en.pdf](http://www.dataprotection.gov.cy/dataprotection/dataprotection.nsf/697e70c0046f7759c2256e8c004a0a49/f8e24ef90a27f34fc2256eb4002854e7/$FILE/138(I)-2001_en.pdf)
- Coahuila de Zaragoza. (2014). Ley de Acceso a la Información Pública y Protección de Datos Personales para el Estado de Coahuila de Zaragoza. *Periódico Oficial [Decreto 532]*, tomo CXXI, 68, de 10 de diciembre de 2015. Recuperado el 30 de marzo de 2017, de http://www.coahuilatr transparente.gob.mx/leyes/documentos_leyes/LEY%20DE%20ACCESO%20REFORMA1.doc
- Colima. (2003). Ley de Protección de Datos del Estado de Colima. *Periódico Oficial [Decreto No. 356]*, tomo LXXXVIII, 27, de 21 de junio de 2003. Recuperado el 30 de marzo de 2017, de <http://148.235.70.104/periodico/peri/21062003/sup01/13062101.PDF>
- Colima. (2013). Ley de Transparencia y Acceso a la Información Pública del Estado de Colima. *Periódico Oficial [Decreto No. 178]*, tomo XCVIII, 55, de 26 de octubre de 2013. Recuperado el 30 de marzo de 2017, de

<http://148.235.70.104/periodico/peri/26102013/sup03/33102601.pdf>

Colombia. (2012). Ley Estatutaria 1581 de 2012 por la cual se dictan disposiciones generales para la protección de datos personales. *Diario Oficial*, 48.587, de 18 de octubre de 2012. Recuperado el 30 de marzo 2017, de <http://www.alcaldiabogota.gov.co/sisjur/normas/Norma1.jsp?i=49981>

Colombia. (2013). Decreto 1377 de 2013. *Diario Oficial*, 48834, de 27 de junio de 2013. Recuperado el 30 de marzo de 2017, de <http://www.alcaldiabogota.gov.co/sisjur/normas/Norma1.jsp?i=53646>

Comisión Europea. (1974). Sentencia del Tribunal de Justicia de 14 de mayo de 1974 (Caso J. Nold, Kohlen-und Baustoffgrosshandlung contra Comisión de las Comunidades Europeas). Recuperado el 10 de abril de 2017, de <http://curia.europa.eu/juris/showPdf.jsf?jsessionid=9ea7d2dc30dbf685c1f1a85348609e96802a930d5fb5.e34KaxiLc3qMb40Rch0SaxuLa3z0?text=&docid=88488&pageIndex=0&doclang=ES&mode=req&dir=&occ=first&part=1&cid=917687>

Comisión Europea. (1999). Afirmación de los derechos fundamentales en la Unión Europea. Ha llegado el momento de actuar. Informe del grupo de expertos sobre derechos fundamentales. *Oficina de Publicaciones Oficiales de las Comunidades Europeas, febrero de 1999*. Recuperado el 10 de octubre de 2016, de

Consejo de la Unión Europea. (2000). Carta de los derechos fundamentales de la Unión Europea. *Diario Oficial de las Comunidades Europeas [C364]*, 1, de 18 de diciembre de 2000. Recuperado el 11 de abril de 2017, de <http://eur-lex.europa.eu/legal-content/ES/TXT/?uri=OJ:C:2000:364:TOC>

Corea del Sur. (2011). Personal Information Protection Act. *National Law Information Center (Ministry of the Interior), versión 24 de marzo de*

- 2014, 12504, de 29 de marzo de 2011. Recuperado el 29 de marzo de 2017, de <http://www.law.go.kr/eng/engLsSc.do?menuId=1&query=personal+information+protection+act&x=0&y=0#liBgcolor15>
- Costa de Marfil. (2013). Loi n° 2013-450 du 19 juin 2013 relative à la protection des données à caractère personnel. *Journal Officiel de la République de Cote D'Ivoire*, de 8 de agosto de 2013. Recuperado el 7 de abril de 2017, de http://media.mofo.com/files/PrivacyLibrary/3979/Cote-d-voire-loi_2013_450.pdf
- Costa Rica. (2011). Ley N° 8968 de Protección de la Persona Frente al Tratamiento de sus Datos Personales. *La Gaceta Digital*, 170, de 5 de septiembre de 2011. Recuperado el 30 de marzo de 2017, de http://www.gaceta.go.cr/pub/2011/09/05/comp_05_09_2011.pdf
- Costa Rica. (2013). Reglamento a la Ley de Protección de la Persona frente al Tratamiento de sus Datos Personales. *El Alcance n° 42 a La Gaceta n° 45 [37554-JP]*, de 5 de marzo del 2013. Recuperado el 30 de marzo de 2017, de <http://www.prodhab.go.cr//documentacion/Download.aspx?32>
- Costello, A. B., & Osborne, J. W. (2005). Best practices in exploratory factor analysis: Four recommendations for getting the most from your analysis. *Practical Assessment, Research & Evaluation*, 10(7) Recuperado el 6 de febrero de 2017, de <http://pareonline.net/getvn.asp?v=10&n=7>
- Criado Álvarez, J. J., Repullo Labrador, J. R., & García Millán, Á. (2011). Vigencia de la Ley General de Sanidad tras veinticinco años. *Revista Española De Salud Pública*, 85(5), 437-448. Recuperado el 3 de octubre de 2016, de http://www.scielosp.org/pdf/resp/v85n5/03_colaboracion2.pdf
- Croacia. (2003). The Act on Personal Data Protection. *Croatian Parliament [01-081-03-2253/2]*, de 18 de junio de 2003. Recuperado el 28 de marzo de 2017, de http://www.ceecprivacy.org/doc/law_croatia.pdf

- Dadisho, S. (2005). Identity Theft and the Police Response: The Problem. *The Police Chief*, 72(1), 25-29. Recuperado el 3 octubre de 2016, de http://www.policechiefmagazine.org/magazine/index.cfm?fuseaction=display_arch&article_id=493&issue_id=12005
- Dinamarca. (2000). Act Nº 429 of 31 May 2000 Act on Processing of Personal Data. *Lovtidende. Versión traducida por Danish Data Protection Agency.*, de 2 de junio de 2000. Recuperado el 28 de marzo de 2017, de <http://www.datatilsynet.dk/english/the-act-on-processing-of-personal-data/read-the-act-on-processing-of-personal-data/compiled-version-of-the-act-on-processing-of-personal-data/>
- Dixon, P. (2006). Medical identity theft: The information crime that can kill you. Recuperado el 3 de octubre de 2016, de <https://www.worldprivacyforum.org/2006/05/report-medical-identity-theft-the-information-crime-that-can-kill-you/>
- Dubai. (2007). Data Protection Law. *Dubai International Financial Center [Law nº 1 of 2007]* Recuperado el 1 de abril de 2016, de https://www.difc.ae/files/5814/5448/9177/Data_Protection_Law_DIFC_Law_No._1_of_2007.pdf
- Durango. (2008). Ley de Transparencia y Acceso a la Información Pública del Estado de Durango. *Periódico Oficial [Decreto No. 157]*, tomo CCXIX, 4, de 13 de junio de 2008. Recuperado el 30 de marzo de 2017, de <http://www.durango.gob.mx/file/17653>
- Durango. (2013). Ley de Protección de Datos Personales del Estado de Durango. *Periódico Oficial*, 104bis, de 28 de diciembre de 2014. Recuperado el 1 de abril de 2017, de http://inicio.ifai.org.mx/EstadosyMunicipios/LPDED_Dgo.pdf
- Durango. (2016). Ley de Transparencia y Acceso a la Información Pública del Estado de Durango. *Periódico Oficial*, 4 extr. dec. 553, de 4 de mayo de 2016. Recuperado el 30 de marzo de 2017, de [http://congresodurango.gob.mx/Archivos/legislacion/LEY%20DE%20TRANSPARENCIA%20Y%20ACCESO%20A%20LA%20INFORMACION%20\(NUEVA\).docx](http://congresodurango.gob.mx/Archivos/legislacion/LEY%20DE%20TRANSPARENCIA%20Y%20ACCESO%20A%20LA%20INFORMACION%20(NUEVA).docx)

- Egipto. (2014). Constitution of the Arab Republic of Egypt 2014. *Wipo.int [EG060]*, de 15 de enero de 2014. Recuperado el 28 de marzo de 2017, de <http://www.wipo.int/edocs/lexdocs/laws/en/eg/eg060en.pdf>
- El Emam, K., Jabbouri, S., Sams, S., Drouet, Y., & Power, M. (2006). Evaluating common de-identification heuristics for personal health information. *J Med Internet Res*, 8(4) doi:10.2196/jmir.8.4.e28
- El Emam, K. (2014). The de-identification of personally identifiable information. Recuperado el 28 de enero de 2017, de <http://docplayer.net/8387277-The-de-identification-of-personally-identifiable-information.html>
- Eslovaquia. (2002). Act No. 428/2002 Coll. on Protection of Personal Data, as amended by the Act No. 602/2003 Coll., Act No. 576/2004 Coll. and the Act No. 90/2005 Coll. *National Council of the Slovak Republic* Recuperado el 8 de abril de 2017, de http://ec.europa.eu/justice/data-protection/law/files/implementation/slovakia_428_02_en.pdf
- España. (1978). Constitución Española de 27 de diciembre de 1978. *Boletín Oficial del Estado [BOE-A-1978-31229]*, 311, de 29 de diciembre de 1978. Recuperado el 3 de abril de 2017, de https://www.boe.es/diario_boe/txt.php?id=BOE-A-1978-31229
- España. (1986). Ley 14/1986, de 25 de abril, General de Sanidad. *Boletín Oficial del Estado [BOE-A-2002-22188]*, 102, de 29 de abril de 1986. Recuperado el 9 de abril de 2017, de http://www.boe.es/diario_boe/txt.php?id=BOE-A-1986-10499
- España. (1992). Ley Orgánica 5/1992, de 29 de octubre, de regulación del Tratamiento automatizado de los Datos de carácter personal. *Boletín Oficial del Estado [BOE-A-1992-24189]*, 262, de 31 de octubre de 1992. Recuperado el 5 de abril de 2017, de <https://www.boe.es/buscar/doc.php?id=BOE-A-1992-24189>
- España. (1999). Ley Orgánica 15/1999, de 13 de diciembre, de Protección de Datos de Carácter Personal. *Boletín Oficial del Estado [BOE-A-*

2003-23936], 298, de 14 de diciembre de 1999. Recuperado el 8 de abril de 2017, de https://www.boe.es/diario_boe/txt.php?id=BOE-A-1999-23750

España. (2002). Ley 41/2002, de 14 de noviembre, básica reguladora de la autonomía del paciente y de derechos y obligaciones en materia de información y documentación clínica. *Boletín Oficial del Estado [BOE-A-2002-22188]*, 274, de 15 de noviembre de 2002. Recuperado el 5 de abril de 2017, de http://www.boe.es/diario_boe/txt.php?id=BOE-A-2002-22188

España. (2007). Real Decreto 1720/2007, de 21 de diciembre, por el que se aprueba el Reglamento de desarrollo de la Ley Orgánica 15/1999, de 13 de diciembre, de protección de datos de carácter personal. *Boletín Oficial del Estado [BOE-A-2008-979]*, 17, de 19 de enero de 2008. Recuperado el 9 de abril de 2017, de <https://www.boe.es/buscar/act.php?id=BOE-A-2008-979>

España. (2010). Real Decreto 1093/2010, de 3 de septiembre, por el que se aprueba el conjunto mínimo de datos de los informes clínicos en el Sistema Nacional de Salud. *Boletín Oficial del Estado [BOE-A-2010-14199]*, 225, de 16 de septiembre de 2010. Recuperado el 10 de abril de 2017, de http://www.boe.es/diario_boe/txt.php?id=BOE-A-2010-14199

España. (2017). Constituido el Grupo de Trabajo sobre Derechos Digitales de los Ciudadanos. Recuperado el 25 de junio de 2017, de <http://www.red.es/redes/es/magazin-red/noticias/constituido-el-grupo-de-trabajo-sobre-derechos-digitales-de-los-ciudadanos>

Espino Timón, C. (2017). Análisis predictivo: técnicas y modelos utilizados y aplicaciones del mismo - herramientas Open Source que permiten su uso. *Universitat Oberta De Catalunya*, , 11-12. Recuperado el 28 de junio de 2017, de <http://openaccess.uoc.edu/webapps/o2/bitstream/10609/59565/6/caresptimTFG0117mem%C3%B2ria.pdf>

Estados Unidos. (1996). Health Insurance Portability and Accountability Act of 1996. *U.S. Government Publishing Office [Public Law 104-*

- 191], de 21 de agosto de 1996. Recuperado el 9 de abril de 2017, de <https://www.gpo.gov/fdsys/pkg/PLAW-104publ191/pdf/PLAW-104publ191.pdf>
- Estonia. (2007). Personal Data Protection Act. *Riigi Teataja I 2007-24-127.*, de 15 de febrero de 2007. Recuperado el 2 de abril de 2017, de <https://www.riigiteataja.ee/en/eli/512112013011/consolide>
- Fernández Barbudo, C. (2015). El papel de la smart city en el concepto de la privacidad. Trabajo presentado en la *11th International Conference on Internet, Law & Politics*, Barcelona. 93-112. Recuperado el 28 de enero de 2017, en http://edcp.uoc.edu/proceedings_idp2015.pdf
- Ferrer Guerra, J., Pascual Miguel, F. J., Guerrero Barrón, A., Ríos Manríquez, M., & Pérez Carmona, M. T. (2011). *Evaluación de Servicios de Tecnologías de Información en Instituciones de Educación Superior*. México: McGraw-Hill Open-publishing.
- Filipinas. (2011). An Act Protection Individual Personal Information in Information and Communications Systems in the Government and the Private Sector, Creating for this Purpose a National Privacy Commission, and for other Purposes. *Official Gazette [Republic Act Nº 10173]*, de 25 de julio de 2011. Recuperado el 5 de abril de 2017, de http://www.lawphil.net/statutes/repacts/ra2012/ra_10173_2012.html
- Finlandia. (1999). Personal Data Act (523/1999). *FINLEX Data Bank*, de 1 de junio de 1999. Recuperado el 4 de abril de 2017, de <http://www.finlex.fi/en/laki/kaannokset/1999/en19990523.pdf>
- Fischer, P. E., & Morte Ferrer, R. (2013). Big data: a challenge for data protection. Trabajo presentado en la *Proceedings of the 9th International Conference on Internet, Law & Politics*, Barcelona. 205-222. Recuperado el 28 de enero de 2017, en http://edcp.uoc.edu/proceedings_idp2013.pdf

- Francia. (2004). Loi n° 2004-801 du 6 août 2004 relative à la protection des personnes physiques à l'égard des traitements de données à caractère personnel et modifiant la loi n° 78-17 du 6 janvier 1978 relative à l'informatique, aux fichiers et aux libertés. *Journal Officiel of République Française [JUSX0100026L]*, 182, de 7 de agosto de 2004. Recuperado el 8 de abril de 2017, de <http://www.legifrance.gouv.fr/affichTexte.do?cidTexte=JORFTEXT00000441676&dateTexte=&categorieLien=id>
- Francia. (2016). Windows 10: CNIL publicly serves formal notice to Microsoft Corporation to comply with the French Data Protection Act within three months. Recuperado el 3 de abril de 2017, de <https://www.cnil.fr/en/windows-10-cnil-publicly-serves-formal-notice-microsoft-corporation-comply-french-data-protection>
- Fundación Ramón Areces, & OCDE. (2017). La nueva revolución de la producción: la transformación digital. Recuperado el 25 de junio de 2017, de <http://www.fundacionareces.es/fundacionareces/cargarAplicacionAgendaEventos.do?idTipoEvento=1&fechaInicio=01%2F03%2F2017&identificador=1992&fechaFinalizacion=31%2F03%2F2017&nivelAgenda=2>
- George, D., & Mallery, P. (2016). *IBM SPSS Statistics 23 Step by Step: A Simple Guide and Reference*. Nueva York: Taylor & Francis Group.
- Gerbing, D. W., & Hamilton, J. G. (1996). Viability of exploratory factor analysis as a precursor to confirmatory factor analysis. *Structural Equation Modeling: A Multidisciplinary Journal*, 3(1), 62-72. doi:10.1080/10705519609540030
- Ghana. (2012). Data Protection Act (Act 843). *Data Protection Commission* Recuperado el 7 de abril de 2017, de [https://www.dataprotection.org.gh/sites/default/files/Data%20Protection%20Act%20,%202012%20\(Act%20843\).pdf](https://www.dataprotection.org.gh/sites/default/files/Data%20Protection%20Act%20,%202012%20(Act%20843).pdf)
- Gibraltar. (2004). Data Protection Act 2004. Recuperado el 2 de abril de 2017, de <http://www.gibraltarlaws.gov.gi/articles/2004-01o.pdf>

- González Pedraz, J. (2014). Desafíos que para la privacidad y la protección de datos implica el Big Data. Trabajo presentado en la *VII Jornada Profesional de la RBIC: "Big data y bibliotecas: convertir datos en conocimiento"*, Madrid. Recuperado el 4 de octubre de 2016, en http://www.cervantes.es/imagenes/File/biblioteca/jornadas/jornada_7/judith_gonzalez_ponencia_7_jornad_rbic.pdf
- Grasso, L. (2006). *Encuestas. elementos para su diseño y análisis*. Córdoba: Editorial Brujas.
- Grecia. (1997). Law 2472/1997 on the Protection of Individuals with regard to the Processing of Personal Data. *Hellenic Data Protection Authority* Recuperado el 7 de abril de 2017, de https://www.dpa.gr/pls/portal/docs/PAGE/APDPX/ENGLISH_INDEX/LEGAL%20FRAMEWORK/LAW%202472-97-APRIL010-EN%20_2_.PDF
- Grupo de Trabajo sobre Protección de Datos del Artículo 29. (2014). Dictamen 05/2014 sobre técnicas de anonimización. Recuperado el 4 de abril de 2017, de http://ec.europa.eu/justice/data-protection/article-29/documentation/opinion-recommendation/files/2014/wp216_es.pdf
- Guanajuato. (2012). Ley de Transparencia y Acceso a la Información Pública para el Estado y los Municipios de Guanajuato. *Periódico Oficial del Gobierno del Estado*, 154, tercera parte, de 25 de septiembre de 2012. Recuperado el 1 de abril de 2017, de http://www.congresogto.gob.mx/uploads/ley/pdf/106/LEY_DE_TRANSPARENCIA_Y_ACCESO_A_LA_INFORMACION_PUBLICA_PARA_EL_EDO._ABROGADA.pdf
- Guanajuato. (2016a). Ley de Protección de Datos Personales para el Estado y los Municipios de Guanajuato. *Periódico Oficial [Decreto núm. 266]*, 77, de 13 de mayo de 2016. Recuperado el 31 de marzo de 2017, de http://www.congresogto.gob.mx/uploads/ley/pdf/49/Ley_de_Proteccion_de_DPersonales_pEdoGtoyM_Ref13May2016.pdf

- Guanajuato. (2016b). Ley de Transparencia y Acceso a la Información Pública para el Estado de Guanajuato. *Periódico Oficial [Decreto núm. 88]*, 77, de 13 de mayo de 2016. Recuperado el 31 de marzo de 2017, de <https://transparencia.guanajuato.gob.mx/archivos/Itaip.pdf>
- Guernsey. (2002). The Data Protection (Bailiwick of Guernsey) Law 2001. *Records of the Island of Guernsey*, de 29 de abril de 2002. Recuperado el 29 de marzo de 2017, de <http://www.guernseylegalresources.gg/CHttpHandler.ashx?id=71706&p=0>
- Guerrero. (2010). Ley N° 374 de Transparencia y Acceso a la Información Pública del Estado de Guerrero. *Periódico Oficial del Estado de Guerrero*, 48, de 15 de junio de 2010. Recuperado el 31 de marzo de 2017, de http://www.sitios.scjn.gob.mx/sinaj/sites/default/files/normativas_transparencia_archivos/LEY%20NUMERO%20374%20DE%20TRANSPARENCIA%20Y%20ACCESO%20A%20LA%20INFORMACION%20PUBLICA%20DEL%20ESTADO%20DE%20GUERRERO.pdf
- Hair, J. F., Tatham, R. L., Anderson, R. E., & Black, W. (1998). *Multivariate data analysis*. Harlow: Prentice Hall 5th Ed.
- Hauge, M. V., Stevenson, M. D., Kim Rossmo, D., & Le Comber, S. C. (2016). Tagging Banksy: using geographic profiling to investigate a modern art mystery. *Journal of Spatial Science*, 61(1), 185-190. doi:10.1080/14498596.2016.1138246
- Health Level Seven International. (2017). HL7 standards. Recuperado el 14 de abril de 2017, de <http://www.hl7.org/>
- Hidalgo. (2006). Ley de Transparencia y Acceso a la Información Pública Gubernamental para el Estado de Hidalgo. *Periódico Oficial [Decreto Núm. 217]*, 53, de 29 de diciembre de 2006. Recuperado el 29 de marzo de 2017, de http://intranet.e-hidalgo.gob.mx/NormatecaE/Archivos/LeyTAIPG_29122006.pdf

- Holanda. (1999). Personal Data Protection Act (Wet bescherming persoonsgegevens). *Session 1999-2000*, 92, de 23 de noviembre de 1999. Recuperado el 28 de marzo de 2017, de <http://www.legislationline.org/documents/id/5342>
- Honduras. (2006). Ley de Transparencia y Acceso a la Información Pública. *La Gaceta [Decreto nº 170-2006]*, 31193, de 30 de diciembre de 2006. Recuperado el 30 de marzo de 2017, de [http://www.palermo.edu/cele/pdf/LeyesdeAccesoalaInformacion/HondurasLeydeAccesoalaInformacion\(2006\).pdf](http://www.palermo.edu/cele/pdf/LeyesdeAccesoalaInformacion/HondurasLeydeAccesoalaInformacion(2006).pdf)
- Hong Kong. (1995). Personal Data (Privacy) Ordinance. *Gazette*, chapter 486, E.R. 1 of 2013, de 25 de abril de 2013. Recuperado el 29 de marzo de 2017, de https://www.pcpd.org.hk/tc_chi/files/pdpo.pdf
- Hripcsak, G., Bloomrosen, M., FlatelyBrennan, P., Chute, C. G., Cimino, J., Detmer, D. E., . . . Wilcox, A. B. (2014). Health data use, stewardship, and governance: ongoing gaps and challenges: a report from AMIA's 2012 Health Policy Meeting. *Journal of the American Medical Informatics Association*, 21(2), 204-211. doi:10.1136/amiajnl-2013-002117
- Hu, L., & Bentler, P. M. (1999). Cutoff criteria for fit indexes in covariance structure analysis: Conventional criteria versus new alternatives. *Structural Equation Modeling: A Multidisciplinary Journal*, 6(1), 1-55. doi:10.1080/10705519909540118
- Hungría. (2011). Act CXII of 2011 on the Right of Informational Self-Determination and on Freedom of Information. *Nemzeti Adatvédelmi és Információszabadság Hatóság (NAIH)* Recuperado el 29 de marzo de 2017, de http://www.naih.hu/files/Privacy_Act-CXII-of-2011_EN_201310.pdf
- Identity Theft Resource Center. (2015a). Data Breach Report. Recuperado el 1 de abril de 2017, de http://www.idtheftcenter.org/images/breach/DataBreachReports_2015.pdf

- Identity Theft Resource Center. (2015b). Identity Theft Resource Center Breach Report Hits Record High in 2014. Recuperado el 31 de marzo de 2017, de http://www.idtheftcenter.org/images/breach/DataBreachReports_2015.pdf
- Igualada Menor, A. (2003). La Ley de Protección de Datos y la comunicación científica. *MEDIFAM - Revista De Medicina Familiar Y Comunitaria*, 13(3), 129-132. doi:10.4321/S1131-57682003000300001
- INCIBE, Red.es & Huawei. (2017). Creación de un mundo IoT fiable y gestionado. Recuperado el 27 de junio de 2017, de <http://huaweispain15.com/notas/libro-blanco-ciberseguridad-iot.pdf>
- India. (2000). The Information Technology Act, 2000. *The Gazette of India [No. 21 of 2000]*, 27, de 9 de junio de 2000. Recuperado el 27 de marzo de 2017, de http://www.dot.gov.in/sites/default/files/itbill2000_0.pdf
- Indonesia. (2012). Government Regulation on Electronic System and Transaction Operation. *Indonesian Government [82/2012]*, de 12 de octubre de 2012. Recuperado el 29 de enero de 2017, de http://www.flevin.com/id/lgso/translations/JICA%20Mirror/english/4902_PP_82_2012_e.html
- Institute of Medicine (IOM). (2015). Appendix B: Concepts and Methods for De-identifying Clinical Trial Data. *Sharing Clinical Trial Data: Maximizing Benefits, Minimizing Risk* (The National Academies Press ed., pp. 203-256). Washington, DC: doi:10.17226/18998
- International Organization for Standardization, and International Electrotechnical Commission. (2001). ISO/IEC 9126-1:2001: Software engineering – product quality – Part 1: quality model. *ISO/IEC*, 1
- International Working Group on Data Protection in Telecommunications. (2014). Working Paper on Big Data and Privacy: Privacy principles under pressure in the age of Big Data analytics. Trabajo presentado

- en la *55th Meeting [675.48.12]*, Skopje. Recuperado el 30 de marzo de 2017, en http://www.datenschutz-berlin.de/attachments/1052/WP_Big_Data_final_clean_675.48.12.pdf?1407931243
- Irlanda. (1988). Data Protection Act, 1988. *Irish Statute Book, Government Publications [Number 25 of 1988]*. Recuperado el 7 de abril de 2017, de <http://www.irishstatutebook.ie/1988/en/act/pub/0025/print.html>
- Irlanda. (2003). Data Protection (amendment) Act 2003. *Irish Statute Book, Government Publications [Number 6 of 2003]*. Recuperado el 8 de abril de 2017, de <http://www.irishstatutebook.ie/2003/en/act/pub/0006/print.html>
- Islandia. (2000). Act on the Protection of Privacy as regards the Processing of Personal Data, No. 77/2000. *The Data Protection Authority*, de 10 de mayo de 2000. Recuperado el 29 de marzo de 2017, de <http://www.personuvernd.is/information-in-english/greinar/nr/438>
- Israel. (1981). Protection of Privacy Law, 5741-1981. *World Intellectual Property Organization [WIPO IL084]*, de 11 de marzo de 1981. Recuperado el 9 de abril de 2017, de http://www.wipo.int/wipolex/en/text.jsp?file_id=347462
- Italia. (2003). Decreto Legislativo 30 giugno 2003, n. 196 Codice in materia di protezione dei dati personali. *Gazzetta Ufficiale*, suplemento ordinario nº 123, 174, de 29 de julio de 2003. Recuperado el 28 de marzo de 2017, de <http://www.parlamento.it/parlam/leggi/deleghe/03196dl.htm>
- Jalisco. (2013). Ley de Transparencia y Acceso a la Información Pública del Estado de Jalisco y sus Municipios. *Periodico Oficial [Decreto 24450/LX/13]*, 41, de 8 de agosto de 2013. Recuperado el 31 de marzo de 2017, de <http://periodicooficial.jalisco.gob.mx/sites/periodicooficial.jalisco.gob.mx/files/08-08-13-ii.pdf>

- Japón. (2003). Act on the Protection of Personal Information. *Official Gazette [Act No. 57 of 2003]*, de 1 de abril de 2005. Recuperado el 28 de marzo de 2017, de <http://www.cas.go.jp/jp/seisaku/hourei/data/APPI.pdf>
- Jersey. (2005). Data Protection (Jersey) Law 2005. *Jersey Legal Information Board [15.240]*, de 1 de enero de 2017. Recuperado el 29 de marzo de 2017, de <https://www.jerseylaw.je/laws/revised/PDFs/15.240.pdf>
- Kaiser, H. F. (1974). An index of factorial simplicity. *Psychometrika*, 39(1), 31-36. doi:10.1007/BF02291575
- Kim, J., & Mueller, C. W. (1994). Factor Analysis, Statistical Methods and Practical Issues. (pp. 75-155). Londres: Sage Publications.
- Kline, P. (2000). *A psychometrics primer*. Londres: Free Association Books.
- Lang, J. C., De Sterck, H., & Abrams, D. M. (2016). The statistical mechanics of human weight change. *ArXiv E-Prints*, Recuperado el 26 de junio de 2017, de <http://adsabs.harvard.edu/abs/2016arXiv161000656L>
- Letonia. (2000). Protection of Personal Data of Natural Persons. *Data State Inspectorate* Recuperado el 2 de abril de 2017, de http://www.dvi.gov.lv/en/wp-content/uploads/legal-acts/Personal_Data_Protection_Law.doc
- Lévy Mangin, J., & Varela Mallou, J. (2006). *Modelización con estructuras de covarianzas en Ciencias Sociales: temas esenciales, avanzados y aportaciones especiales*. España: Netbiblo.
- Lituania. (1996). Law on Legal Protection of Personal Data of June 11, 1996 No. I-1374 (new version of February 1, 2008, Law No. X-1444). *Organización Mundial de la Propiedad Intelectual [WIPO LT077]*, de 11 de junio de 1996. Recuperado el 9 de abril de 2017, de http://www.wipo.int/wipolex/en/text.jsp?file_id=202094

- Lozoya de Diego, A., Villalba de Benito, María Teresa, & Arias Pou, M. (2016a). Análisis sobre la heterogeneidad en la legislación de protección de datos personales de carácter médico. *Diario La Ley*, 8688, 1-13.
- Lozoya de Diego, A., Villalba de Benito, María Teresa, & Arias Pou, M. (2016b). Seudonimización de datos en el entorno sanitario. Trabajo presentado en la *Building a European Digital Space, 12th International Conference on Internet, Law & Politics*, Barcelona. 325-342. Recuperado el 6 de febrero de 2017, en https://1drv.ms/b/s!AoKgP3w1Rj9WbQbZUoXD-V_WvUk
- Lozoya de Diego, A., Villalba de Benito, María Teresa, & Arias Pou, M. (2017). Taxonomía de información personal de salud para garantizar la privacidad de los individuos. *El Profesional De La Información*, 26(2), 293-302. JCR impact factor (2016): 1.063 (Q3, Information Science). SJR (2016): Q1 (Information Sciences). doi:<https://doi.org/10.3145/epi.2017.mar.16>
- Lucas Murillo de la Cueva, Pablo. (2008). El derecho a la autodeterminación informativa y la protección de datos personales. *Cuadernos De Derecho*, 20, 43-58. Recuperado el 5 de octubre de 2016, <http://www.euskomedia.org/PDFAnlt/azpilcueta/20/20043058.pdf>
- Luxemburgo. (2002). Loi du 2 août 2002 relative à la protection des personnes à l'égard du traitement des données à caractère personnel. *Memorial: Journal Officiel du Grand-Duché de Luxembourg*, de 13 de agosto de 2002. Recuperado el 30 de marzo de 2017, <http://www.legilux.public.lu/leg/a/archives/2002/0091/a091.pdf#page=2&pagemode=none=zoom=125,0,0>
- Macao. (2005). Lei nº 8/2005 da Protecção de Dados Pessoais. *Boletim Oficial*, série I, 34/2005, de 22 de agosto de 2005. Recuperado el 29 de marzo de 2017, de <http://bo.io.gov.mo/bo/i/2005/34/lei08.asp>
- Macedonia. (2005). Law on Personal Data Protection. *Official Gazette of the Republic of Macedonia [07-378/1]*, 7/2005, de 25 de enero de

2005. Recuperado el 28 de marzo de 2017, de <http://www.ceecprivacy.org/pdf/Law%20on%20Personal%20Data%20Protection.pdf>

Madagascar. (2014). Loi n° 2014–038 Sur la protection des données à caractère personne. *Association Francophone des Autorités de Protection des Données Personnelles*, de 9 enero de 2015. Recuperado el 27 de marzo de 2017, de <http://www.afapdp.org/wp-content/uploads/2015/01/Madagascar-L-2014-038-du-09-01-15-sur-la-protection-des-donn%C3%A9es-%C3%A0-caract%C3%A8re-personnel.pdf>

Malasia. (2010). Personal Data Protection Act 2010. *Laws of Malaysia [Act 709]*, de 10 de junio de 2010. Recuperado el 29 de marzo de 2017, de http://www.kkmm.gov.my/akta_kpkk/Personal_Data_Protection_Act_2010.pdf

Mali. (2013). Loi n° 2013-015 Portant Protection des Donnees a Caractere Personnel en Republique du Mali. *Journal Officiel de la Republique du Mali*, 26, de 28 de junio de 2013. Recuperado el 4 de abril de 2017, de <http://www.sgg-mali.ml/JO/2013/mali-jo-2013-26.pdf>

Malta. (2001). Data Protection Act. *Office of the Information and Data Protection Commissioner, Chapter 440 [ACT XXVI of 2001]* Recuperado el 5 de abril de 2017, de <http://idpc.gov.mt/en/Legislation/CAP%20440.pdf>

Manolov, R., Solanas, A., Leiva, D., Losada, J. L., Però, M., & Guàrdia, J. (2015). *Actividades de aprendizaje autónomo y presencial en Metodología de las Ciencias del comportamiento*. Barcelona: Edicions Universitat Barcelona.

Marimón, S. (1999). *La sanidad en la sociedad de la información: Sistemas y tecnologías de la información para la gestión y la reforma de los servicios de salud*. Madrid: Ediciones Díaz de Santos.

- Marruecos. (2009a). Décret n° 2-09-165 du 25 jourmada I 1430 (21 mai 2009) pris pour l'application de la loi n° 09-08 relative à la protection des personnes physiques à l'égard des traitements des données à caractère personnel. *Bulletin Officiel*, 24 jourmada II 1430, 5744, de 18 de junio de 2009. Recuperado el 3 de abril de 2017, de http://www.cdm.gov.ma/sites/default/files/Decret_n209165.pdf
- Marruecos. (2009b). Loi n° 09-08 relative à la protection des personnes physiques à l'égard du traitement des données à caractère personnel. *Bulletin Officiel*, 5714, de 5 de marzo de 2009. Recuperado el 10 de abril de 2017, de <http://www.droit-afrique.com/images/textes/Maroc/Maroc%20-%20Loi%202009%20traitement%20des%20donnees%20protection%20vie%20personnelle.pdf>
- Martín Martín, Q., & del Rosario de Paz Santana, Yanira. (2007). *Tratamiento estadístico de datos con SPSS*. Madrid: Editorial Paraninfo.
- Mauricio. (2004). Data Protection Act 2004. *Government Gazette of Mauritius [Act No. 13 of 2004]*, de 17 de junio de 2004. Recuperado el 9 de abril de 2017, de <http://unpan1.un.org/intradoc/groups/public/documents/unpan/unpan034963.pdf>
- México. (2008a). Ley de Protección de Datos Personales para el Distrito Federal. *Gaceta Oficial*, 434, de 3 de octubre de 2008. Recuperado el 30 de marzo de 2017, de http://iibi.unam.mx/archivistica/ley_datos_personales_df.pdf
- México. (2008b). Ley de Transparencia y Acceso a la Información Pública del Distrito Federal. *Gaceta Oficial del Distrito Federal*, 28, de 28 de marzo de 2008. Recuperado el 30 de marzo de 2017, de http://www.fimevic.df.gob.mx/documentos/transparencia/ley_local/LTAIPDF.pdf
- México. (2010). Ley Federal de Protección de Datos Personales en Posesión de los Particulares. *Diario Oficial de la Federación*, tomo DCLXXXII, 3, de 5 de julio de 2010. Recuperado el 30 de marzo de

2017, de <http://dof.gob.mx/abrirPDF.php?archivo=05072010-MAT.pdf&anio=2010&repo=repositorio/>

México. (2015). Ley General de Transparencia y Acceso a la Información Pública. *Diario Oficial de la Federación*, de 4 de mayo de 2015. Recuperado el 31 de marzo de 2017, de <http://inicio.ifai.org.mx/MarcoNormativoDocumentos/LGTAIP.pdf>

México. (2016). Decreto por el que se expide la Ley de Transparencia, Acceso a la Información Pública y Rendición de cuentas de la Ciudad de México. *Gaceta Oficial de la Ciudad de México*, 66bis, de 6 de mayo de 2016. Recuperado el 30 de marzo de 2017, de <http://www.infodf.org.mx/documentospdf/Ley%20de%20Transparencia,%20Acceso%20a%20la%20Informaci%C3%B3n%20P%C3%BAblica%20y%20Rendici%C3%B3n%20de%20Cuentas%20de%20la%20Ciudad%20de%20M%C3%A9xico.pdf>

Meyers, L. S., Gamst, G., & Guarino, A. J. (2012). *Applied Multivariate Research: Design and Interpretation*. SAGE Publications

Meystre, S. M., Friedlin, F. J., South, B. R., Shen, S., & Samore, M. H. (2010). Automatic de-identification of textual documents in the electronic health record: a review of recent research. *BMC Medical Research Methodology*, 10(1), 10-70. doi:10.1186/1471-2288-10-70

Michoacán. (2008). Ley de Transparencia y Acceso a la Información Pública del Estado de Michoacán de Ocampo. *Periódico Oficial [Decreto N° 29]*, tomo CXLV, 14, de 7 de noviembre de 2008. Recuperado el 31 de marzo de 2017, de http://www.redipd.org/legislacion/common/legislacion/mexico/por_estados/michoacan_1.pdf

Michoacán. (2016). Ley de Transparencia y Acceso a la Información Pública del Estado de Michoacán de Ocampo. *Periódico Oficial [Decreto N° 144]*, tomo CLXIV, 90, de 18 de mayo de 2016. Recuperado el 31 de marzo de 2017, de http://transparencia.congresomich.gob.mx/media/documentos/trabajo_legislativo/LEY_DE_TRANSPARENCIA_ACCESO_INFORMACION_15-06-16.pdf

- Mónaco. (1993). Act n° 1.165 on the protection of personal data. *Journal de Monaco - Official Gazette of the Principality*, 7.918, de 26 de junio de 2009. Recuperado el 22 de marzo de 2017, de <http://www.legimonaco.mc/305/legismclois.nsf/db3b0488a44ebcf9c12574c7002a8e84/28a1a1d90812e249c125773f003beebb!OpenDocument>
- Montenegro. (2008). Personal Data Protection Law. *Official Journal of Montenegro*, 79/08, 70/09 Recuperado el 8 de abril de 2017, de <http://www.afapdp.org/wp-content/uploads/2012/01/Mont%C3%A9n%C3%A9gro-Personal-Data-Protection-Law-79-08-and-70-09.pdf>
- Morelos. (2003). Ley de Información Pública, Estadística y Protección de Datos Personales del Estado de Morelos. *Periódico Oficial*, 4274, de 27 de agosto de 2003. Recuperado el 31 de marzo de 2017, de <http://periodico.morelos.gob.mx/periodicos/2003/4274.pdf>
- Morte Ferrer, R. (2017). ¿Protección de datos/privacidad en la época del Big Data, IoT, wearables...? Sí, más que nunca. *DILEMATA: Revista Internacional De Éticas Aplicadas*, 9(24), 219-233. Recuperado el 27 de junio de 2017, de <http://www.dilemata.net/revista/index.php/dilemata/article/download/412000108/498>
- National Electrical Manufacturers Association. (2016). Digital Imaging and Communications in Medicine. Recuperado el 14 de abril de 2017, de <http://dicom.nema.org/standard.html>
- Nayarit. (2007). Ley de Transparencia y Acceso a la Información Pública del Estado de Nayarit. *Periódico Oficial*, tomo CLXXXI, 114, de 22 de diciembre de 2007. Recuperado el 31 de marzo de 2017, de <http://www.uv.mx/transparencia/files/2012/11/NayaritAI.pdf>
- Nayarit. (2016). Ley de Transparencia y Acceso a la Información Pública del Estado de Nayarit. *Periódico Oficial*, tomo CXCVIII, 84, de 3 de mayo de 2016. Recuperado el 3 de abril de 2017, de http://www.nayarit.gob.mx/transparenciafiscal/des/1_marco_regulatorio/ley_transparencia.pdf

- Nicaragua. (2012). Ley de protección de datos personales. *La Gaceta Diario Oficial [No. 787]*, 61, de 29 de marzo de 2012. Recuperado el 31 de marzo de 2017, de http://www.pgr.gob.ni/PDF/2012/GACETAS/MARZO/GACETA_61_29-03-2012.pdf
- Nigeria. (2013). Guidelines on Data Protection. *National Information Technology Development Agency*, 4, de septiembre de 2013. Recuperado el 27 de marzo de 2017, de <http://nitdademo.azurewebsites.net/wp-content/uploads/2016/06/Guidelines-on-Data-Protection-Final-Draft-3.5.pdf>
- Noruega. (2000). Act of 14 April 2000 No. 31 relating to the processing of personal data (Personal Data Act). *Datatilsynet* Recuperado el 28 de marzo de 2017, de <https://datatilsynet.no/English/Regulations/Personal-Data-Act/>
- Noruega. (2001). Act of 18 May 2001 No. 24 on Personal Health Data Filing Systems and the Processing of Personal Health Data (Personal Health Data Filing System Act). *Datatilsynet* Recuperado el 28 de marzo de 2017, de https://www.datatilsynet.no/Global/english/Personal_Health_Data_Filing_System_Act_20100907.pdf
- Nueva Zelanda. (1993). Privacy Act 1993 No. 28. *New Zealand Legislation*, de 17 de mayo de 1993. Recuperado el 28 de marzo de 2017, de <http://www.legislation.govt.nz/act/public/1993/0028/latest/DLM296639.html>
- Nuevo León. (2008). Ley de Transparencia y Acceso a la Información del Estado de Nuevo León. *Periódico Oficial [Decreto núm. 256]*, de 19 de julio de 2008. Recuperado el 3 de abril de 2017, de http://transparencia.uanl.mx/secciones/disposiciones_legales/archivos/LEY_DE_TRANSPARENCIA_Y_ACCESO_A_LA_INF_DEL_EDO_DE_NUEVO_LEON_v-17-09-12.pdf
- Nuevo León. (2016). Ley de Transparencia y Acceso a la Información Pública del Estado de Nuevo León. *Periódico Oficial [Decreto núm.*

- 119], 83, de 1 de julio de 2016. Recuperado el 31 de marzo de 2017, de <http://inicio.ifai.org.mx/LeyesEstados/NL.pdf>
- O'Rourke, N., & Hatcher, L. (2013). *A Step-by-Step Approach to Using SAS® for Factor Analysis and Structural Equation Modeling*. Carolina del Norte: SAS Institute Inc. Recuperado el 7 de febrero de 2017, de
- Oaxaca. (2008a). Ley de protección de datos personales del estado de Oaxaca. *Periódico Oficial del Estado [Decreto nº 672]*, de 23 de agosto de 2008. Recuperado el 31 de marzo de 2017, de <http://www.congresoaxaca.gob.mx/legislatura/legislacion/leyes/056.pdf>
- Oaxaca. (2008b). Ley de transparencia y acceso a la información pública para el estado de Oaxaca. *Periódico Oficial del Estado [Decreto nº 221]*, de 15 de marzo de 2008. Recuperado el 31 de marzo de 2017, de <http://www.congresoaxaca.gob.mx/legislatura/legislacion/leyes/061.pdf>
- Oaxaca. (2016). Ley de transparencia y acceso a la información pública para el estado de Oaxaca. *Extra del Periódico Oficial del Estado de Oaxaca*, de 11 de marzo de 2016. Recuperado el 31 de marzo de 2017, de <http://www.infodf.org.mx/documentospdf/Ley%20de%20Transparencia%20de%20Oaxaca.pdf>
- Organización de las Naciones Unidas. (1948). Declaración Universal de Derechos Humanos. Recuperado el 10 de abril de 2017, de <http://www.un.org/es/documents/udhr/>
- Paraguay. (2001). Ley Nº 1682, que reglamenta la información de carácter privado. *Agencia Española de Protección de Datos* Recuperado el 30 de marzo de 2017, de https://www.agpd.es/portalwebAGPD/canaldocumentacion/legislacion/iberoamerica/leyes_nacionales/common/pdfs/C.3-cp--Ley-paraguaya-de-proteccion-de-datos.pdf

- Paraguay. (2002). Ley Nº 1.969 que modifica, amplía y deroga varios artículos de la Ley Nº 1682/2001 "Que reglamenta la información de carácter privado". *Red Iberoamericana de Protección de Datos* Recuperado el 30 de marzo de 2017, de http://www.redipd.org/legislacion/common/legislacion/paraguay/Ley_1969_de_2002.pdf
- Pérez Marzabal, J. M. (2014). Internet de las cosas y protección de los consumidores orientada a procesos máquina a máquina. Trabajo presentado en la *Proceedings of the 10th International Conference on Internet, Law & Politics*, Barcelona. 447-463. Recuperado el 8 de abril de 2017, en http://edcp.uoc.edu/proceedings_idp2014.pdf
- Perú. (2011). Ley de Protección de Datos Personales. *Diario Oficial El Peruano [Nº 29733]*, 11455, de 3 de julio de 2011. Recuperado el 31 de marzo de 2017, de <http://www.redipd.org/legislacion/common/legislacion/peru/Ley-29733.pdf>
- Perú. (2013). Reglamento de la Ley Nº 29733. *El Peruano [Decreto Supremo 003-2013-JUS]*, de 22 de marzo de 2013. Recuperado el 30 de marzo de 2017, de http://www.minjus.gob.pe/wp-content/uploads/2013/04/DS-3-2013-JUS.REGLAMENTO.LPDP_.pdf
- Peter, J. P. (1981). Construct validity: A review of basic issues and marketing practices. *Journal of Marketing Research*, v. 18, n. 2, pp. 133-145. doi:10.2307/3150948
- Pianyk, O. S. (2009). *Digital Imaging and Communications in Medicine (DICOM): A Practical Introduction and Survival Guide*. Boston Massachusetts: Springer Science & Business Media.
- Plasencia Cano, M. (2014). Análisis de contingencias en el uso de las formas orales de hierro y su correlación con efectos adversos. *Farmacéuticos Comunitarios*, 6(1) Recuperado el 26 de junio de 2017, de <http://www.farmaceticoscomunitarios.org/es/journal-article/analisis-contingencias-uso-formas-orales-hierro-su-correlacion-con-efectos-adversos>

- Polonia. (1996). Protection of Personal Data. *Journal of Laws*, 133, de agosto de 1997. Recuperado el 28 de marzo de 2017, de http://www.giodo.gov.pl/plik/id_p/180/j/en/
- Ponemon Institute LLC. (2015). Fifth Annual Benchmark Study on Privacy & Security of Healthcare Data. Recuperado el 8 de abril de 2017, de https://media.scmagazine.com/documents/121/healthcare_privacy_security_be_30019.pdf
- Portugal. (1998). Act on the Protection of Personal Data (transposing into the Portuguese legal system Directive 95/46/EC of the European Parliament and of the Council of 24 October 1995 on the protection of individuals with regard to the processing of personal data and on the free movement of such data). *Diário da República I [Act 67/98]*, série A, 65, de 18 de marzo de 1998. Recuperado el 28 de marzo de 2017, de <http://www.cnpd.pt/english/bin/legislation/Law6798EN.HTM>
- President's Council of Advisors on Science and Technology. (2014). Big data and privacy: a technological perspective. *Executive Office of the President*, de mayo de 2014. Recuperado el 5 de abril de 2017, de <http://en.youscribe.com/catalogue/tous/big-data-and-privacy-a-technological-perspective-white-house-2586707>
- Puebla. (2011). Ley de transparencia y acceso a la información pública del estado de Puebla. *Periódico Oficial*, tomo CDXL, de 31 de diciembre de 2011. Recuperado el 2 de abril de 2017, de <http://www.caip.org.mx/documentos/LeyTransparenciaAccesoInformacionPublicaEstadoPuebla2011.pdf>
- Punter, T., Solingen, R. V., & Trienekens, J. (1997). Software product evaluation – current status and future needs for customers and industry. Trabajo presentado en la *Proceedings of 4th Conference on Evaluation of Information Technology*, 57-66. Recuperado el 8 de abril de 2017, en https://www.researchgate.net/profile/Teade_Punter/publication/2508499_Software_Product_Evaluation_-_Current_status_and_future_needs_for_customers_and_industry/li

nks/Ofcfd507d7668c1ff4000000/Software-Product-Evaluation-Current-status-and-future-needs-for-customers-and-industry.pdf

Qatar. (2005). Regulation Nº. 6 of 2005 QFC Data Protection Regulations. *Qatar Financial Centre* Recuperado el 7 de abril de 2016, de http://www.complinet.com/file_store/pdf/rulebooks/QFCRA_1559.pdf

Querétaro. (2002). Ley estatal de acceso a la información gubernamental en el estado de Querétaro. *Periódico Oficial del Estado*, tomo CXXXV, 44, de 27 de septiembre de 2002. Recuperado el 3 de abril de 2017, de <http://www2.queretaro.gob.mx/disco2/servicios/LaSombradeArtega/2002/20020944-01.pdf>

Quitana Roo. (2008). Ley de Transparencia y Acceso a la Información Pública del Estado de Quintana Roo. *Periódico Oficial*, de 2 de julio de 2008. Recuperado el 29 de marzo de 2017, de <http://www.congresoqroo.gob.mx/leyes/administrativo/ley109/L1320130409252.pdf>

Quitana Roo. (2016). Ley de Transparencia y Acceso a la Información Pública del Estado de Quintana Roo. *Periódico Oficial del Estado*, de 3 de mayo de 2016. Recuperado el 29 de marzo de 2017, de <http://www.congresoqroo.gob.mx/leyes/administrativo/ley153/L1410160503398.pdf>

Real Academia Española. (2014). Diccionario de la lengua española. Recuperado el 5 de abril de 2017, de <http://www.rae.es/>

Recio Gayo, M. (2015). Se busca líder para dirigir la Economía Digital basada en el Big Data. Recuperado el 28 de enero de 2017, de <http://red.computerworld.es/actualidad/se-busca-lider-para-dirigir-la-economia-digital-basada-en-el-big-data>

Reino Unido. (1998). Data protection act of 1998. *The Stationery Office chapter*, de 29 de julio de 1998. Recuperado el 8 de abril de 2017, de

http://www.legislation.gov.uk/ukpga/1998/29/pdfs/ukpga_19980029_en.pdf

República Checa. (2000). Personal Data Protection Act on the Protection of Personal Data and on Amendment to Some Acts. *Official Journal [Act 101]*, de 4 de abril de 2000. Recuperado el 1 de abril de 2017, de http://ec.europa.eu/justice/data-protection/law/files/implementation/czech_republic_act_101_en.pdf

República de Eslovenia. (2004). Personal Data Protection Act (ZVOP-1). *Official Gazette [Nº 001-22-148/04]*, 86, de 5 de agosto de 2004. Recuperado el 31 de marzo de 2017, de http://ec.europa.eu/justice/data-protection/law/files/implementation/personal_data_protection_act_rs_2004.pdf

República Dominicana. (2013). Ley No. 172-13 que tiene por objeto la protección integral de los datos personales asentados en archivos, registros públicos, bancos de datos u otros medios técnicos de tratamiento de datos destinados a dar informes, sean estos públicos o privados. *Gaceta Oficial*, 10737, de 15 de diciembre de 2013. Recuperado el 31 de marzo de 2017, de http://www.poderjudicial.gob.do/documentos/PDF/leyes/LEY_172_13.pdf

Rocha, K. B., Pérez, K., Rodríguez-Sanz, M., Borrell, C., & Obiols, J. E. (2011). Propiedades psicométricas y valores normativos del General Health Questionnaire (GHQ-12) en población general española. *International Journal of Clinical and Health Psychology*, 11(1), 125-139. Recuperado el 8 de abril de 2017, de http://www.aepc.es/ijchp/articulos_pdf/ijchp-374.pdf

Ruiz, M. A., Pardo, A., & San Martín, R. (2010). Modelos de ecuaciones estructurales. *Papeles Del Psicólogo*, V. 31, N.1, Pp. 34-45., 31(1), 34-45. Recuperado el 6 de abril de 2017, de <http://www.papelesdelpsicologo.es/pdf/1794.pdf>

- Rumania. (2001). Law No. 677/2001 of 21st of November 2001 on the Protection of Individuals with Regard to the Processing of Personal Data and the Free Movement of Such Data. *Official Journal of Romania*, part I, 790, de 12 de diciembre de 2001. Recuperado el 28 de marzo de 2017, de http://ec.europa.eu/justice/policies/privacy/docs/implementation/r_o_law_677_2001_en_unofficial.pdf
- Rusia. (2006). Federal Law of 27 July 2006 N 152-FZ on Personal Data. [*ny-720231*], de 27 de julio de 2006. Recuperado el 29 de marzo de 2017, de <http://media.mofo.com/docs/mofoprivacy/Federal%20Law%20of%2027%20July%202006%20N152-FZ%20on%20Personal%20Data%20%20%28English%29.pdf>
- San Luis de Potosí. (2007). Ley de transparencia y acceso a la información pública del estado de San Luis de Potosí. *Periódico Oficial del Estado [Decreto 234]*, de 18 de octubre de 2007. Recuperado el 31 de marzo de 2017, de http://www.redipd.org/legislacion/common/legislacion/mexico/por_estados/san_luis_potosi.pdf
- San Luis de Potosí. (2016). Ley de transparencia y acceso a la información pública del estado de San Luis de Potosí. *Periódico Oficial del Gobierno del Estado [Decreto 0217]*, tomo II, de 9 de mayo de 2016. Recuperado el 31 de marzo de 2017, de <http://www.infodf.org.mx/documentospdf/Ley%20de%20Transparencia%20de%20San%20Luis%20Potosi.pdf>
- Schermelleh-Engel, K., Moosbrugger, H., & Müller, H. (2003). Evaluating the Fit of Structural Equation Models: Tests of Significance and Descriptive Goodness-of-Fit Measures. *Methods of Psychological Research Online*, V. 8, N. 2, Pp. 23-74., 8(2), 23-74. Recuperado el 6 de abril de 2017, de https://www.researchgate.net/profile/Karin_Schermelleh-Engel/publication/251060246_Evaluating_the_Fit_of_Structural_Equation_Models_Tests_of_Significance_and_Descriptive_Goodness-of-Fit_Measures/links/0c96051ee93e8c1279000000.pdf

- Senegal. (2008). Loi n° 2008-12 du 25 janvier 2008 portant sur la Protection des données à caractère personnel. *Journal Officiel du Senegal*, de 25 de enero de 2008. Recuperado el 7 de abril de 2017, de http://data.over-blog-kiwi.com/0/88/55/36/20140310/ob_cffb13_loi-protection.pdf
- Serbia. (2008). Law on Personal Data Protection. *Official Gazette of the Republic of Serbia*, 97/08 Recuperado el 28 de marzo de 2017, de <http://www.refworld.org/pdfid/4b5718f52.pdf>
- Seychelles. (2003). Data Protection Act, 2003 (No. 9 of 2003). *Official Gazette [SYC-2003-L-66655]*, 9, de 29 de diciembre de 2003. Recuperado el 9 de abril de 2017, de <http://www.seylli.org/sc/legislation/act/2003/9>
- Sinaloa. (2008). Ley de Acceso a la Información Pública del Estado de Sinaloa. *Periódico Oficial [Decreto No. 141]*, tomo XCIX, 100, de 20 de agosto de 2008. Recuperado el 31 de marzo de 2017, de http://www.laipsinaloa.gob.mx/po_files/2008/agosto/POE-20-AGO-2008-100-II.pdf
- Singapur. (2012). Personal Data Protection Act 2012. *Government Gazette [No. 26 of 2012]*, 25, de 20 de noviembre de 2012. Recuperado el 29 de marzo de 2017, de <http://statutes.agc.gov.sg/aol/search/display/view.w3p?page=0;query=DocId%3Aea8b8b45-51b8-48cf-83bf-81d01478e50b%20Depth%3A0%20Status%3Ainforce;rec=0;whole=yes>
- Sirera, C. (2015). Nuevos Negocios y Aplicaciones: Big Data y la evolución del Internet de las Cosas. Recuperado el 27 de junio de 2017, de <http://www.a pep.es/nuevos-negocios-y-aplicaciones-big-data-y-la-evolucion-del-internet-de-las-cosas/>
- Sonora. (2005). Ley de Acceso a la Información Pública y de Protección de Datos Personales del Estado de Sonora. *Boletín Oficial [Número 156]*, tomo CLXXV, 16, de 25 de febrero de 2005. Recuperado el 31 de marzo de 2017, de http://www.transparencia.uson.mx/ley_estatal.pdf

- Sonora. (2016). Ley de Transparencia y Acceso a la Información Pública del Estado de Sonora. *Boletín Oficial del Estado de Sonora [Ley nº 90]*, sección II, de 28 de abril de 2016. Recuperado el 31 de marzo de 2017, de <http://inicio.ifai.org.mx/LeyesEstados/son.pdf>
- Sudáfrica. (2002). No. 68 of 2002: Electronic Communications Security. *Government Gazette [Act No. 68, 2002]*, 452(24356), de 6 de febrero de 2003. Recuperado el 1 de abril de 2017, de http://us-cdn.creamermedia.co.za/assets/articles/attachments/00466_a68-02.pdf
- Sudáfrica. (2013). Protection of Personal Information Act, 2013. *Government Gazette [Act No. 4 of 2013]*, 581(37067), de 26 de noviembre de 2013. Recuperado el 1 de abril de 2017, de <http://www.justice.gov.za/legislation/acts/2013-004.pdf>
- Suecia. (1998). Personuppgiftslag (1998:204). *Regeringskansliets rättsdatabaser*, de 29 de marzo de 1998. Recuperado el 28 de marzo de 2017, de <http://rkrattsbaser.gov.se/sfst?bet=1998:204>
- Suiza. (1992). Federal Act on Data Protection (FADP). *SR.235.1 [Bundesgesetz über den Datenschutz (DSG)]*, de 19 de junio de 1992. Recuperado el 1 de abril de 2017, de <http://www.admin.ch/ch/e/rs/2/235.1.en.pdf>
- Supervisor Europeo de Protección de Datos. (2014). Privacy and competitiveness in the age of big data: The interplay between data protection, competition law and consumer protection in the Digital Economy. *Supervisor Europeo de Protección de Datos (EDPS)*, de marzo de 2014. Recuperado el 5 de abril de 2017, de https://secure.edps.europa.eu/EDPSWEB/webdav/shared/Documents/Consultation/Opinions/2014/14-03-26_competition_law_big_data_EN.pdf
- Sweeney, L. (1996). Replacing personally-identifying information in medical records, the Scrub system. Trabajo presentado en la *Proceedings of the AMIA annual fall symposium. American Medical Informatics Association [PMC2233179]*, 333-337. Recuperado el 5 de abril de 2017, en

https://www.ncbi.nlm.nih.gov/pmc/articles/PMC2233179/pdf/proc_amiaafs00002-0370.pdf

- Sweeney, L. (2002). k-anonymity: a model for protecting privacy. *International Journal on Uncertainty, Fuzziness and Knowledge-Based Systems*, de octubre de 2002, 10(5), 557-570. doi:10.1142/S0218488502001648
- Symantec. (2015). Internet Security Threat Report. Recuperado el 14 de abril de 2017, de http://www.symantec.com/security_response/publications/threatreport.jsp
- Tabachnick, B. G., & Fidell, L. S. (2006). *Using Multivariate Statistics*. Estados Unidos de América: Allyn & Bacon, Inc. Needham Heights, 5th Ed.
- Tabasco. (2015). Ley de transparencia y acceso a la información pública del estado de Tabasco. *Periódico Oficial del Estado*, 618, de 15 de diciembre de 2015. Recuperado el 31 de marzo de 2017, de <http://www.ordenjuridico.gob.mx/Documentos/Estatal/Tabasco/w0109907.pdf>
- Tailandia. (1997). Official Information Act. *Government Gazette*, 114, de 10 de septiembre 1997. Recuperado el 29 de marzo de 2017, de [http://www.track.unodc.org/LegalLibrary/LegalResources/Thailand/Laws/Thailand%20Official%20Information%20Act%20\(1997\).pdf](http://www.track.unodc.org/LegalLibrary/LegalResources/Thailand/Laws/Thailand%20Official%20Information%20Act%20(1997).pdf)
- Taiwán. (2010). Personal Information Protection Act. *Privacy Law in Greater China*, de 26 de mayo de 2010. Recuperado el 29 de marzo de 2017, de <http://www.law.hku.hk/cprivacy/archives/150>
- Tamaulipas. (2013). Ley de Información Pública del Estado de Tamaulipas. *Periódico Oficial [Decreto LIX-958]*, tomo CXXIX, 63, de 23 de mayo de 2013. Recuperado el 31 de marzo de 2017, de <http://transparencia.tamaulipas.gob.mx/wp-content/uploads/2014/03/TF-I-LEY-DE-TRANSPARENCIA-Y-ACCESO-A-LA-INFO-PUBLICA-23052013.pdf>

- Tamaulipas. (2016). Ley de Transparencia y Acceso a la Información Pública del Estado de Tamaulipas. *Periódico Oficial del Estado [Decreto LXII-948]*, tomo CXLI, 50, de 27 de abril de 2016. Recuperado el 31 de marzo de 2017, de <http://www.tamaulipas.gob.mx/wp-content/uploads/2016/09/cxli-50-270416F-ANEXO.pdf>
- Thomas, S. M., Mamlin, B., Schadow, G., & McDonald, C. (2002). A successful technique for removing names in pathology reports using an augmented search and replace method. Trabajo presentado en la *Proceedings of the AMIA Symposium. American Medical Informatics Association*, 777-781. Recuperado el 5 de abril de 2017, en <https://www.ncbi.nlm.nih.gov/pmc/articles/PMC2244188/pdf/procamiasymp00001-0818.pdf>
- Tlaxcala. (2012). Ley de Protección de Datos Personales para el estado de Tlaxcala. *Periódico Oficial del Estado*, 31, de 14 de mayo de 2012. Recuperado el 3 de abril de 2017, de http://www.redipd.org/legislacion/common/legislacion/mexico/por_estados/Tlaxcala_ley_proteccion_datos.pdf
- Tribunal Constitucional de España. (2000). Pleno. Sentencia 292/2000, de 30 de noviembre de 2000. Recurso de inconstitucionalidad 1.463/2000. Promovido por el Defensor del Pueblo respecto de los arts. 21.1 y 24.1 y 2 de la Ley Orgánica 15/1999, de 13 de diciembre, de Protección de Datos de Carácter Personal. Vulneración del derecho fundamental a la protección de datos personales. Nulidad parcial de varios preceptos de la Ley Orgánica. *Boletín Oficial del Estado [BOE-T-2001-332]*, 4, de 4 de enero de 2001. Recuperado el 10 de abril de 2017, de <https://www.boe.es/buscar/doc.php?id=BOE-T-2001-332>
- Tribunal Europeo de Derechos Humanos. (1950). *Convenio Europeo de Derechos Humanos*. Estrasburgo: Council of Europe. Recuperado el 30 de marzo de 2017, de http://www.echr.coe.int/Documents/Convention_SPA.pdf

- Tribunal Europeo de Derechos Humanos. (1978). Case of Klass and Others v. Germany. Recuperado el 2 de abril de 2017, de <http://hudoc.echr.coe.int/eng?i=001-57510>
- Tribunal Europeo de Derechos Humanos. (1984). Case of Malone v. The United Kingdom. Recuperado el 1 de abril de 2017, de <http://hudoc.echr.coe.int/eng?i=001-57533>
- Tribunal Europeo de Derechos Humanos. (1987). Case of Leander v. Sweden. Recuperado el 7 de abril de 2017, de <http://hudoc.echr.coe.int/eng?i=001-57519>
- Tribunal Europeo de Derechos Humanos. (2007). Case of Copland v. The United Kingdom. Recuperado el 9 de abril de 2017, de <http://hudoc.echr.coe.int/eng?i=001-79996>
- Tribunal Europeo de Derechos Humanos. (2008). Case of S. and Marper v. The United Kingdom. Recuperado el 10 de abril de 2017, de <http://hudoc.echr.coe.int/eng?i=001-90051>
- Tribunal Europeo de Derechos Humanos. (2010). Case of Uzun v. Germany. Recuperado el 10 de abril de 2017, de <http://hudoc.echr.coe.int/eng?i=001-57533>
- Trinidad y Tobago. (2011). Data Protection Act. *First Session Tenth Parliament Republic of Trinidad and Tobago [Act No. 13 of 2011]* Recuperado el 31 de marzo de 2017, de <http://www.ttparliament.org/legislations/a2011-13.pdf>
- Tunisia. (2004). Organic Act n°2004-63 of July 27th 2004 on the protection of personal data. *Tunisia National Data Protection Authority*, de 27 de julio de 2004. Recuperado el 7 de abril de 2017, de <http://desiderando.com/Docs/UNPAN042957.pdf>
- Turquía. (2016). Data Protection Law number 6698. *Official Gazette*, 29677, de 7 de abril de 2016. Recuperado el 29 de marzo de 2017, de <http://www.resmigazete.gov.tr/eskiler/2016/04/20160407-8.pdf>

- Ucrania. (2011). Law of Ukraine on Protection of Personal Data. *State Service of Ukraine on Personal Data Protection [No.2297-VI]* Recuperado el 28 de marzo de 2017, de http://www.legislationline.org/download/action/download/id/5569/file/Ukraine_law_protection_personal_data_2011_en.pdf
- Unión Europea. (1992). Tratado de la Unión Europea. *Diario Oficial de las Comunidades Europeas [C191/1]*, de 29 de julio de 1992. Recuperado el 10 de abril de 2017, de <http://eur-lex.europa.eu/legal-content/ES/TXT/PDF/?uri=OJ:C:1992:191:FULL&from=ES>
- Unión Europea. (1995). Directiva 95/46/CE del Parlamento Europeo y del Consejo de 24 de octubre de 1995 relativa a la protección de las personas físicas en lo que respecta al tratamiento de datos personales ya la libre circulación de estos datos. *Diario Oficial de las Comunidades Europeas [L281]*, 38, de 23 de noviembre de 1995. Recuperado el 15 de abril de 2017, de <http://eur-lex.europa.eu/legal-content/ES/ALL/?uri=OJ:L:1995:281:TOC>
- Unión Europea. (2004). Tratado por el que se establece una Constitución para Europa. *Diario Oficial de la Unión Europea [C310]*, 47, de 16 de diciembre de 2004. Recuperado el 15 de abril de 2017, de <http://eur-lex.europa.eu/legal-content/ES/TXT/PDF/?uri=OJ:C:2004:310:FULL&from=ES>
- Unión Europea. (2007). Tratado de Lisboa por el que se modifican el Tratado de la Unión Europea y el Tratado Constitutivo de la Comunidad Europea. *Diario Oficial de la Unión Europea [C306/01]*, 50, de 17 de diciembre 2007. Recuperado el 12 de abril de 2017, de <http://eur-lex.europa.eu/legal-content/en/ALL/?uri=OJ:C:2007:306:TOC>
- Unión Europea. (2016). Reglamento (UE) 2016/679 del Parlamento Europeo y del Consejo, de 27 de abril de 2016, 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 y por el que se deroga la Directiva 95/46/CE (Reglamento general de protección de datos). *Diario Oficial de la Unión Europea [L119/1]*,

59, de 4 de mayo de 2016. Recuperado el 11 de abril de 2017, de <http://eur-lex.europa.eu/legal-content/ES/TXT/PDF/?uri=OJ:L:2016:119:FULL&from=ES>

- Uruguay. (2008). Ley Nº 18.331 de Protección de datos personales y acción de "Habeas Data". *Diario Oficial*, 27549, de 18 de agosto de 2008. Recuperado el 31 de marzo de 2017, de https://datospersonales.gub.uy/wps/wcm/connect/urcdp/f085d1b8-0a24-4070-9dad-adc87b7595f2/18.331_con_modificaciones_de_la_19.355..pdf?MOD=AJPERES&CONVERT_TO=url&CACHEID=f085d1b8-0a24-4070-9dad-adc87b7595f2
- Uruguay. (2009). Decreto nº 414/009. *Diario Oficial*, 27815, de 15 de septiembre de 2009. Recuperado el 31 de marzo de 2017, de <http://www.dgi.gub.uy/wdgi/agxppdwn?6,4,167,O,S,0,12552%3B%3B2%3B877>
- Uzuner, Ö, Luo, Y., & Szolovits, P. (2007). Evaluating the state-of-the-art in automatic de-identification. *Journal of the American Medical Informatics Association*, 14(5), 550-563. doi:10.1197/jamia.M2444
- Veracruz. (2007). Ley No. 848 de Transparencia y Acceso a la Información Pública para el Estado de Veracruz de Ignacio de la Llave. *Gaceta Oficial del Estado*, tomo CLXXVI, 61, de 27 de febrero de 2007. Recuperado el 31 de marzo de 2017, de http://veracruz2014.mx/portal_transparencia/l/Ley_Transparencia_y_Acceso_Inf_Pub_EdodeVer_ILlave.pdf
- Verbeek, J. (2016). Critical technological dependency requires a revised privacy policy of major service providers. *CEPIS Legal & Security Issues*, Recuperado el 14 de junio de 2017, de <https://www.cepis.org/media/LSI.Statement.57Council.20161.pdf>
- Vietnam. (2005). Law on E-Transactions. *National Assembly [51/2005/QH11]*, de 29 de noviembre de 2005. Recuperado el 29 de marzo de 2017, de http://www.moj.gov.vn/vbpq/en/lists/vn%20bn%20php%20lut/view_detail.aspx?itemid=6109

- Villalba de Benito, María Teresa. (2016). Red sobre temas legales y seguridad CEPIS LSI SIN. *Novatica*, 238, 3-4. Recuperado el 14 de junio de 2017, de <http://www2.ati.es/novatica/2017/238/Nv238-BloqueEditorial.pdf>
- Villalba de Benito, María Teresa, De Buenaga Rodríguez, M., Gachet Páez, D., & Aparicio Galisteo, F. (2015). Security analysis of an IoT architecture for Healthcare. Trabajo presentado en la *In Healthy IoT: 2nd EAI International Conference on IoT Technologies for HealthCare*, , 169 454-460. doi:10.1007/978-3-319-47063-4_48
- Villalba, M. T., Fernández-Sanz, L., Cuadrado-Gallego, J., & Martínez, J. J. (2010). Software quality evaluation for security COTS products. Trabajo presentado en la *International Journal of Software Engineering and Knowledge Engineering*, World Scientific Publishing Co, , 20(1) 27-48. doi:10.1142/S0218194010004633
- Villalba, M. T., Fernández-Sanz, L., & Martínez, J. J. (2010). Empirical support for the generation of domain-oriented quality models. *IET Digital Library*, 4(1), 1-14. doi:10.1049/iet-sen.2009.0040
- Walters, K., Hardoon, S., Petersen, I., Iliffe, S., Omar, R. Z., Nazareth, I., & Rait, G. (2016). Predicting dementia risk in primary care: development and validation of the Dementia Risk Score using routinely collected data. *BMC Medicine*, 14(1) doi:10.1186/s12916-016-0549-y
- Wotton, C. J., & Goldacre, M. J. (2014). Age at obesity and association with subsequent dementia: record linkage study. *Postgraduate Medical Journal*, doi:10.1136/postgradmedj-2014-132571
- Zacatecas. (2011). Ley de Transparencia y Acceso a la Información Pública del Estado de Zacatecas. *Periódico Oficial del Estado [Decreto No. 149]*, tomo CXXI, 52, de 29 de junio de 2011. Recuperado el 30 de marzo de 2017, de http://omg.zacatecas.gob.mx/periodicozac/bajarArchivo.php?id_clasificacion=2&id_documento=1153

Zimbabwe. (2002). Access to Information and Protection of Privacy Act 5/2002. *Gazette Newspaper*, 116/2002, Chapter 10:27, de 15 de marzo de 2002. Recuperado el 4 de abril de 2017, de <http://www.africanplatform.org/uploads/media/Zimbabwe-Access-to-Information-and-Public-Privacy-Act-2002.pdf>

7. ANEXOS

ANEXOS

ANEXO 1 GLOSARIO DE TÉRMINOS RELACIONADOS CON LA PROTECCIÓN DE DATOS	263
ANEXO 2 ÍNDICE DE TABLAS	275
ANEXO 3 ÍNDICE DE FIGURAS	279
ANEXO 4 CUESTIONARIO	285
ANEXO 5 HISTOGRAMAS DE LAS VARIABLES ESTADÍSTICAS POR CATEGORÍAS QUE NO SE HAN ELIMINADO EN LA FASE DE ANÁLISIS PREVIO	309
ANEXO 6 MATRICES DE CORRELACIÓN POR CATEGORÍAS DE DATOS	335
ANEXO 7 RESULTADO DE LAS ENTREVISTAS.....	359

ANEXO 1 Glosario de términos relacionados con la protección de datos

A lo largo de este trabajo de tesis se hace uso de un gran número de términos. Por ello, y de forma alfabética, se listan los términos que a su vez de definen de forma cronológica por cada una de las siguientes normas:

- Directiva 95/46/CE del Parlamento Europeo y del Consejo de 24 de octubre de 1995, relativa a la protección de las personas físicas en lo que respecta al tratamiento de datos personales y a la libre circulación de estos datos (Unión Europea, 1995). En adelante Directiva.
- Ley Orgánica 15/1999, de 13 de diciembre, de protección de datos de carácter personal (España, 1999). En adelante LOPD.
- Real Decreto 1720/2007, de 21 de diciembre, que desarrolla la citada Ley (España, 2007). En adelante RLOPD.
- Reglamento (UE) 2016/679 del Parlamento Europeo y del Consejo de 27 de abril de 2016 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 y por el que se deroga la Directiva 95/46/CE (Reglamento general de protección de datos) (Unión Europea, 2016). En adelante RGPD.

Accesos autorizados

Autorizaciones concedidas a un usuario para la utilización de los diversos recursos. En su caso, incluirán las autorizaciones o funciones que tenga

atribuidas un usuario por delegación del responsable del fichero o tratamiento o del responsable de seguridad (RLOPD art. 5.2.a).

Afectado o interesado

Persona física titular de los datos que sean objeto del tratamiento a que se refiere el apartado c) del presente artículo (LOPD art. 3.e).

Persona física titular de los datos que sean objeto del tratamiento (RLOPD art. 5.1.a).

Autenticación

Procedimiento de comprobación de la identidad de un usuario (RLOPD 5.2.b).

Cancelación

Procedimiento en virtud del cual el responsable cesa en el uso de los datos. La cancelación implicará el bloqueo de los datos, consistente en la identificación y reserva de los mismos con el fin de impedir su tratamiento excepto para su puesta a disposición de las Administraciones públicas, Jueces y Tribunales, para la atención de las posibles responsabilidades nacidas del tratamiento y sólo durante el plazo de prescripción de dichas responsabilidades. Transcurrido ese plazo deberá procederse a la supresión de los datos (RLOPD art. 5.1b).

Cesión o comunicación de datos

Toda revelación de datos realizada a una persona distinta del interesado (LOPD art. 3.i).

Tratamiento de datos que supone su revelación a una persona distinta del interesado (RLOPD art. 5.1c).

Comunicación de datos

Toda revelación de datos realizada a una persona distinta del interesado (LOPD 3.i).

Consentimiento del interesado

Toda manifestación de voluntad, libre, específica e informada, mediante la que el interesado consienta el tratamiento de datos personales que le conciernan (Directiva 2.h).

Toda manifestación de voluntad, libre, inequívoca, específica e informada, mediante la que el interesado consienta el tratamiento de datos personales que le conciernen (LOPD art. 3.h).

Toda manifestación de voluntad, libre, inequívoca, específica e informada, mediante la que el interesado consienta el tratamiento de datos personales que le conciernen (RLOPD art. 5.1.d).

Toda manifestación de voluntad libre, específica, informada e inequívoca por la que el interesado acepta, ya sea mediante una declaración o una clara acción afirmativa, el tratamiento de datos personales que le conciernen (RGPD art. 4.11).

Datos genéticos

Datos personales relativos a las características genéticas heredadas o adquiridas de una persona física que proporcionen una información única sobre la fisiología o la salud de esa persona, obtenidos en particular del análisis de una muestra biológica de tal persona (RGPD art. 4.13).

Datos biométricos

Datos personales obtenidos a partir de un tratamiento técnico específico, relativos a las características físicas, fisiológicas o conductuales de una persona física que permitan o confirmen la identificación única de dicha persona, como imágenes faciales o datos dactiloscópicos (RGPD art. 4.14).

Dato disociado

Aquél que no permite la identificación de un afectado o interesado (RLOPD art. 5.1.e).

Datos de carácter personal o dato personal

Toda información sobre una persona física identificada o identificable (el «interesado»); se considerará identificable toda persona cuya identidad pueda determinarse, directa o indirectamente, en particular mediante un número de identificación o uno o varios elementos específicos, característicos de su identidad física, fisiológica, psíquica, económica, cultural o social (Directiva art. 2.a).

Cualquier información concerniente a personas físicas identificadas o identificables (LOPD art. 3.a).

Cualquier información numérica, alfabética, gráfica, fotográfica, acústica o de cualquier otro tipo concerniente a personas físicas identificadas o identificables (RLOPD art. 5.1.f).

Toda información sobre una persona física identificada o identificable («el interesado»); se considerará persona física identificable toda persona cuya identidad pueda determinarse, directa o indirectamente, en particular mediante un identificador, como por ejemplo un nombre, un número de identificación, datos de localización, un identificador en línea o uno o varios elementos propios de la identidad física, fisiológica, genética, psíquica, económica, cultural o social de dicha persona (RGPD art.4.1).

Datos de carácter personal relacionados con la salud

Las informaciones concernientes a la salud pasada, presente y futura, física o mental, de un individuo. En particular, se consideran datos relacionados con la salud de las personas los referidos a su porcentaje de discapacidad y a su información genética (RLOPD art. 5.1.g).

Datos personales relativos a la salud física o mental de una persona física, incluida la prestación de servicios de atención sanitaria, que revelen información sobre su estado de salud (RGPD art. 14.15).

Documento

Todo escrito, gráfico, sonido, imagen o cualquier otra clase de información que puede ser tratada en un sistema de información como una unidad diferenciada (RLOPD art. 5.2.f).

Encargado del tratamiento

Persona física o jurídica, autoridad pública, servicio o cualquier otro organismo que, solo o conjuntamente con otros, trate datos personales por cuenta del responsable del tratamiento (Directiva 2.e).

Persona física o jurídica, autoridad pública, servicio o cualquier otro organismo que, sólo o conjuntamente con otros, trate datos personales por cuenta del responsable del tratamiento (LOPD 3.g).

Persona física o jurídica, pública o privada, u órgano administrativo que, solo o conjuntamente con otros, trate datos personales por cuenta del responsable del tratamiento o del responsable del fichero, como consecuencia de la existencia de una relación jurídica que le vincula con el mismo y delimita el ámbito de su actuación para la prestación de un servicio. Podrán ser también encargados del tratamiento los entes sin personalidad jurídica que actúen en el tráfico como sujetos diferenciados (RLOPD art. 5.1.i).

La persona física o jurídica, autoridad pública, servicio u otro organismo que trate datos personales por cuenta del responsable del tratamiento (RGPD art. 4.8).

Fichero

Todo conjunto estructurado de datos personales, accesibles con arreglo a criterios determinados, ya sea centralizado, descentralizado o repartido de forma funcional o geográfica (Directiva art. 2.c).

Todo conjunto organizado de datos de carácter personal, cualquiera que fuere la forma o modalidad de su creación, almacenamiento, organización y acceso (LOPD 3.b).

Todo conjunto organizado de datos de carácter personal, que permita el acceso a los datos con arreglo a criterios determinados, cualquiera que fuere la forma o modalidad de su creación, almacenamiento, organización y acceso (RLOPD art. 5.1.k).

Todo conjunto estructurado de datos personales, accesibles con arreglo a criterios determinados, ya sea centralizado, descentralizado o repartido de forma funcional o geográfica (RGPD art. 4.6).

Identificación

Procedimiento de reconocimiento de la identidad de un usuario (RLOPD art. 5.2.h).

Interesado

Persona física titular de los datos que sean objeto del tratamiento (LOPD art. 3.e).

Persona identificable

Toda persona cuya identidad pueda determinarse, directa o indirectamente, mediante cualquier información referida a su identidad física, fisiológica, psíquica, económica, cultural o social. Una persona física no se considerará identificable si dicha identificación requiere plazos o actividades desproporcionados (RLOPD art. 5.1.o).

Procedimiento de disociación

Todo tratamiento de datos personales de modo que la información que se obtenga no pueda asociarse a persona identificada o identificable (LOPD 3.f).

Todo tratamiento de datos personales que permita la obtención de datos disociados (RLOPD art. 5.1.p).

Recurso

Cualquier parte componente de un sistema de información (RLOPD art. 5.2.k).

Responsable del fichero o del tratamiento

La persona física o jurídica, autoridad pública, servicio o cualquier otro organismo que sólo o conjuntamente con otros determine los fines y los medios del tratamiento de datos personales; en caso de que los fines y los medios del tratamiento estén determinados por disposiciones legislativas o reglamentarias nacionales o comunitarias, el responsable del tratamiento o los criterios específicos para su nombramiento podrán ser fijados por el Derecho nacional o comunitario (Directiva art. 2.d).

Persona física o jurídica, de naturaleza pública o privada, u órgano administrativo, que decida sobre la finalidad, contenido y uso del tratamiento (LOPD art. 3.d).

Persona física o jurídica, de naturaleza pública o privada, u órgano administrativo, que sólo o conjuntamente con otros decida sobre la finalidad, contenido y uso del tratamiento, aunque no lo realizase

materialmente. Podrán ser también responsables del fichero o del tratamiento los entes sin personalidad jurídica que actúen en el tráfico como sujetos diferenciados (RLOPD art. 5.1.q).

La persona física o jurídica, autoridad pública, servicio u otro organismo que, solo o junto con otros, determine los fines y medios del tratamiento; si el Derecho de la Unión o de los Estados miembros determina los fines y medios del tratamiento, el responsable del tratamiento o los criterios específicos para su nombramiento podrá establecerlos el Derecho de la Unión o de los Estados miembros (RGPD art. 4.7).

Responsable de seguridad

Persona o personas a las que el responsable del fichero ha asignado formalmente la función de coordinar y controlar las medidas de seguridad aplicables (RLOPD art. 5.2.l).

Seudonimización

El tratamiento de datos personales de manera tal que ya no puedan atribuirse a un interesado sin utilizar información adicional, siempre que dicha información adicional figure por separado y esté sujeta a medidas técnicas y organizativas destinadas a garantizar que los datos personales no se atribuyan a una persona física identificada o identificable (RGPD art. 4.5).

Sistema de información

Conjunto de ficheros, tratamientos, programas, soportes y en su caso, equipos empleados para el tratamiento de datos de carácter personal (RLOPD art. 5.2.m).

Soporte

Objeto físico que almacena o contiene datos o documentos, u objeto susceptible de ser tratado en un sistema de información y sobre el cual se pueden grabar y recuperar datos (RLOPD art. 5.2.ñ).

Transferencia internacional de datos

Tratamiento de datos que supone una transmisión de los mismos fuera del territorio del Espacio Económico Europeo, bien constituya una cesión o comunicación de datos, bien tenga por objeto la realización de un tratamiento de datos por cuenta del responsable del fichero establecido en territorio español (RLOPD art. 5.1.s).

Transmisión de documentos

Cualquier traslado, comunicación, envío, entrega o divulgación de la información contenida en el mismo (RLOPD art. 5.2.o).

Tratamiento de datos

Cualquier operación o conjunto de operaciones, efectuadas o no mediante procedimientos automatizados, y aplicadas a datos personales, como la recogida, registro, organización, conservación, elaboración o modificación, extracción, consulta, utilización, comunicación por transmisión, difusión o cualquier otra forma que facilite el acceso a los

mismos, cotejo o interconexión, así como su bloqueo, supresión o destrucción (Directiva art. 2.b).

Operaciones y procedimientos técnicos de carácter automatizado o no, que permitan la recogida, grabación, conservación, elaboración, modificación, bloqueo y cancelación, así como las cesiones de datos que resulten de comunicaciones, consultas, interconexiones y transferencias (LOPD art. 3.c).

Cualquier operación o procedimiento técnico, sea o no automatizado, que permita la recogida, grabación, conservación, elaboración, modificación, consulta, utilización, modificación, cancelación, bloqueo o supresión, así como las cesiones de datos que resulten de comunicaciones, consultas, interconexiones y transferencias (RLOPD art. 5.1.t).

Cualquier operación o conjunto de operaciones realizadas sobre datos personales o conjuntos de datos personales, ya sea por procedimientos automatizados o no, como la recogida, registro, organización, estructuración, conservación, adaptación o modificación, extracción, consulta, utilización, comunicación por transmisión, difusión o cualquier otra forma de habilitación de acceso, cotejo o interconexión, limitación, supresión o destrucción (RGPD art. 4.2).

Usuario

Sujeto o proceso autorizado para acceder a datos o recursos. Tendrán la consideración de usuarios los procesos que permitan acceder a datos o recursos sin identificación de un usuario físico (RLOPD art. 5.2.p).

ANEXO 2 Índice de tablas

TABLA 1 PROYECTO DE CARÁCTER PRIVADO EN EL QUE SE ENMARCA ESTA INVESTIGACIÓN	35
TABLA 2 PROYECTO DE CARÁCTER PÚBLICO EN EL QUE SE ENMARCA ESTA INVESTIGACIÓN.....	35
TABLA 3 CONTENIDO DE LA HISTORIA CLÍNICA EXTRAÍDA DEL ARTÍCULO 15 APARTADO 2	48
TABLA 4 MUESTRA DE ATRIBUTOS MÉDICOS DEL PACIENTE RECOGIDOS EN EL R.D. 1093/2010 (ESPAÑA, 2010).....	50
TABLA 5 ASPECTOS A TENER EN CUENTA EN UN PROCESO DE SEUDONIMIZACIÓN (UNIÓN EUROPEA, 2016).....	53
TABLA 6 RELACIÓN DE PAÍSES DE ÁFRICA Y SU NORMATIVA EN PROTECCIÓN DE DATOS.....	56
TABLA 7 CAMPOS DE INFORMACIÓN PERSONAL PROTEGIDOS EN SUDÁFRICA (SUDÁFRICA, 2002)	60
TABLA 8 NUEVOS CAMPOS DE INFORMACIÓN PERSONAL PROTEGIDOS EN SUDÁFRICA (SUDÁFRICA, 2013)	60
TABLA 9 DATOS SENSIBLES SEGÚN MARRUECOS (MARRUECOS, 2009B)	61
TABLA 10 RELACIÓN DE PAÍSES DE EUROPA Y NORMATIVA NACIONAL DE PROTECCIÓN DE DATOS	63
TABLA 11 DATOS SENSIBLES APORTADOS POR FINLANDIA RELACIONADOS CON DATOS LA SALUD (FINLANDIA, 1999).....	63
TABLA 12 DATOS SANITARIOS DEFINIDOS POR LITUANIA (LITUANIA, 1996)	64
TABLA 13 DATOS NUEVOS APORTADOS POR ISLANDIA (ISLANDIA, 2000).....	64
TABLA 14 ATRIBUTOS IDENTIFICABLES CATEGORIZADOS POR AUSTRALIA (AUSTRALIA, 1988) ..	65
TABLA 15 CAMPOS MÉDICOS DEFINIDOS POR AUSTRALIA (AUSTRALIA, 1988).....	65
TABLA 16 GRUPO 1 DE PAÍSES ASIÁTICOS	67
TABLA 17 GRUPO 2 DE PAÍSES ASIÁTICOS	67
TABLA 18 RELACIÓN DE PAÍSES DE AMÉRICA Y NORMATIVA DE PROTECCIÓN DE DATOS	71
TABLA 19 INFORMACIÓN PERSONAL DEFINIDA POR CANADÁ (CANADÁ, 1985).....	72
TABLA 20 INFORMACIÓN RELATIVA CON LA SALUD SEGÚN CANADÁ (CANADÁ, 2000)	72

TABLA 21 DATOS RELATIVOS CON LA SALUD SEGÚN TRINIDAD Y TOBAGO (TRINIDAD Y TOBAGO, 2011)	75
TABLA 22 DATOS RELATIVOS CON LA SALUD SEGÚN NUEVO LEÓN (NUEVO LEÓN, 2008).....	76
TABLA 23 DATOS RELATIVOS CON LA SALUD CON LA NUEVA NORMATIVA DE NUEVO LEÓN (NUEVO LEÓN, 2016)	77
TABLA 24 CUASI-IDENTIFICADORES COMUNES (INSTITUTE OF MEDICINE (IOM), 2015).....	84
TABLA 25 POSIBLES ATRIBUTOS IDENTIFICABLES (EL EMAM, 2014).....	84
TABLA 26 ALGUNAS DE LAS ENTIDADES RECONOCIDAS POR SCRUB SE ENUMERAN MÁS ARRIBA EN ORDEN RELATIVO DE PRIORIDAD (SWEENEY, 1996).....	85
TABLA 27 LISTADO DE DATOS MÉDICOS PROTEGIDOS A NIVEL INTERNACIONAL.....	89
TABLA 28 RELACIÓN DE CONTEXTOS DE VALIDACIÓN Y TÉCNICAS APLICADAS	108
TABLA 29 DATOS MÉDICOS ELIMINADOS EN EL PROCESO DE REVISIÓN INTERNA	125
TABLA 30 CATÁLOGO DE DATOS MÉDICOS DEL ÁMBITO SANITARIO DE ESPAÑA	129
TABLA 31 RESULTADOS DE DATOS DEMOGRÁFICOS RELACIONADOS CON LA PROFESIÓN DE LOS ENCUESTADOS	135
TABLA 32 ANÁLISIS DESCRIPTIVO PARA LAS VARIABLES ENGLOBADAS DENTRO DE DATOS PERSONALES DE CARÁCTER GENERAL.....	138
TABLA 33 ANÁLISIS DESCRIPTIVO PARA LAS VARIABLES ENGLOBADAS DENTRO DE DATOS PERSONALES DE CONTACTO	140
TABLA 34 ANÁLISIS DESCRIPTIVO PARA LAS VARIABLES ENGLOBADAS DENTRO DE DATOS PERSONALES ASOCIADOS A GEOLOCALIZACIÓN	142
TABLA 35 ANÁLISIS DESCRIPTIVO PARA LAS VARIABLES ENGLOBADAS DENTRO DE DATOS PERSONALES RELACIONADOS CON CREENCIAS, CULTURA, RAZA, ETNIA U OPINIONES	145
TABLA 36 ANÁLISIS DESCRIPTIVO PARA LAS VARIABLES ENGLOBADAS DENTRO DE DATOS PERSONALES RELACIONADOS CON LA PROFESIÓN.....	148
TABLA 37 ANÁLISIS DESCRIPTIVO PARA LAS VARIABLES ENGLOBADAS DENTRO DE DATOS PERSONALES ASOCIADOS A DONACIONES	152
TABLA 38 ANÁLISIS DESCRIPTIVO PARA LAS VARIABLES ENGLOBADAS DENTRO DE DATOS PERSONALES ASOCIADOS A ASEGURADORAS	154

TABLA 39 ANÁLISIS DESCRIPTIVO PARA LAS VARIABLES ENLOBADAS DENTRO DE DATOS DE DISPOSITIVOS IMPLANTADOS.....	157
TABLA 40 ANÁLISIS DESCRIPTIVO PARA LAS VARIABLES ENLOBADAS DENTRO DE DATOS MÉDICOS	162
TABLA 41 CATEGORÍA DE DATOS PERSONALES COMO RESULTADO DEL ANÁLISIS PREVIO	168
TABLA 42 CATEGORÍA DE DATOS MÉDICOS COMO RESULTADO DEL ANÁLISIS PREVIO	169
TABLA 43 ANÁLISIS DE FIABILIDAD.....	174
TABLA 44 PRUEBAS DE VALIDEZ PARA REALIZAR EL ANÁLISIS FACTORIAL EXPLORATORIO	175
TABLA 45 MATRIZ DE COMPONENTE ROTADO PARA DATOS PERSONALES.....	177
TABLA 46 MATRIZ DE COMPONENTE ROTADO PARA DATOS MÉDICOS	181
TABLA 47 TAXONOMÍA DE DATOS MÉDICOS DE CARÁCTER PERSONAL	182
TABLA 48 ANÁLISIS DE BONDAD DE AJUSTE PARA LAS ECUACIONES ESTRUCTURALES INICIALES PARA LA CATEGORÍA "DATOS PERSONALES"	185
TABLA 49 ANÁLISIS DE BONDAD DE AJUSTE PARA LAS ECUACIONES ESTRUCTURALES INICIALES PARA LA CATEGORÍA "DATOS MÉDICOS"	186
TABLA 50 PRUEBA DEL MULTIPLICADOR DE LAGRANGE PARA LA INCLUSIÓN DE NUEVAS RELACIONES ENTRE FACTORES E INDICADORES PARA LA CATEGORÍA "DATOS PERSONALES"	187
TABLA 51 PRUEBA DEL MULTIPLICADOR DE LAGRANGE PARA LA INCLUSIÓN DE NUEVAS RELACIONES ENTRE FACTORES E INDICADORES PARA LA CATEGORÍA "DATOS MÉDICOS"	189
TABLA 52 MODIFICACIONES REALIZADAS PARA MEJORAR LA BONDAD DE AJUSTE DEL MODELO	190
TABLA 53 BONDAD DE AJUSTE DEL AJUSTE FINAL DEL MODELO PARA LA CATEGORÍA "DATOS PERSONALES"	191
TABLA 54 COMPARACIÓN DE ÍNDICES DE BONDAD DE AJUSTE PARA LA CATEGORÍA "DATOS PERSONALES"	192
TABLA 55 BONDAD DE AJUSTE DEL MODELO FINAL PARA LA CATEGORÍA "DATOS MÉDICOS" ...	192
TABLA 56 COMPARACIÓN DE ÍNDICES DE BONDAD DE AJUSTE PARA LA CATEGORÍA "DATOS MÉDICOS"	193

TABLA 57 COMPARATIVA DEL MODELO PROPORCIONADO POR HIPAA Y EL MODELO VALIDADO 204

TABLA 58 MATRIZ DE CORRELACIÓN PARA LA CATEGORÍA "DATOS PERSONALES" PARTE 1 DE 2 337

TABLA 59 MATRIZ DE CORRELACIÓN PARA LA CATEGORÍA "DATOS PERSONALES" PARTE 2 DE 2 338

TABLA 60 MATRIZ DE CORRELACIÓN PARA LA CATEGORÍA "DATOS MÉDICOS" PARTE 1 DE 6... 341

TABLA 61 MATRIZ DE CORRELACIÓN PARA LA CATEGORÍA "DATOS MÉDICOS" PARTE 2 DE 6... 344

TABLA 62 MATRIZ DE CORRELACIÓN PARA LA CATEGORÍA "DATOS MÉDICOS" PARTE 3 DE 6... 347

TABLA 63 MATRIZ DE CORRELACIÓN PARA LA CATEGORÍA "DATOS MÉDICOS" PARTE 4 DE 6... 350

TABLA 64 MATRIZ DE CORRELACIÓN PARA LA CATEGORÍA "DATOS MÉDICOS" PARTE 5 DE 6... 353

TABLA 65 MATRIZ DE CORRELACIÓN PARA LA CATEGORÍA "DATOS MÉDICOS" PARTE 6 DE 6... 357

TABLA 66 DATOS MÉDICOS RESULTANTES DE LA REVISIÓN INTERNA..... 363

ANEXO 3 Índice de figuras

FIGURA 1 PRIMEROS CUASI-IDENTIFICADORES IDENTIFICADOS (EL EMAM ET AL., 2006).....	83
FIGURA 2 METODOLOGÍA PARA LA CREACIÓN DE MODELOS DE DATOS MÉDICOS A ANONIMIZAR (ADAPTACIÓN DE LA METODOLOGÍA DUMOD).....	98
FIGURA 3 RESUMEN DEL PROCESO LLEVADO A CABO PARA EXTRAER EL MODELO DE DATOS MÉDICOS	99
FIGURA 4: ETAPAS DEL MÉTODO DE ENCUESTA (AMELLER ET AL., 2015)	102
FIGURA 5 ETAPAS DEL ANÁLISIS CONFIRMATORIO	114
FIGURA 6 DATOS DEMOGRÁFICOS RELACIONADOS CON LA EDAD DE LOS ENCUESTADOS.....	134
FIGURA 7 DATOS DEMOGRÁFICOS RELACIONADOS CON EL PAÍS DE ORIGEN DE LOS ENCUESTADOS	134
FIGURA 8 DATOS DEMOGRÁFICOS RELACIONADOS CON LA FAMILIARIDAD CON PROTECCIÓN DE DATOS.....	136
FIGURA 9 DATOS DEMOGRÁFICOS RELACIONADOS CON LA FORMACIÓN EN PROTECCIÓN DE DATOS	136
FIGURA 10 HISTOGRAMA DEL ATRIBUTO DPCG4 (PERSONA DE CONTACTO DEL PACIENTE) CON ASIMETRÍA NEGATIVA	139
FIGURA 11 HISTOGRAMA DEL ATRIBUTO DPCG5 (ESTADO CIVIL) CON ASIMETRÍA NEGATIVA.	139
FIGURA 12 HISTOGRAMA DEL ATRIBUTO DPC2 (FAX) CON ASIMETRÍA NEGATIVA	141
FIGURA 13 HISTOGRAMA DEL ATRIBUTO DPC3 (CORREO ELECTRÓNICO) CON ASIMETRÍA NEGATIVA	141
FIGURA 14 HISTOGRAMA DEL ATRIBUTO DPAG1 (CALLE) CON ASIMETRÍA NEGATIVA	143
FIGURA 15 HISTOGRAMA DEL ATRIBUTO DPAG4 (DISTRITO) CON ASIMETRÍA NEGATIVA	143
FIGURA 16 HISTOGRAMA DEL ATRIBUTO DPAG5 (DIRECCIÓN) CON ASIMETRÍA NEGATIVA....	144
FIGURA 17 HISTOGRAMA DEL ATRIBUTO DPAG8 (CÓDIGO POSTAL) CON ASIMETRÍA NEGATIVA	144
FIGURA 18 HISTOGRAMA DEL ATRIBUTO DPRCCREO2 (NECESIDADES PERSONALES) CON ASIMETRÍA NEGATIVA	146

FIGURA 19 HISTOGRAMA DEL ATRIBUTO DPRCCREO3 (ORIENTACIÓN SEXUAL) CON ASIMETRÍA NEGATIVA	146
FIGURA 20 HISTOGRAMA DEL ATRIBUTO DPRCCREO4 (ORIGEN SOCIAL) CON ASIMETRÍA NEGATIVA	147
FIGURA 21 HISTOGRAMA DEL ATRIBUTO DPRCCREO5 (PLURALIDAD DE NACIMIENTO) CON ASIMETRÍA NEGATIVA	147
FIGURA 22 HISTOGRAMA DEL ATRIBUTO DPRP1 (AÑOS DE ESCOLARIDAD) CON ASIMETRÍA NEGATIVA	149
FIGURA 23 HISTOGRAMA DEL ATRIBUTO DPRP2 (HISTORIAL DE EDUCACIÓN) CON ASIMETRÍA NEGATIVA	149
FIGURA 24 HISTOGRAMA DEL ATRIBUTO DPRP3 (HISTORIAL DE EMPLEO) CON ASIMETRÍA NEGATIVA	150
FIGURA 25 HISTOGRAMA DEL ATRIBUTO DPRP4 (POSICIÓN DENTRO DE LA FAMILIA) CON ASIMETRÍA NEGATIVA	150
FIGURA 26 HISTOGRAMA DEL ATRIBUTO DPRP5 (NOMBRE DEL ÚLTIMO EMPLEO) CON ASIMETRÍA NEGATIVA	151
FIGURA 27 HISTOGRAMA DEL ATRIBUTO DPRP6 (ORIGEN EN LA QUE TRABAJA) CON ASIMETRÍA NEGATIVA	151
FIGURA 28 HISTOGRAMA DEL ATRIBUTO DPAD1 (DONACIÓN DE SUSTANCIA CORPORAL) CON ASIMETRÍA NEGATIVA	153
FIGURA 29 HISTOGRAMA DEL ATRIBUTO DPAD2 (DONACIÓN DE PARTE CORPORAL) CON ASIMETRÍA NEGATIVA	153
FIGURA 30 HISTOGRAMA DEL ATRIBUTO DPAA1 (INGRESOS TOTALES) CON ASIMETRÍA NEGATIVA	155
FIGURA 31 HISTOGRAMA DEL ATRIBUTO DPAA2 (PROVEEDOR DE ATENCIÓN MÉDICA) CON ASIMETRÍA NEGATIVA	155
FIGURA 32 HISTOGRAMA DEL ATRIBUTO DPAA3 (NÚMERO DE LA SEGURIDAD SOCIAL) CON ASIMETRÍA NEGATIVA	156

FIGURA 33 HISTOGRAMA DEL ATRIBUTO DPAA4 (NÚMERO DE TARJETA SANITARIA) CON ASIMETRÍA NEGATIVA	156
FIGURA 34 HISTOGRAMA DEL ATRIBUTO DDI1 CON ASIMETRÍA NEGATIVA	158
FIGURA 35 HISTOGRAMA DEL ATRIBUTO DDI2 CON ASIMETRÍA NEGATIVA	158
FIGURA 36 HISTOGRAMA DEL ATRIBUTO DDI3 CON ASIMETRÍA NEGATIVA	159
FIGURA 37 HISTOGRAMA DEL ATRIBUTO DDI5 CON ASIMETRÍA NEGATIVA	159
FIGURA 38 HISTOGRAMA DEL ATRIBUTO DM1 (TIPO SANGUÍNEO) CON ASIMETRÍA NEGATIVA	163
FIGURA 39 HISTOGRAMA DEL ATRIBUTO DM4 (ÁCIDO DESOXIRRIBONUCLEICO) CON ASIMETRÍA NEGATIVA	163
FIGURA 40 HISTOGRAMA DEL ATRIBUTO DM23 (IMÁGENES FOTOGRÁFICAS IDENTIFICABLES) CON ASIMETRÍA NEGATIVA	164
FIGURA 41 HISTOGRAMA DEL ATRIBUTO DM25 (GENÉTICA DE FAMILIAR CERCANO) CON ASIMETRÍA NEGATIVA	164
FIGURA 42 HISTOGRAMA DEL ATRIBUTO DM26 (GENÉTICA PROPIA) CON ASIMETRÍA NEGATIVA	165
FIGURA 43 HISTOGRAMA DEL ATRIBUTO DM29 (NÚMERO DE PLACA MÉDICA) CON ASIMETRÍA NEGATIVA	165
FIGURA 44 HISTOGRAMA DEL ATRIBUTO DM30 (NÚMERO DE REGISTRO MÉDICO) CON ASIMETRÍA NEGATIVA	166
FIGURA 45 HISTOGRAMA DEL ATRIBUTO DM31 (NÚMERO DE SERIE DE DISPOSITIVOS IDENTIFICADORES) CON ASIMETRÍA NEGATIVA	166
FIGURA 46 HISTOGRAMA DEL ATRIBUTO DM33 (CÓDIGOS DE EVENTOS ADVERSOS) CON ASIMETRÍA NEGATIVA	167
FIGURA 47 DISTRIBUCIÓN DE LAS RESPUESTAS A LA PREGUNTA SOBRE LA IMPORTANCIA DE UN MODELO APLICABLE DE PROTECCIÓN DE DATOS MÉDICOS.....	171
FIGURA 48 REPRESENTACIÓN GRÁFICA DEL MODELO OBTENIDO PARA LA CATEGORÍA "DATOS PERSONALES" EN NOTACIÓN LISREL.....	194
FIGURA 49 REPRESENTACIÓN GRÁFICA DEL MODELO OBTENIDO PARA LA CATEGORÍA "DATOS MÉDICOS" EN NOTACIÓN LISREL	195

FIGURA 50 HISTOGRAMA DEL ATRIBUTO DPCG1 (NOMBRE)	309
FIGURA 51 HISTOGRAMA DEL ATRIBUTO DPCG2 (PRIMER APELLIDO).....	309
FIGURA 52 HISTOGRAMA DEL ATRIBUTO DPCG3 (SEGUNDO APELLIDO).....	310
FIGURA 53 HISTOGRAMA DEL ATRIBUTO DPCG6 (GÉNERO)	310
FIGURA 54 HISTOGRAMA DEL ATRIBUTO DPCG7 (NACIONALIDAD).....	311
FIGURA 55 HISTOGRAMA DEL ATRIBUTO DPCG8 (FECHA DE NACIMIENTO).....	311
FIGURA 56 HISTOGRAMA DEL ATRIBUTO DPCG9 (EDAD O ELEMENTO INDICATIVO DE EDAD).	312
FIGURA 57 HISTOGRAMA DEL ATRIBUTO DPCG10 (SEXO)	312
FIGURA 58 HISTOGRAMA DEL ATRIBUTO DPC1 (TELÉFONO)	313
FIGURA 59 HISTOGRAMA DEL ATRIBUTO DPAG2 (CIUDAD)	313
FIGURA 60 HISTOGRAMA DEL ATRIBUTO DPAG3 (MUNICIPIO)	314
FIGURA 61 HISTOGRAMA DEL ATRIBUTO DPAG6 (COMUNIDAD AUTÓNOMA).....	314
FIGURA 62 HISTOGRAMA DEL ATRIBUTO DPAG7 (PAÍS DE NACIMIENTO)	315
FIGURA 63 HISTOGRAMA DEL ATRIBUTO DPRCCREO1 (ORIGEN ÉTNICO).....	315
FIGURA 64 HISTOGRAMA DEL ATRIBUTO DPRP7 (PROFESIÓN).....	316
FIGURA 65 HISTOGRAMA DEL ATRIBUTO DDI4 (PRÓTESIS)	316
FIGURA 66 HISTOGRAMA DEL ATRIBUTO DM2 (SERVICIO PRESTADO)	317
FIGURA 67 HISTOGRAMA DEL ATRIBUTO DM3 (TRATAMIENTO)	317
FIGURA 68 HISTOGRAMA DEL ATRIBUTO DM5 (ALERGIAS)	318
FIGURA 69 HISTOGRAMA DEL ATRIBUTO DM6 (CENTRO DE SALUD).....	318
FIGURA 70 HISTOGRAMA DEL ATRIBUTO DM7 (CONSUMO DE SUSTANCIAS TÓXICAS)	319
FIGURA 71 HISTOGRAMA DEL ATRIBUTO DM8 (DIAGNOSIS)	319
FIGURA 72 HISTOGRAMA DEL ATRIBUTO DM9 (ENFERMEDADES)	320
FIGURA 73 HISTOGRAMA DEL ATRIBUTO DM10 (ESTADO DE EMBARAZO)	320
FIGURA 74 HISTOGRAMA DEL ATRIBUTO DM11 (ESTADO DE ENFERMEDAD).....	321
FIGURA 75 HISTOGRAMA DEL ATRIBUTO DM12 (HISTORIAL CLÍNICO).....	321
FIGURA 76 HISTOGRAMA DEL ATRIBUTO DM13 (INFORMACIÓN PSICOLÓGICA)	322
FIGURA 77 HISTOGRAMA DEL ATRIBUTO DM14 (INFORMACIÓN PSIQUIÁTRICA)	322

FIGURA 78 HISTOGRAMA DEL ATRIBUTO DM15 (INFORMACIÓN RECOLECTADA DURANTE UN SERVICIO SANITARIO)	323
FIGURA 79 HISTOGRAMA DEL ATRIBUTO DM16 (INTERVENCIONES QUIRÚRGICAS)	323
FIGURA 80 HISTOGRAMA DEL ATRIBUTO DM17 (LABORATORIOS QUE ANALIZARON LAS PRUEBAS MÉDICAS)	324
FIGURA 81 HISTOGRAMA DEL ATRIBUTO DM18 (MEDIDAS CORPORALES)	324
FIGURA 82 HISTOGRAMA DEL ATRIBUTO DM19 (PROGNOSIS)	325
FIGURA 83 HISTOGRAMA DEL ATRIBUTO DM20 (RESIDENCIAS DE ANCIANOS).....	325
FIGURA 84 HISTOGRAMA DEL ATRIBUTO DM21 (VACUNAS)	326
FIGURA 85 HISTOGRAMA DEL ATRIBUTO DM22 (INCAPACIDADES MÉDICAS).....	326
FIGURA 86 HISTOGRAMA DEL ATRIBUTO DM24 (ESTADO DE DISCAPACIDAD)	327
FIGURA 87 HISTOGRAMA DEL ATRIBUTO DM27 (NÚMERO DE HISTORIA CLÍNICA)	327
FIGURA 88 HISTOGRAMA DEL ATRIBUTO DM28 (NÚMERO DE IDENTIFICACIÓN)	328
FIGURA 89 HISTOGRAMA DEL ATRIBUTO DM32 (CÓDIGOS DE DIAGNOSIS).....	328
FIGURA 90 HISTOGRAMA DEL ATRIBUTO DM34 (CÓDIGOS DE TRATAMIENTO).....	329
FIGURA 91 HISTOGRAMA DEL ATRIBUTO DM35 (FECHA DE ADMISIÓN)	329
FIGURA 92 HISTOGRAMA DEL ATRIBUTO DM36 (FECHA DE ALTA)	330
FIGURA 93 HISTOGRAMA DEL ATRIBUTO DM37 (FECHA DE DECESO).....	330
FIGURA 94 HISTOGRAMA DEL ATRIBUTO DM38 (FECHA DE RECOGIDA DE MUESTRAS)	331
FIGURA 95 HISTOGRAMA DEL ATRIBUTO DM39 (FECHA DE TRATAMIENTO).....	331
FIGURA 96 HISTOGRAMA DEL ATRIBUTO DM40 (FECHA DE VISITA).....	332
FIGURA 97 HISTOGRAMA DEL ATRIBUTO DM41 (NÚMERO DE REGISTROS DE ENSAYOS CLÍNICOS)	332
FIGURA 98 HISTOGRAMA DEL ATRIBUTO DM42 (SALUD FÍSICA)	333
FIGURA 99 HISTOGRAMA DEL ATRIBUTO DM43 (SALUD MENTAL)	333
FIGURA 100 HISTOGRAMA DEL ATRIBUTO DM44 (TÉRMINOS MÉDICOS QUE IDENTIFICAN PATOLOGÍAS NO HABITUALES)	334

ANEXO 4 Cuestionarios

Como se ha indicado en el punto anterior, existen dos cuestionarios orientados a diferentes profesionales vinculados al campo sanitario. Ambos cuestionarios presentan un acceso inicial común a los dos perfiles de usuarios. En este acceso común se muestra el propósito del estudio junto con una breve descripción. Posteriormente el cuestionario se divide en tres secciones:

- Antes de empezar.
- Cuestionario sobre el grado de importancia.
- Para terminar.

En la sección “Antes de empezar” se solicita información de datos demográficos para que más tarde se utilicen para verificar el perfil del participante en el estudio. En esta sección, los cuestionarios difieren en las siguientes preguntas:

- Entendiendo que en todos los casos trabaja en relación con el área médica, ¿qué tipo de actividad refleja mejor el sector de la organización para la que trabaja?
- ¿Cuál es su especialidad en la organización para la que trabaja?

En la sección “Cuestionario sobre el grado de importancia” se le solicita al participante que responda sobre el grado de importancia que tienen los campos de información mostrados en su trabajo diario.

En la sección “Para terminar” el usuario puede sugerir otro atributo que utiliza y que puede ser de interés su evaluación en el estudio. Además se le solicita qué grado de importancia cree que debería tener el saber los

atributos que deben ser eliminados, seudonimizados o anonimizados para garantizar la confidencialidad de los mismos.

CUESTIONARIO ATI

Estudio de investigación sobre los atributos clínicos con necesidad de anonimizar antes del tratamiento de datos

De antemano muchas gracias por su participación. La encuesta no le llevará más de 10 minutos.

¿CUÁL ES EL OBJETIVO DEL ESTUDIO?

En el sector sanitario se encuentra almacenada una gran cantidad de datos en formato digital que permite o permitiría a profesionales e investigadores extraer conocimiento. Pero estos datos clínicos se encuentran bajo la protección de la Ley Orgánica 15/1999 de 13 de diciembre de Protección de Datos de Carácter Personal clasificados como “datos especialmente protegidos”. En 2007 se aprobó la Ley 14/2007 de Investigación Biomédica que permite su utilización siempre y cuando se realice un proceso de anonimización que no permita la identificación de los usuarios. En ningún caso se especifican qué datos o atributos se deben anonimizar aunque existen numerosos estudios que demuestran que eliminar datos del tipo nombre y DNI no es suficiente.

Los atributos que se muestran a continuación han sido extraídos de artículos y normativa de protección de datos aplicada en otros países como datos que permiten identificar a un individuo ya sea

individualmente, combinados o cruzados con otros que se pueden obtener fácilmente. El objetivo de este estudio es conocer qué datos son esenciales y, por tanto, no pueden ser objeto de procesos de anonimización.

ANTES DE EMPEZAR... (1/3)

Por favor, proporciónenos información general sobre el tipo de organización para la que trabaja, su función en la misma y sus conocimientos/experiencia profesional.

¿En qué rango de edad se encuentra? Seleccione una de las siguientes opciones

- 25 años o menos
- Entre 26 y 35 años
- Entre 36 y 49 años
- Entre 50 y 65 años
- Más de 65 años

¿Cuál es su país? Seleccione una de las siguientes opciones

Entendiendo que en todos los casos trabaja en relación con el área médica, ¿qué tipo de actividad refleja mejor el sector de la organización para la que trabaja? Seleccione una de las siguientes opciones

- Servicios y consultoría informática
- Acción humanitaria

- Educación y Formación
- I+D
- Organismos oficiales
- Salud pública
- Salud privada
- Salud militar
- Farmacia
- Gestión hospitalaria
- Recursos humanos
- Administraciones públicas
- Otro:

ANTES DE EMPEZAR... (2/3)

¿Cuál es su especialidad en la organización en la que trabaja? Seleccione una de las siguientes opciones

¿Qué grado de experiencia considera que tiene en dicha función o puesto?

- Ninguna
- Menos de 5 años
- Entre 5 y 9 años
- Entre 10 y 15 años

- Más de 15 años

¿Cuál es su grado de familiaridad con la protección de datos? Seleccione una de las siguientes opciones

- Ninguno
- Nivel usuario/paciente
- Responsable de fichero de datos personales
- Experto

ANTES DE EMPEZAR... (3/3)

SOBRE LA INFORMACIÓN PROPORCIONADA

Durante el proceso no se pedirá ninguna información personal, tan solo la relacionada con sus conocimientos y experiencia profesional. Al finalizar el estudio, si así lo desea, le enviaremos un detallado informe de los resultados obtenidos a la dirección de correo electrónico que nos facilite.

La información aquí proporcionada se tratará únicamente para los fines del estudio anteriormente descritos, será estrictamente confidencial y en ningún caso se tratará de forma individual.

Si lo desea podemos enviarle los resultados del estudio una vez concluido. Para ello, por favor, introduzca una cuenta de correo electrónico a la que desea que se lo remitamos. Le aseguramos que no usaremos dicha dirección electrónica para ninguna otra finalidad.

--

CUESTIONARIO SOBRE EL GRADO DE IMPORTANCIA (1/4)

Indique con qué frecuencia usa los datos que aparecen a continuación en su trabajo o investigaciones (casi nunca, a veces, normalmente, con frecuencia, siempre).

Por ejemplo, para investigaciones sobre enfermedades tropicales ¿qué frecuencia de uso tiene el origen étnico del paciente en los estudios médicos? ¿Podría prescindirse de ese dato o es fundamental?

Los atributos se han agrupado para facilitar la comprensión del encuestado.

Datos personales de carácter general:

	Nunca o casi nunca	A veces	Normalmente	Con frecuencia	Siempre
Nombre					
Primer apellido					
Segundo apellido					
Persona de contacto del paciente					
Estado civil					
Género					
Nacionalidad					
Fecha de nacimiento					
Edad o elemento indicativo de edad					
Sexo					

Datos personales de contacto:

	Nunca o casi nunca	A veces	Normalmente	Con frecuencia	Siempre
Teléfono					
Fax					
Correo electrónico					

Datos personales asociados a geolocalización:

	Nunca o casi nunca	A veces	Normalmente	Con frecuencia	Siempre
Calle					
Ciudad					
Municipio					
Distrito					
Dirección					
Comunidad autónoma					
País de nacimiento					
Código postal					

CUESTIONARIO SOBRE EL GRADO DE IMPORTANCIA (2/4)

¿Con qué frecuencia usa estos datos en su trabajo o investigaciones?

Datos personales asociados a donaciones:

	Nunca o casi nunca	A veces	Normalmente	Con frecuencia	Siempre
Origen étnico					
Necesidades personales					
Orientación sexual					
Origen social					
Pluralidad de nacimiento					

Datos personales relacionados con la profesión:

	Nunca o casi nunca	A veces	Normalmente	Con frecuencia	Siempre
Años de escolaridad					
Historial de educación					
Historial de empleo					
Posición dentro de la familia (padre, hermano mayor, etc)					
Nombre del último empleo					
Organización en la que trabaja					
Profesión					

CUESTIONARIO SOBRE EL GRADO DE IMPORTANCIA (3/4)

¿Con qué frecuencia usa esos datos en su trabajo o investigaciones?

Datos personales asociados a donaciones:

	Nunca o casi nunca	A veces	Normalmente	Con frecuencia	Siempre
Donación de sustancia corporal					
Donación de parte corporal					

Datos personales asociados a aseguradoras:

	Nunca o casi nunca	A veces	Normalmente	Con frecuencia	Siempre
Ingresos totales					
Proveedor de atención médica					
Número de la Seguridad Social					
Número de tarjeta sanitaria					

Dispositivos implantados:

	Nunca o casi nunca	A veces	Normalmente	Con frecuencia	Siempre
Uso de aparatos auditivos					
Uso de aparatos oftalmológicos					
Uso de aparatos ortopédicos					
Prótesis					
Número identificativo de los dispositivos implantados					

CUESTIONARIO SOBRE EL GRADO DE IMPORTANCIA (4/4)

¿Con qué frecuencia usa esos datos en su trabajo o investigaciones?

Datos médicos:

	Nunca o casi nunca	A veces	Normalmente	Con frecuencia	Siempre
Tipo sanguíneo					
Servicio prestado					
Tratamiento					
Ácido					

	Nunca o casi nunca	A veces	Normalmente	Con frecuencia	Siempre
desoxirribonucleico					
Alergias					
Centros de salud					
Consumo de sustancias tóxicas					
Diagnosis					
Enfermedades					
Estado de embarazo					
Estado de enfermedad					
Historial clínico					
Información psicológica					
Información psiquiátrica					
Información recolectada durante un servicio sanitario					
Intervenciones quirúrgicas					
Laboratorios que analizaron las pruebas médicas					
Medidas corporales (talla, peso)					
Prognosis					
Residencias de ancianos					
Vacunas					
Incapacidades médicas					
Imágenes fotográficas identificables					
Estado de					

	Nunca o casi nunca	A veces	Normalmente	Con frecuencia	Siempre
discapacidad					
Genética de familiar cercano					
Genética propia					
Número de historia clínica					
Número de identificación					
Número de placa médica					
Número de registro médico					
Número de serie de dispositivos identificadores					
Códigos de diagnóstico					
Códigos de eventos adversos					
Códigos de tratamiento					
Fecha de admisión					
Fecha de alta					
Fecha de deceso					
Fecha de recogida de muestras					
Fecha de tratamiento					
Fecha de visita					
Número de registros de ensayos clínicos					
Salud física (otra información física no recogida en anteriores campos)					


	Nunca o casi nunca	A veces	Normalmente	Con frecuencia	Siempre
Salud mental (otra información mental no recogida en anteriores campos)					
Términos médicos que identifican patologías no habituales					

PARA TERMINAR...

Si cree que hay algún atributo clínico que sería importante incluir y que no aparece aquí, por favor, indíquelo a continuación:

¿Cree que es importante que profesionales, investigadores y/o docentes conozcan qué atributos deben ser eliminados, deidentificados o anonimizados para asegurar la confidencialidad de los datos de los pacientes? Seleccione una de las siguientes opciones

- No es importante
- Poco importante
- Moderadamente importante
- Importante

 Muy importante

CUESTIONARIO GOOGLE DRIVE

Estudio de investigación sobre los atributos clínicos con necesidad de anonimizar antes del tratamiento de datos

De antemano muchas gracias por su participación. La encuesta no le llevará más de 10 minutos.

¿CUÁL ES EL OBJETIVO DEL ESTUDIO?

En el sector sanitario se encuentra almacenada una gran cantidad de datos en formato digital que permite o permitiría a profesionales e investigadores extraer conocimiento. Pero estos datos clínicos se encuentran bajo la protección de la Ley Orgánica 15/1999 de 13 de diciembre de Protección de Datos de Carácter Personal clasificados como “datos especialmente protegidos”. En 2007 se aprobó la Ley 14/2007 de Investigación Biomédica que permite su utilización siempre y cuando se realice un proceso de anonimización que no permita la identificación de los usuarios. En ningún caso se especifican qué datos o atributos se deben anonimizar aunque existen numerosos estudios que demuestran que eliminar datos del tipo nombre y DNI no es suficiente.

Los atributos que se muestran a continuación han sido extraídos de artículos y normativa de protección de datos aplicada en otros países como datos que permiten identificar a un individuo ya sea individualmente, combinados o cruzados con otros que se pueden obtener fácilmente. El objetivo de este estudio es conocer qué datos son esenciales y, por tanto, no pueden ser objeto de procesos de anonimización.

ANTES DE EMPEZAR... (1/3)

Por favor, proporciónenos información general sobre el tipo de organización para la que trabaja, su función en la misma y sus conocimientos/experiencia profesional.

¿En qué rango de edad se encuentra?

- 25 años o menos
- Entre 26 y 35 años
- Entre 36 y 49 años
- Entre 50 y 65 años
- Más de 65 años

¿Cuál es su país?

Entendiendo que en todos los casos trabaja en relación con el área médica, ¿qué tipo de actividad refleja mejor el sector de la organización para la que trabaja?

- Informática y Telecomunicaciones
- Administraciones públicas
- Acción humanitaria
- Educación y Formación
- Organismos oficiales
- I+D
- Salud pública
- Salud privada
- Salud militar
- Farmacia
- Marketing y comunicación
- Recursos Humanos
- Seguros
- Gestión hospitalaria

ANTES DE EMPEZAR... (2/3)

¿Cuál es su especialidad en la organización en la que trabaja?

¿Qué grado de experiencia considera que tiene en dicha función o puesto?

- Ninguna
- Menos de 5 años
- Entre 5 y 9 años
- Entre 10 y 15 años

- Más de 15 años

¿Qué nivel de conocimiento posee sobre protección de datos?

- Ninguno
- Cursos básicos
- Cursos avanzados
- Máster en protección de datos

¿Cuál es su grado de familiaridad con la protección de datos?

- Ninguno
- Nivel usuario/paciente
- Responsable de fichero de datos personales
- Experto

ANTES DE EMPEZAR... (3/3)

¿Cuál es la tarea que desarrolla en la organización en la que trabaja?

- MIR
- Cuidado de pacientes
- Investigación
- Gestión hospitalaria
- Gestión TIC

Si lo desea podemos enviarle los resultados del estudio una vez concluido. Para ello, por favor, introduzca una cuenta de correo

electrónico a la que desea que se lo remitamos. Le aseguramos que no usaremos dicha dirección electrónica para ninguna otra finalidad.

SOBRE LA INFORMACIÓN PROPORCIONADA

Durante el proceso no se pedirá ninguna información personal, tan solo la relacionada con sus conocimientos y experiencia profesional. Al finalizar el estudio, si así lo desea, le enviaremos un detallado informe de los resultados obtenidos a la dirección de correo electrónico que nos facilite.

La información aquí proporcionada se tratará únicamente para los fines del estudio anteriormente descritos, será estrictamente confidencial y en ningún caso se tratará de forma individual.

CUESTIONARIO SOBRE EL GRADO DE IMPORTANCIA (1/4)

Indique con qué frecuencia usa los datos que aparecen a continuación en su trabajo o investigaciones (casi nunca, a veces, normalmente, con frecuencia, siempre).

Por ejemplo, para investigaciones sobre enfermedades tropicales ¿qué frecuencia de uso tiene el origen étnico del paciente en los estudios médicos? ¿Podría prescindirse de ese dato o es fundamental?

Los atributos se han agrupado para facilitar la comprensión del encuestado.

Datos personales de carácter general:

	Nunca o casi nunca	A veces	Normalmente	Con frecuencia	Siempre
Nombre					
Primer apellido					
Segundo apellido					
Persona de contacto del paciente					
Estado civil					
Género					
Nacionalidad					
Fecha de nacimiento					
Elemento indicativo de edad					
Edad					
Sexo					

Datos personales de contacto:

	Nunca o casi nunca	A veces	Normalmente	Con frecuencia	Siempre
Teléfono					
Fax					
Correo electrónico					

Datos personales asociados a geolocalización:

	Nunca o casi nunca	A veces	Normalmente	Con frecuencia	Siempre
Calle					
Ciudad					
Municipio					
Distrito					
Dirección					
Comunidad autónoma					

	Nunca o casi nunca	A veces	Normalmente	Con frecuencia	Siempre
Estado					
País de nacimiento					
Código postal					

CUESTIONARIO SOBRE EL GRADO DE IMPORTANCIA (2/4)

¿Con qué frecuencia usa estos datos en su trabajo o investigaciones?

Datos personales asociados a donaciones:

	Nunca o casi nunca	A veces	Normalmente	Con frecuencia	Siempre
Origen étnico					
Necesidades personales					
Orientación sexual					
Origen social					
Pluralidad de nacimiento					

Datos personales relacionados con la profesión:

	Nunca o casi nunca	A veces	Normalmente	Con frecuencia	Siempre
Años de escolaridad					
Historial de educación					
Historial de empleo					
Posición dentro de la familia (padre, hermano mayor, etc)					
Nombre del último empleo					
Organización en					

	Nunca o casi nunca	A veces	Normalmente	Con frecuencia	Siempre
la que trabaja					
Profesión					

CUESTIONARIO SOBRE EL GRADO DE IMPORTANCIA (3/4)

¿Con qué frecuencia usa esos datos en su trabajo o investigaciones?

Datos personales asociados a donaciones:

	Nunca o casi nunca	A veces	Normalmente	Con frecuencia	Siempre
Donación de sustancia corporal					
Donación de parte corporal					

Datos personales asociados a aseguradoras:

	Nunca o casi nunca	A veces	Normalmente	Con frecuencia	Siempre
Ingresos totales					
Proveedor de atención médica					
Número de la Seguridad Social					
Número de tarjeta sanitaria					

Dispositivos implantados:

	Nunca o casi nunca	A veces	Normalmente	Con frecuencia	Siempre
Uso de aparatos auditivos					
Uso de aparatos oftalmológicos					

	Nunca o casi nunca	A veces	Normalmente	Con frecuencia	Siempre
Uso de aparatos ortopédicos					
Prótesis					
Número identificativo de los dispositivos implantados					

CUESTIONARIO SOBRE EL GRADO DE IMPORTANCIA (4/4)

¿Con qué frecuencia usa esos datos en su trabajo o investigaciones?

Datos médicos:

	Nunca o casi nunca	A veces	Normalmente	Con frecuencia	Siempre
Tipo sanguíneo					
Servicio prestado					
Tratamiento					
Ácido desoxirribonucleico					
Alergias					
Centros de salud					
Consumo de sustancias tóxicas					
Diagnosís					
Enfermedades					
Estado de embarazo					
Estado de enfermedad					
Historial clínico					
Información psicológica					
Información psiquiátrica					
Información recolectada durante un servicio sanitario					
Intervenciones					

	Nunca o casi nunca	A veces	Normalmente	Con frecuencia	Siempre
quirúrgicas					
Laboratorios que analizaron las pruebas médicas					
Medidas corporales (talla, peso)					
Prognosis					
Residencias de ancianos					
Vacunas					
Incapacidades médicas					
Imágenes fotográficas identificables					
Estado de discapacidad					
Genética de familiar cercano					
Genética propia					
Número de historia clínica					
Número de identificación					
Número de placa médica					
Número de registro médico					
Número de serie de dispositivos identificadores					
Códigos de diagnóstico					
Códigos de eventos adversos					
Códigos de tratamiento					
Fecha de admisión					
Fecha de alta					
Fecha de deceso					
Fecha de recogida de muestras					
Fecha de tratamiento					
Fecha de visita					
Número de registros de ensayos clínicos					

	Nunca o casi nunca	A veces	Normalmente	Con frecuencia	Siempre
Salud física (otra información física no recogida en anteriores campos)					
Salud mental (otra información mental no recogida en anteriores campos)					
Términos médicos que identifican patologías no habituales					

PARA TERMINAR...

Si cree que hay algún atributo clínico que sería importante incluir y que no aparece aquí, por favor, indíquelo a continuación:

¿Cree que es importante para su trabajo de investigador y/o docente utilizar el sistema que mejor asegure la confidencialidad de los datos de sus pacientes?

- No es importante
- Poco importante
- Moderadamente importante
- Importante
- Muy importante

ANEXO 5 Histogramas de las variables estadísticas por categorías que no se han eliminado en la fase de análisis previo

(1) Datos personales de carácter general

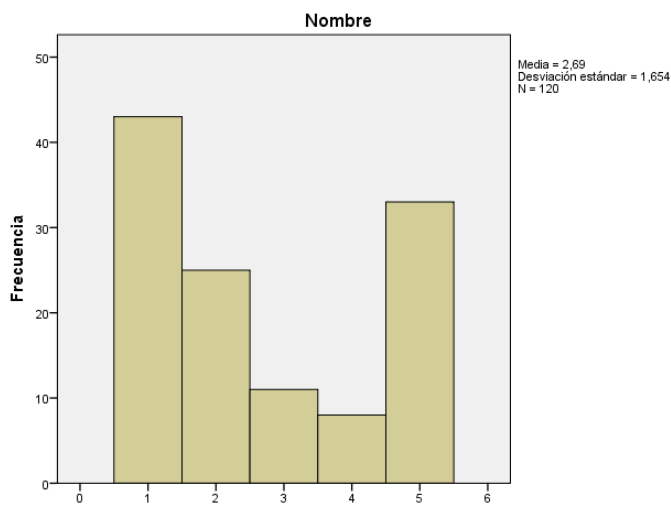


Figura 50 Histograma del atributo DPCG1 (Nombre)

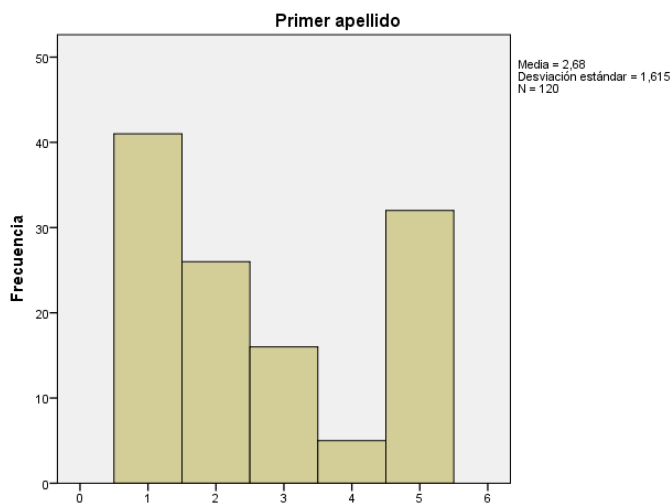


Figura 51 Histograma del atributo DPCG2 (Primer apellido)

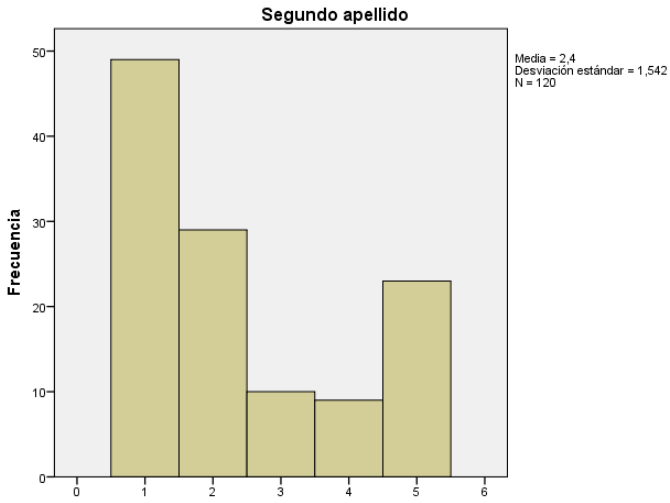


Figura 52 Histograma del atributo DPCG3 (Segundo apellido)

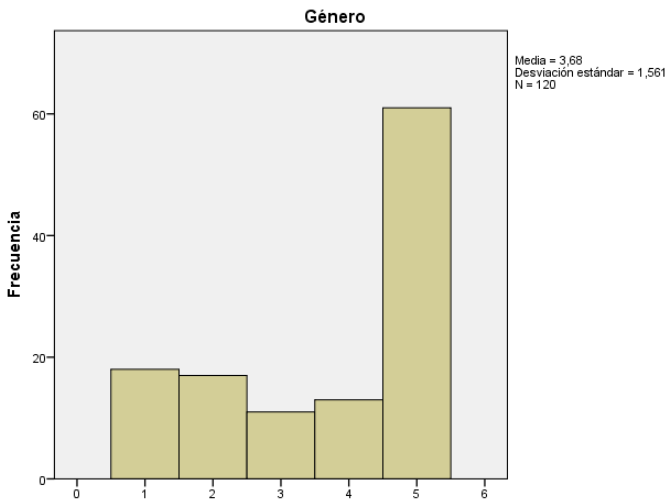


Figura 53 Histograma del atributo DPCG6 (Género)

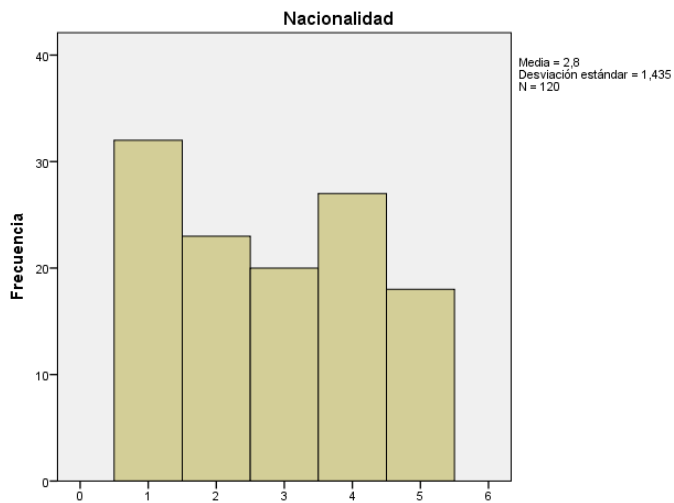


Figura 54 Histograma del atributo DPCG7 (Nacionalidad)

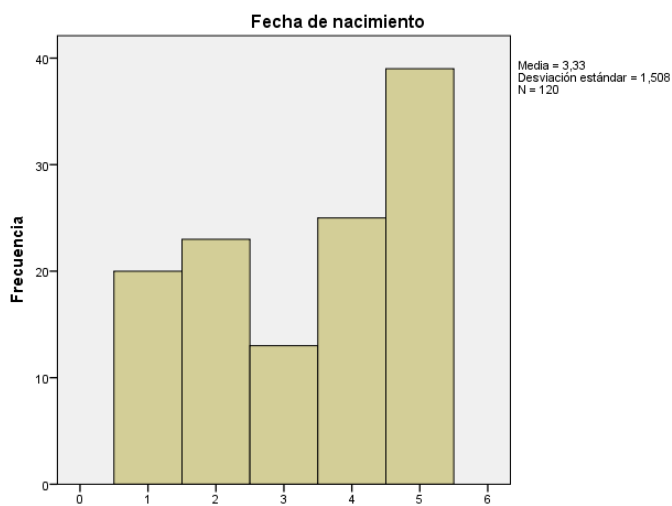


Figura 55 Histograma del atributo DPCG8 (Fecha de nacimiento)

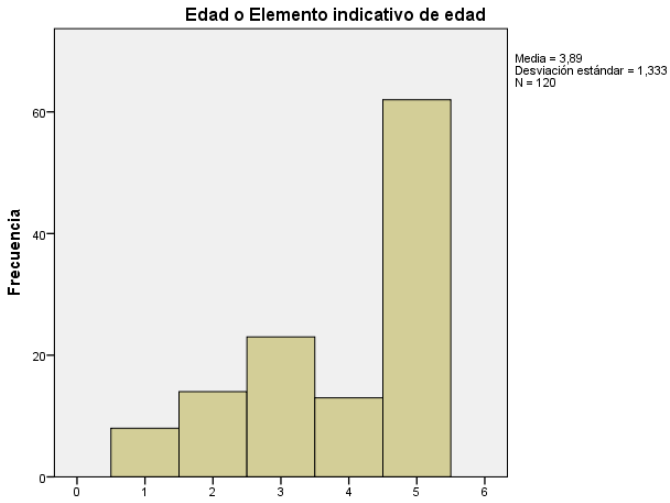


Figura 56 Histograma del atributo DPCG9 (Edad o elemento indicativo de edad)

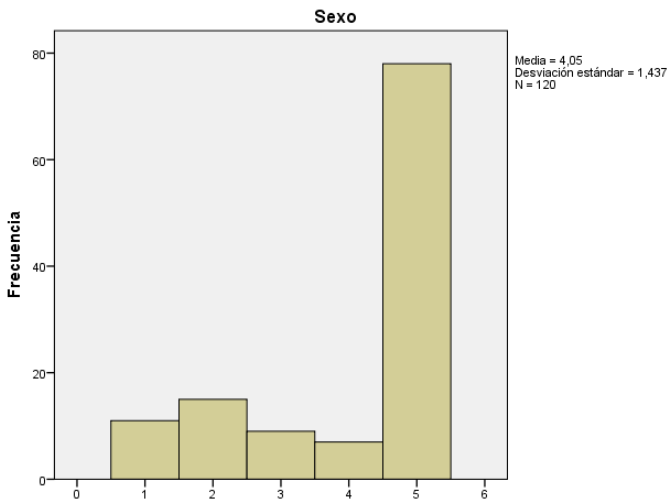


Figura 57 Histograma del atributo DPCG10 (Sexo)

(2) Datos personales de contacto

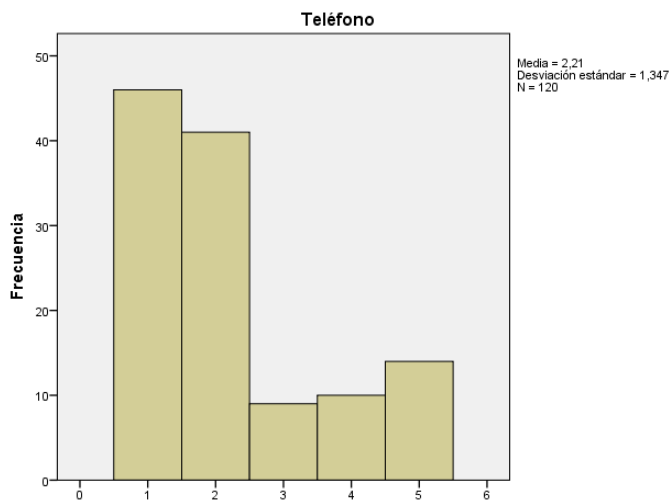


Figura 58 Histograma del atributo DPC1 (Teléfono)

(3) Datos personales asociados a geolocalización

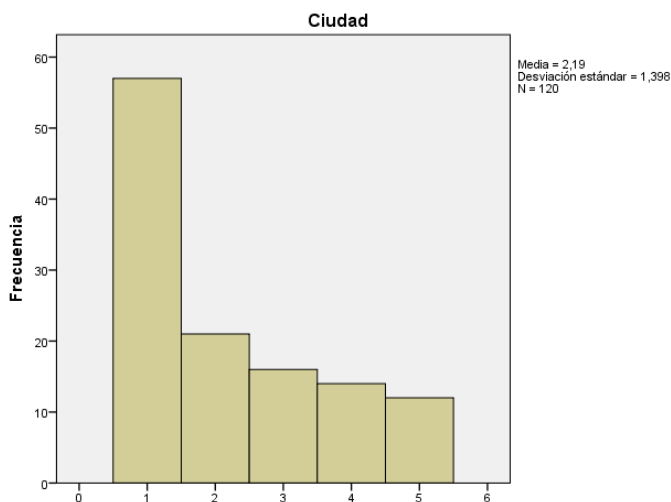


Figura 59 Histograma del atributo DPAG2 (Ciudad)

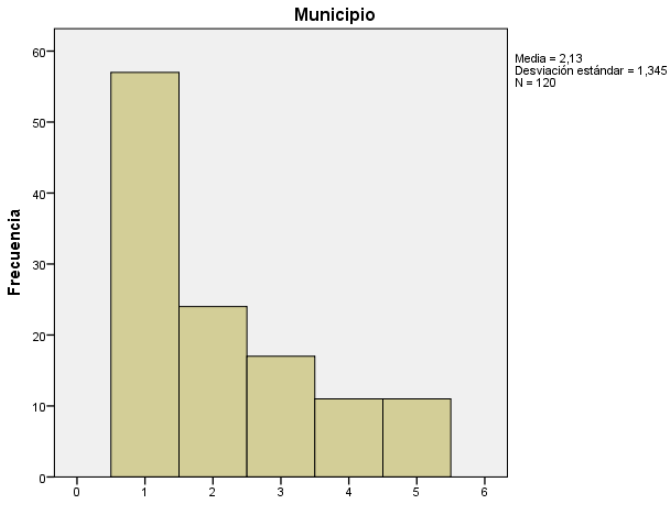


Figura 60 Histograma del atributo DPAG3 (Municipio)

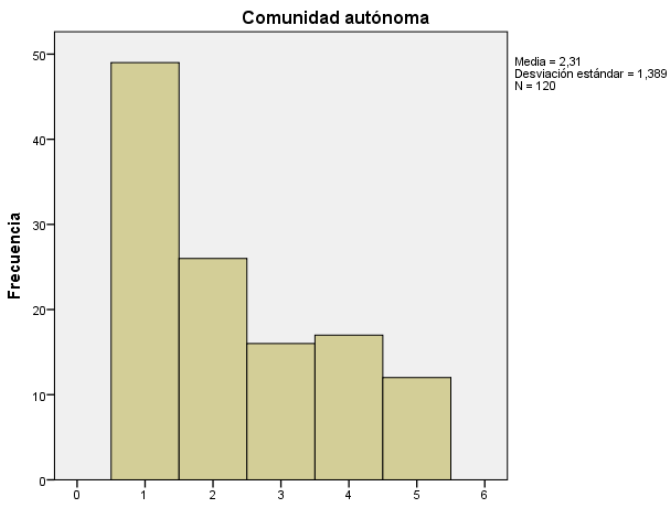


Figura 61 Histograma del atributo DPAG6 (Comunidad autónoma)

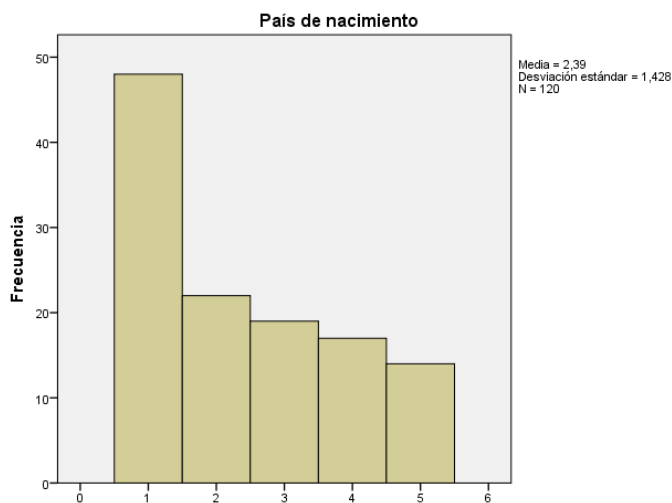


Figura 62 Histograma del atributo DPAG7 (País de nacimiento)

(4) Datos personales relacionados con creencias, cultura, raza, etnia u opiniones

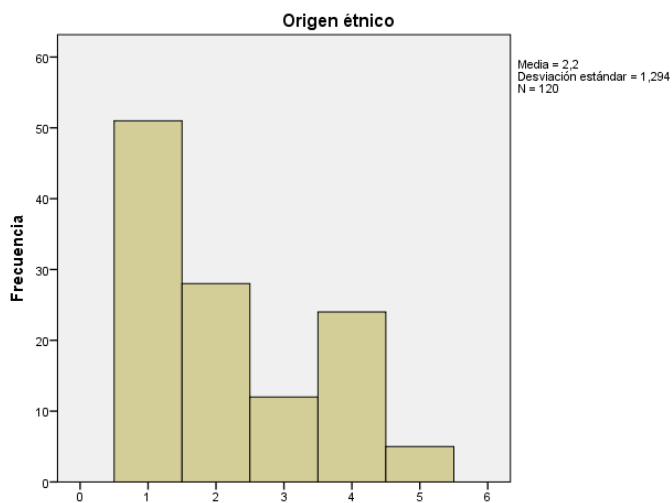


Figura 63 Histograma del atributo DPRCCRE01 (Origen étnico)

(5) Datos personales relacionados con la profesión

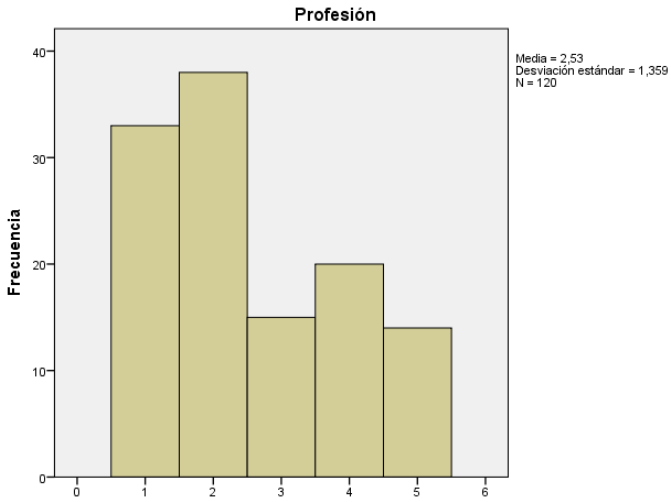


Figura 64 Histograma del atributo DPRP7 (Profesión)

(6) Datos de dispositivos implantados

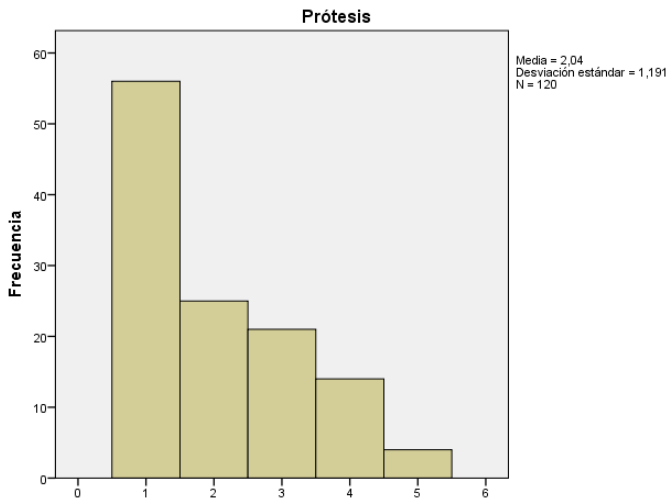
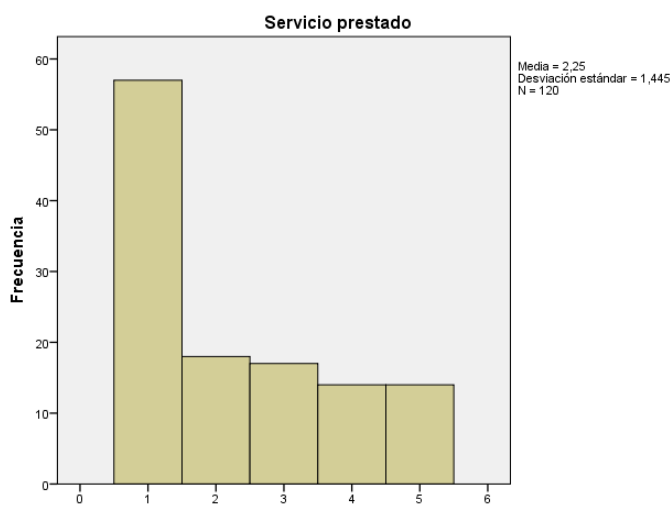
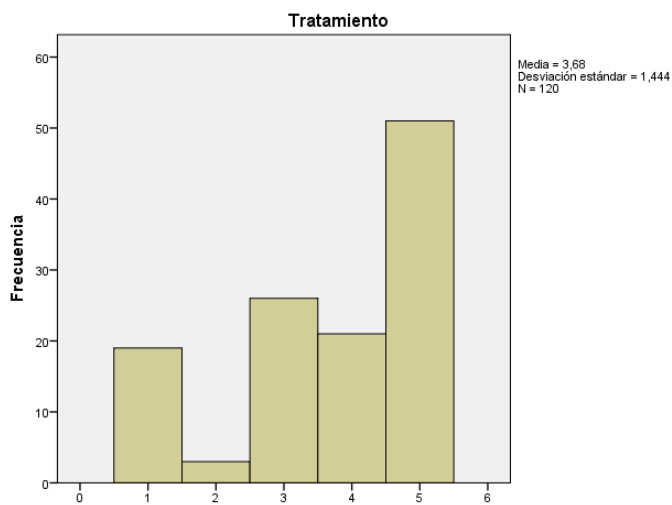


Figura 65 Histograma del atributo DDI4 (Prótesis)

(7) Datos médicos

**Figura 66 Histograma del atributo DM2 (Servicio prestado)****Figura 67 Histograma del atributo DM3 (Tratamiento)**

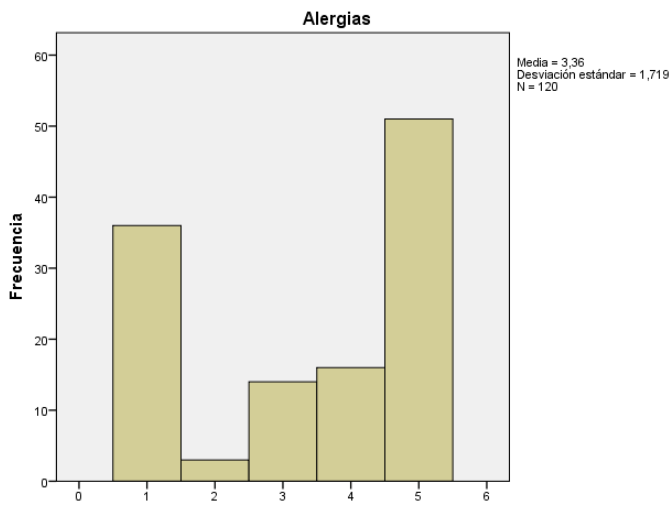


Figura 68 Histograma del atributo DM5 (Alergias)

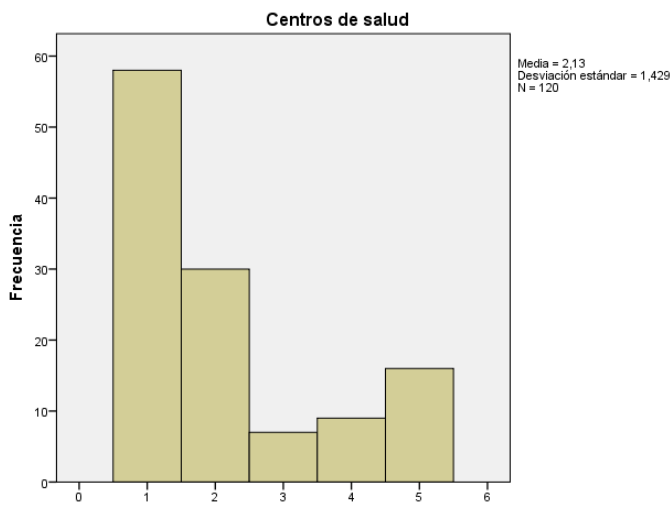


Figura 69 Histograma del atributo DM6 (Centro de salud)

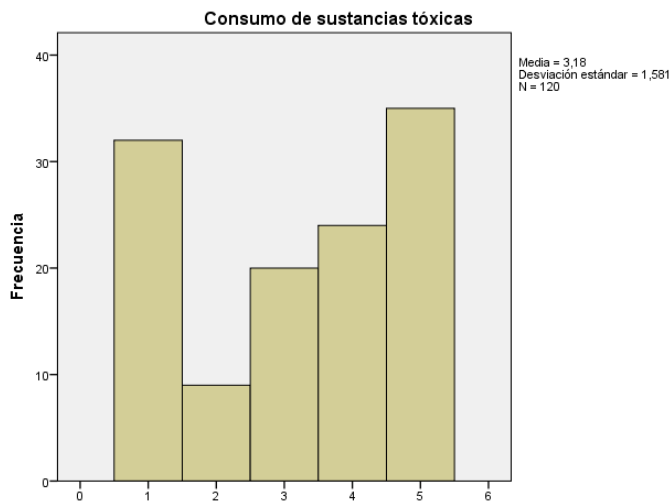


Figura 70 Histograma del atributo DM7 (Consumo de sustancias tóxicas)

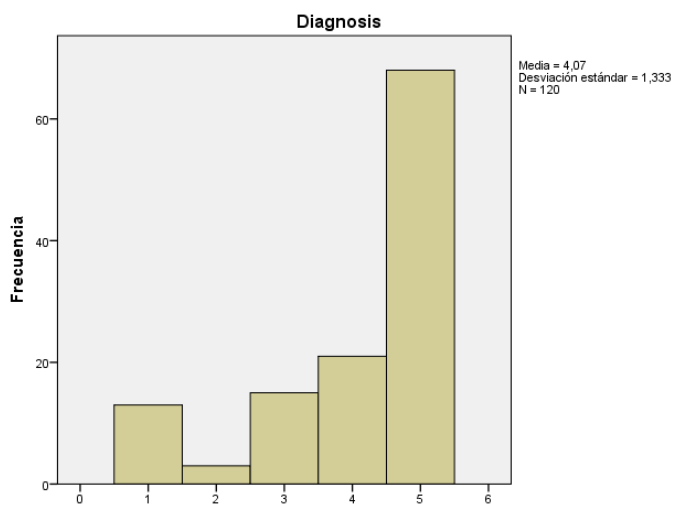


Figura 71 Histograma del atributo DM8 (Diagnosis)

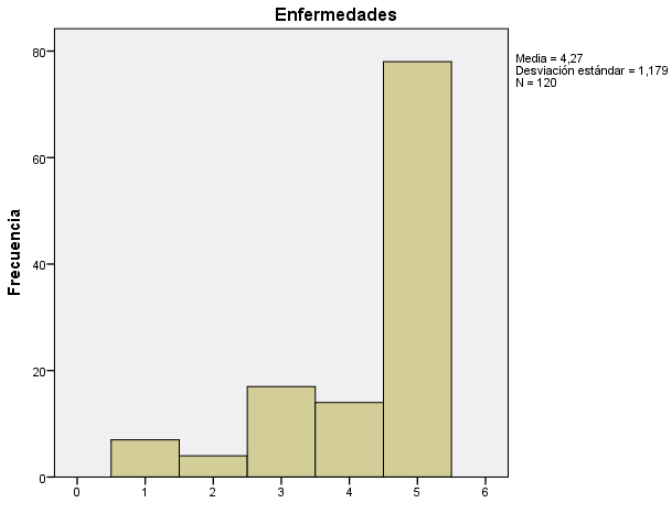


Figura 72 Histograma del atributo DM9 (Enfermedades)

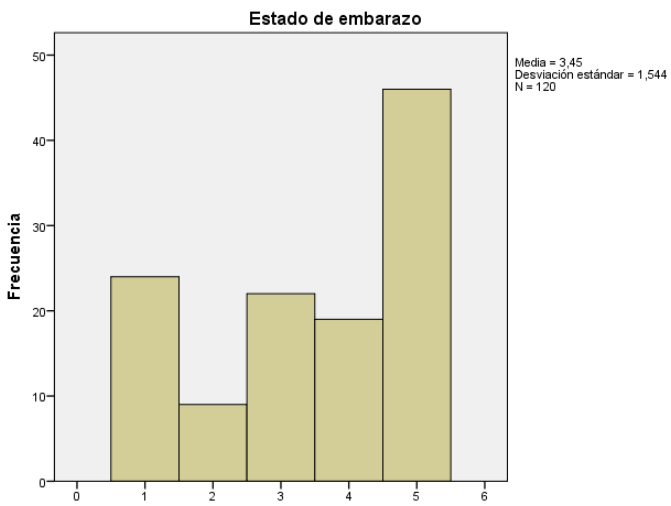


Figura 73 Histograma del atributo DM10 (Estado de embarazo)

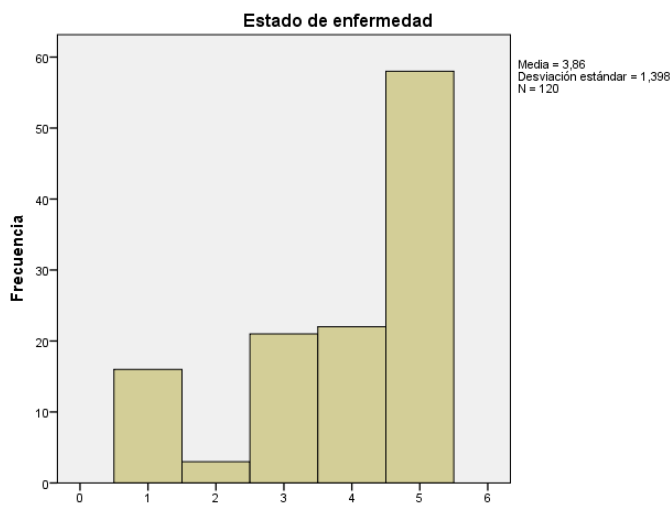


Figura 74 Histograma del atributo DM11 (Estado de enfermedad)

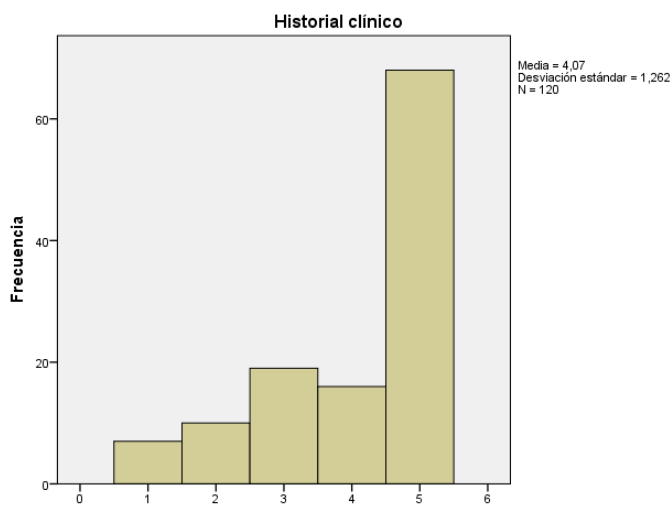


Figura 75 Histograma del atributo DM12 (Historial clínico)

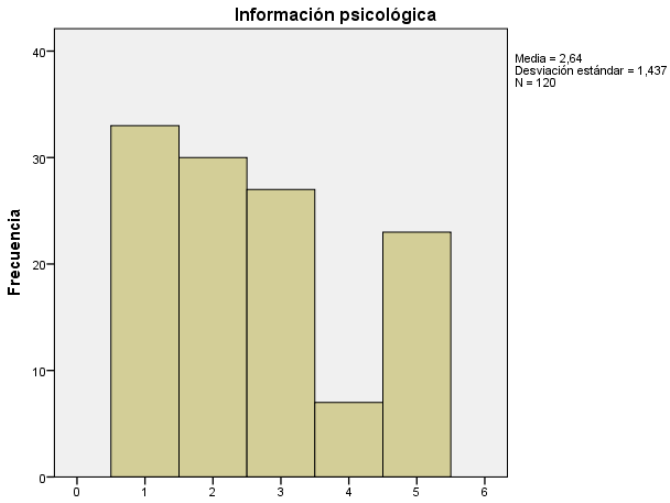


Figura 76 Histograma del atributo DM13 (Información psicológica)

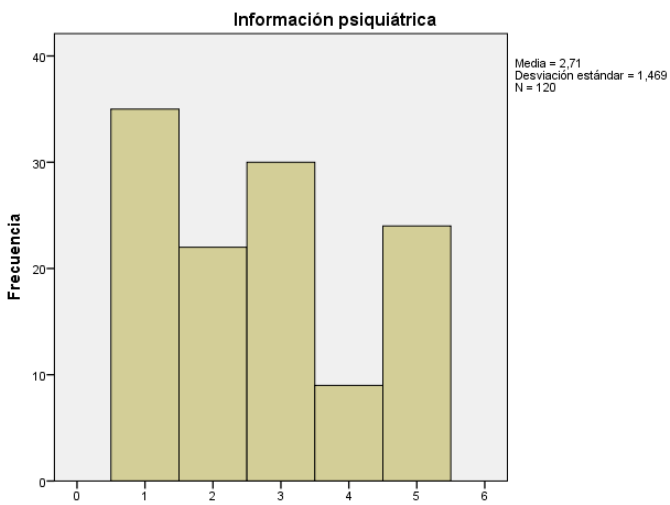


Figura 77 Histograma del atributo DM14 (Información psiquiátrica)

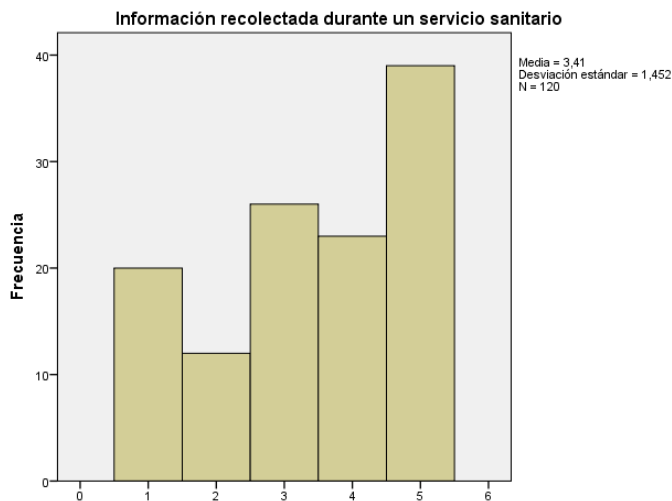


Figura 78 Histograma del atributo DM15 (Información recolectada durante un servicio sanitario)

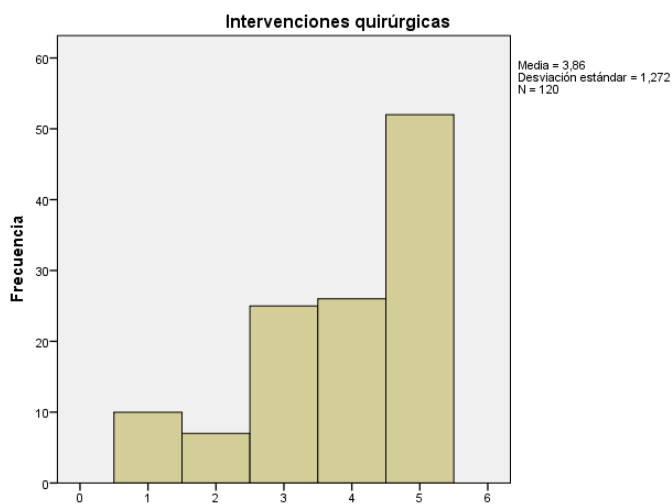


Figura 79 Histograma del atributo DM16 (Intervenciones quirúrgicas)

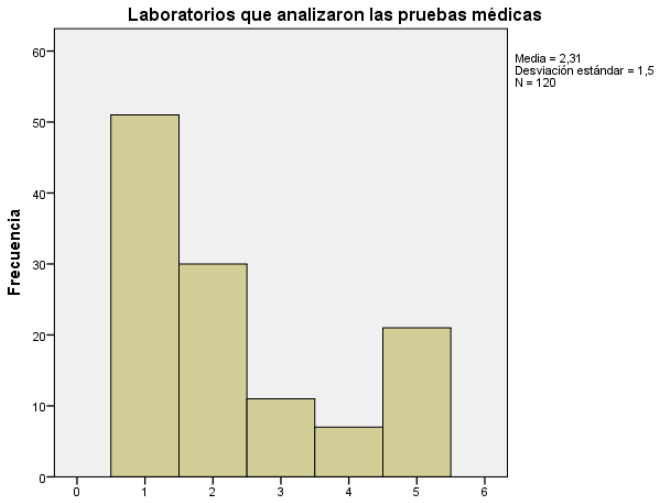


Figura 80 Histograma del atributo DM17 (Laboratorios que analizaron las pruebas médicas)

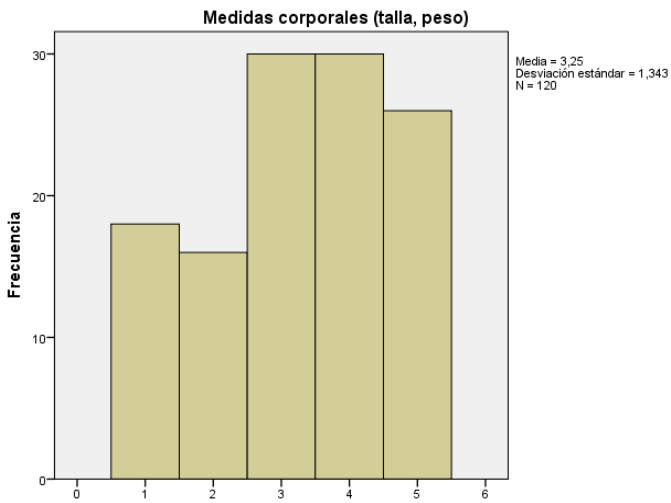


Figura 81 Histograma del atributo DM18 (Medidas corporales)

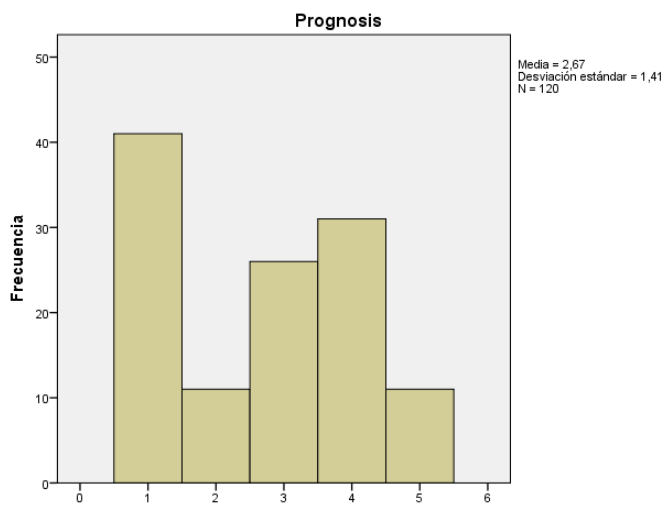


Figura 82 Histograma del atributo DM19 (Prognosis)

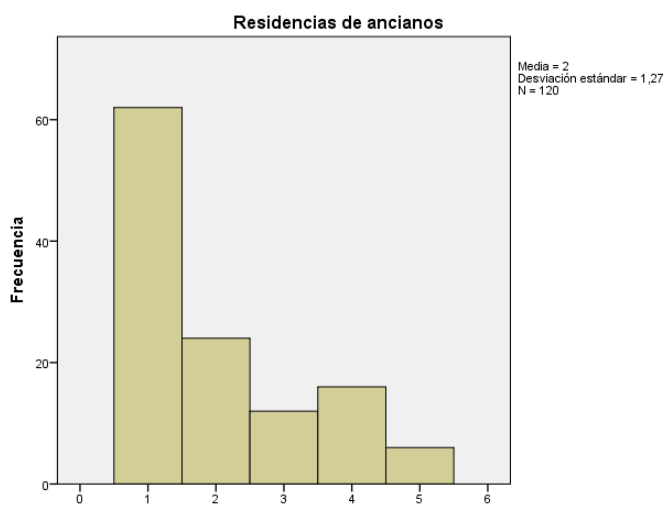


Figura 83 Histograma del atributo DM20 (Residencias de ancianos)

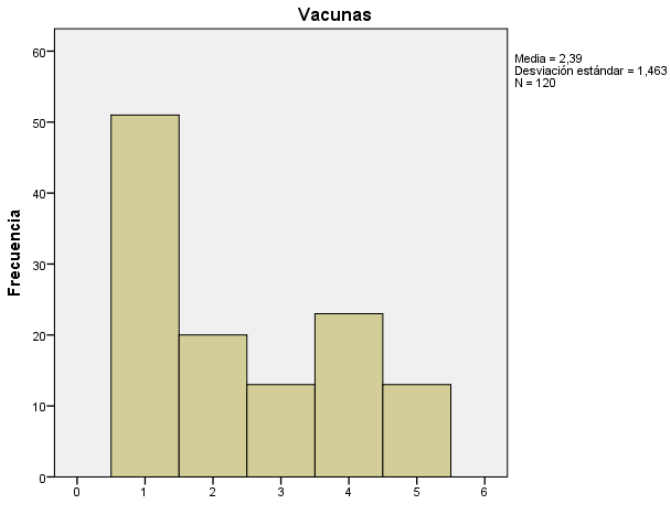


Figura 84 Histograma del atributo DM21 (Vacunas)

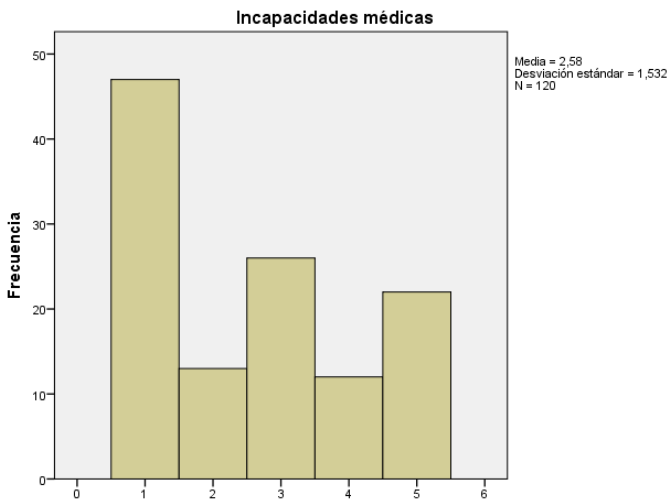


Figura 85 Histograma del atributo DM22 (Incapacidades médicas)

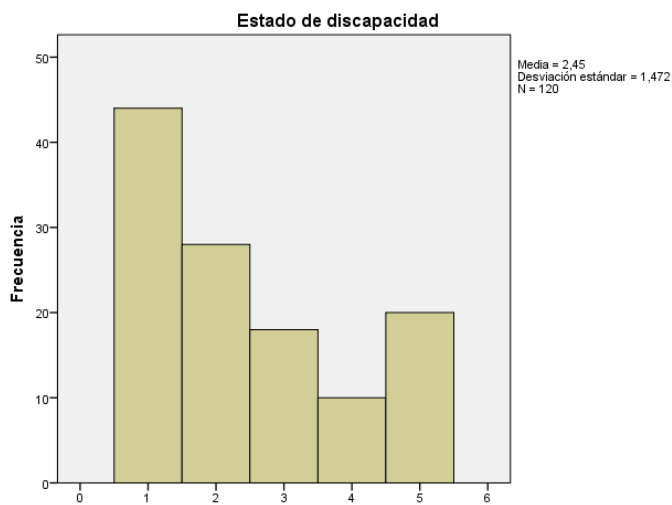


Figura 86 Histograma del atributo DM24 (Estado de discapacidad)

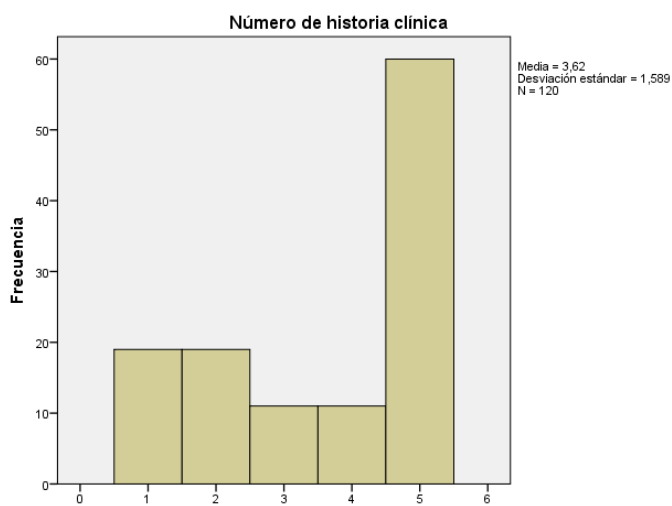


Figura 87 Histograma del atributo DM27 (Número de historia clínica)

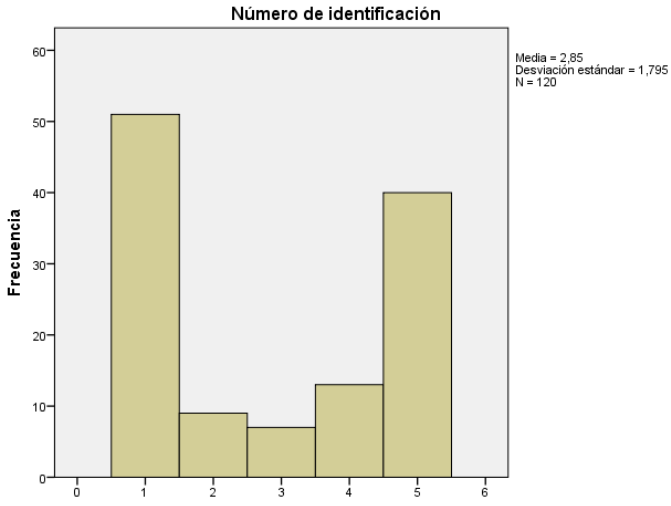


Figura 88 Histograma del atributo DM28 (Número de identificación)

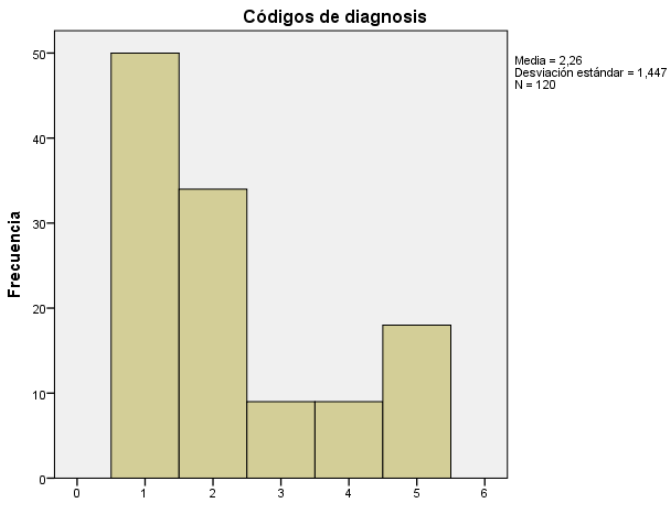


Figura 89 Histograma del atributo DM32 (Códigos de diagnóstico)

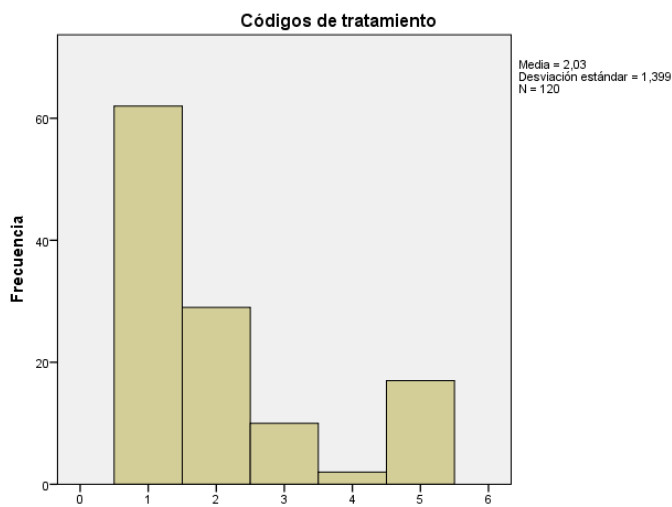


Figura 90 Histograma del atributo DM34 (Códigos de tratamiento)

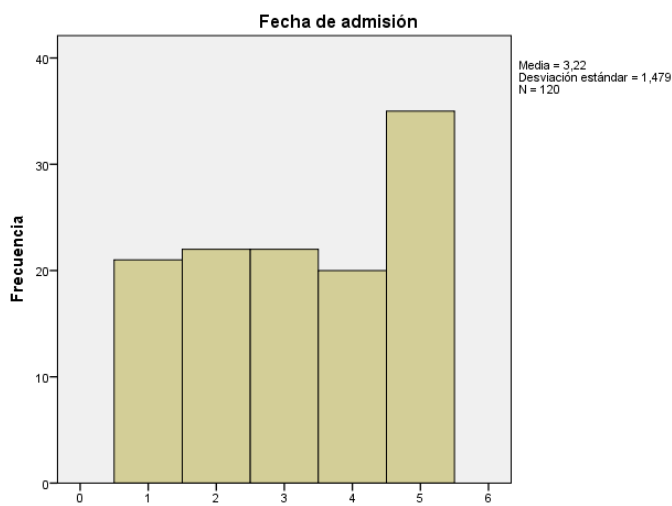


Figura 91 Histograma del atributo DM35 (Fecha de admisión)

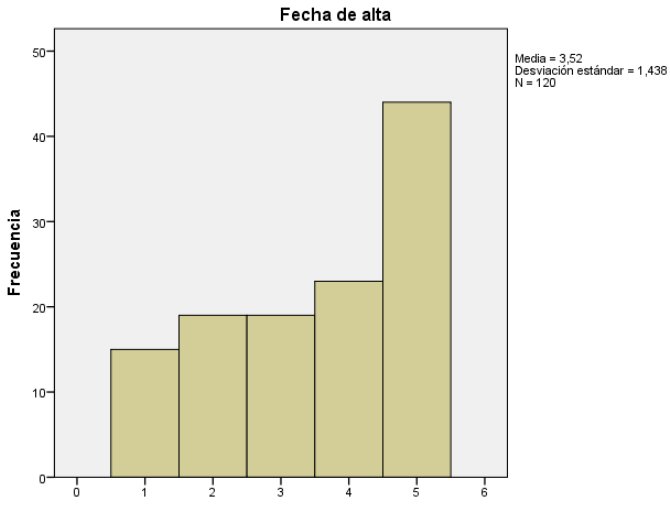


Figura 92 Histograma del atributo DM36 (Fecha de alta)

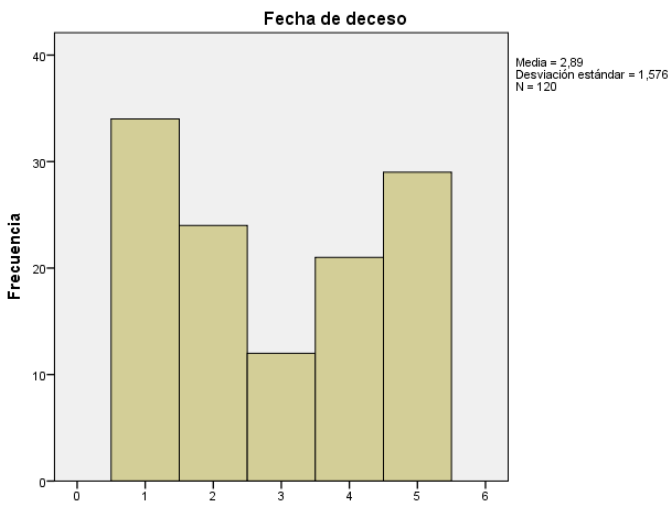


Figura 93 Histograma del atributo DM37 (Fecha de deceso)

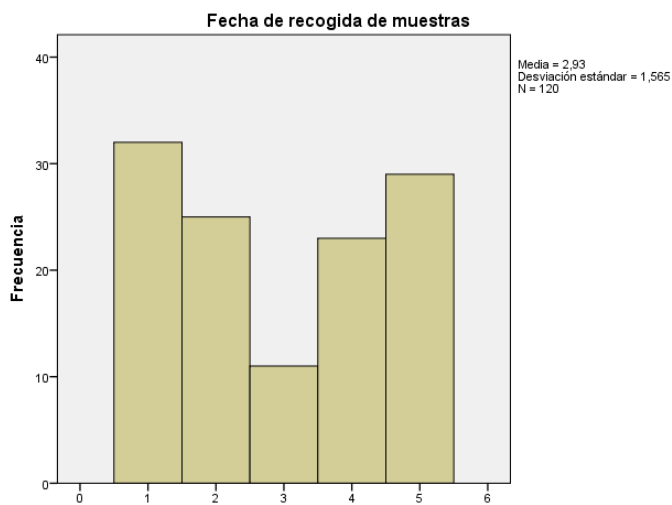


Figura 94 Histograma del atributo DM38 (Fecha de recogida de muestras)

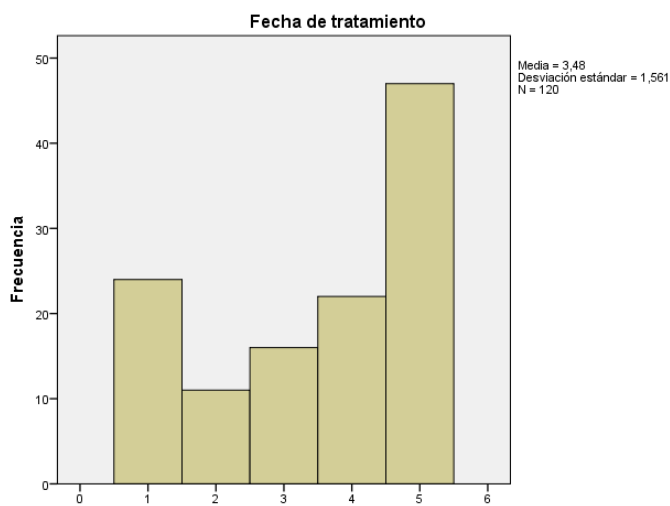


Figura 95 Histograma del atributo DM39 (Fecha de tratamiento)

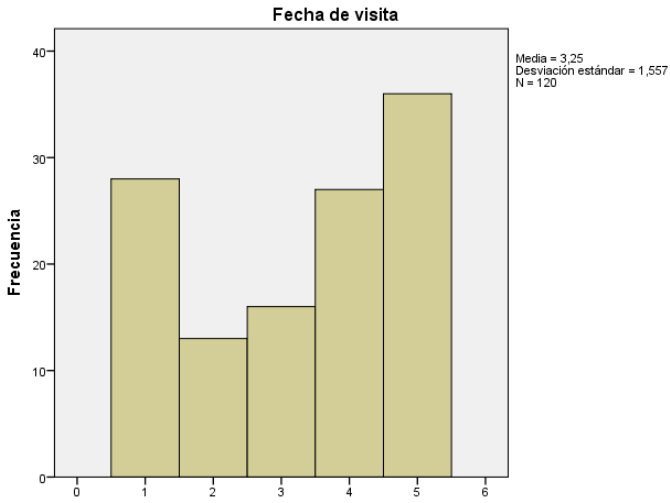


Figura 96 Histograma del atributo DM40 (Fecha de visita)

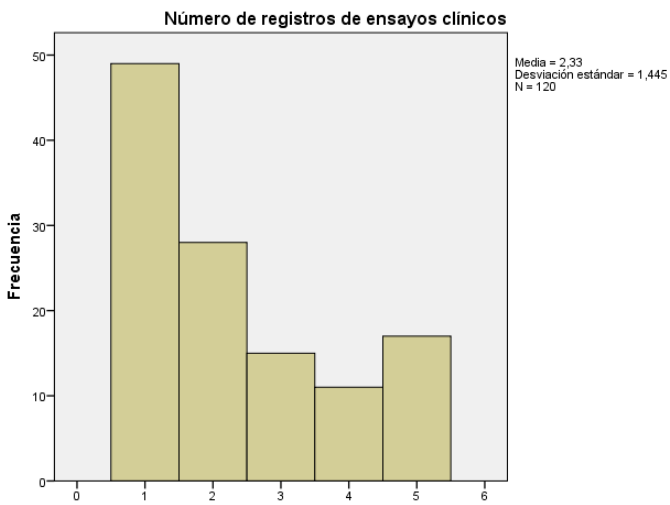


Figura 97 Histograma del atributo DM41 (Número de registros de ensayos clínicos)

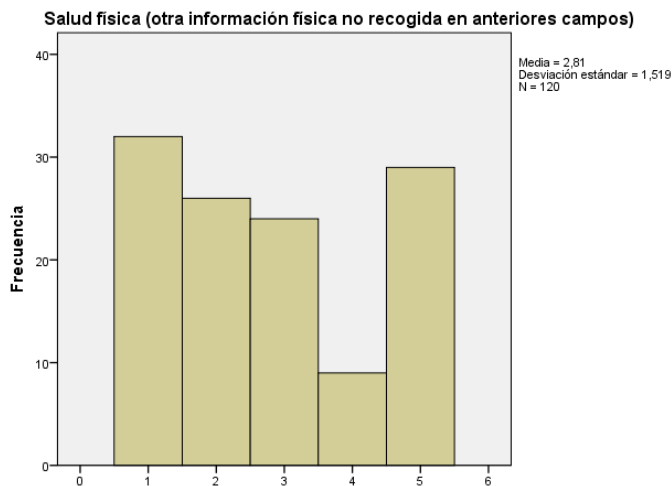


Figura 98 Histograma del atributo DM42 (Salud física)

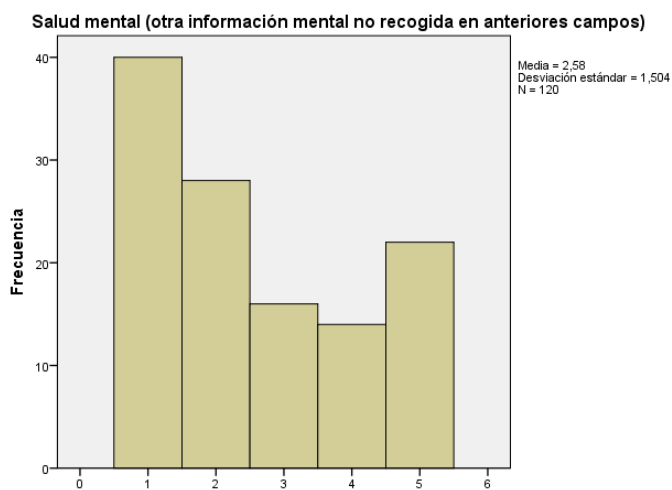


Figura 99 Histograma del atributo DM43 (Salud mental)

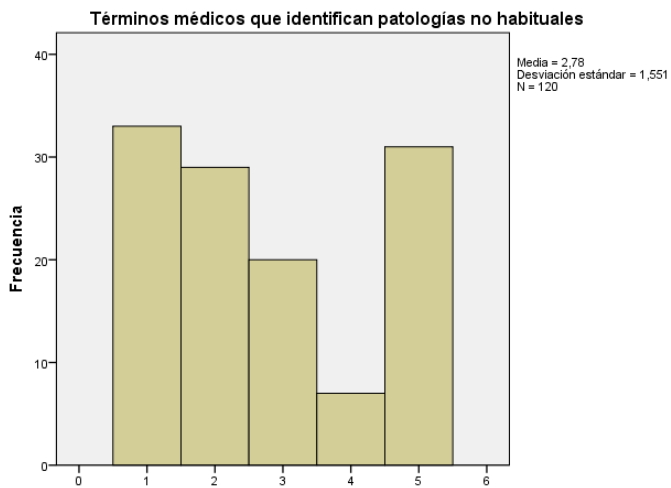


Figura 100 Histograma del atributo DM44 (Términos médicos que identifican patologías no habituales)

ANEXO 6 Matrices de correlación por categorías de datos

(1) Datos personales

A continuación, se recogen las matrices de correlación obtenidas con el programa SAS para la categoría “datos personales” y utilizada en la fase del análisis factorial confirmatorio. Debido al tamaño de su representación, se ha recogido en las tablas: Tabla 58 y Tabla 59.

Pearson Correlation Coefficients, N = 120							
Prob > r under H0: Rho=0							
	DPCG1	DPCG2	DPCG3	DPC1	DPCG6	DPCG7	DPCG8
DPCG1	1.00000	0.98461	0.86936	0.57989	0.01394	0.10122	0.05502
Nombre		<.0001	<.0001	<.0001	0.8799	0.2713	0.5506
DPCG2	0.98461	1.00000	0.91006	0.60718	-0.02450	0.06598	0.02415
Primer apellido	<.0001		<.0001	<.0001	0.7905	0.4740	0.7934
DPCG3	0.86936	0.91006	1.00000	0.67197	-0.09010	-0.03190	0.01807
Segundo apellido	<.0001	<.0001		<.0001	0.3278	0.7294	0.8447
DPC1	0.57989	0.60718	0.67197	1.00000	0.01166	-0.04782	0.01931
Teléfono	<.0001	<.0001	<.0001		0.8994	0.6040	0.8342
DPCG6	0.01394	-0.02450	-0.09010	0.01166	1.00000	0.58655	0.55922
Género	0.8799	0.7905	0.3278	0.8994		<.0001	<.0001
DPCG7	0.10122	0.06598	-0.03190	-0.04782	0.58655	1.00000	0.70647
Nacionalidad	0.2713	0.4740	0.7294	0.6040	<.0001		<.0001
DPCG8	0.05502	0.02415	0.01807	0.01931	0.55922	0.70647	1.00000
Fecha de nacimiento	0.5506	0.7934	0.8447	0.8342	<.0001	<.0001	
DPCG9	0.18177	0.16032	0.12119	0.07292	0.48187	0.42858	0.56834
Edad o elemento indicativo de edad	0.0469	0.0803	0.1873	0.4287	<.0001	<.0001	<.0001
DPCG10	-0.01821	-0.05450	-0.06601	0.10749	0.62154	0.43267	0.51574
Sexo	0.8435	0.5544	0.4738	0.2426	<.0001	<.0001	<.0001
DPAG2	0.51997	0.53776	0.59188	0.58120	0.16667	0.28307	0.26836
Ciudad	<.0001	<.0001	<.0001	<.0001	0.0689	0.0017	0.0030
DPAG3	0.55780	0.57619	0.64463	0.62135	0.14314	0.22641	0.27351
Municipio	<.0001	<.0001	<.0001	<.0001	0.1188	0.0129	0.0025
DPAG6	0.39650	0.41594	0.49135	0.53145	0.25082	0.36836	0.37173
Comunidad autónoma	<.0001	<.0001	<.0001	<.0001	0.0057	<.0001	<.0001

Pearson Correlation Coefficients, N = 120							
Prob > r under H0: Rho=0							
	DPCG1	DPCG2	DPCG3	DPC1	DPCG6	DPCG7	DPCG8
DPAG8	0.01242	0.03380	0.12294	0.07521	0.45201	0.69457	0.60229
País de nacimiento	0.8929	0.7140	0.1810	0.4143	<.0001	<.0001	<.0001
DPRCCREO1	0.05654	0.03539	-0.02359	0.01929	0.51844	0.61446	0.50818
Origen étnico	0.5396	0.7011	0.7981	0.8343	<.0001	<.0001	<.0001
DPRP7	0.33164	0.32082	0.28233	0.27852	0.16740	0.37383	0.16260
Profesión	0.0002	0.0004	0.0018	0.0021	0.0676	<.0001	0.0760

Tabla 58 Matriz de correlación para la categoría "datos personales" parte 1 de 2

Pearson Correlation Coefficients, N = 120								
Prob > r under H0: Rho=0								
	DPCG9	DPCG10	DPAG2	DPAG3	DPAG6	DPAG8	DPRCCREO1	DPRP7
DPCG1	0.18177	-0.01821	0.51997	0.55780	0.39650	0.01242	0.05654	0.33164
Nombre	0.0469	0.8435	<.0001	<.0001	<.0001	0.8929	0.5396	0.0002
DPCG2	0.16032	-0.05450	0.53776	0.57619	0.41594	0.03380	0.03539	0.32082
Primer apellido	0.0803	0.5544	<.0001	<.0001	<.0001	0.7140	0.7011	0.0004
DPCG3	0.12119	-0.06601	0.59188	0.64463	0.49135	0.12294	-0.02359	0.28233
Segundo apellido	0.1873	0.4738	<.0001	<.0001	<.0001	0.1810	0.7981	0.0018
DPC1	0.07292	0.10749	0.58120	0.62135	0.53145	0.07521	0.01929	0.27852
Teléfono	0.4287	0.2426	<.0001	<.0001	<.0001	0.4143	0.8343	0.0021
DPCG6	0.48187	0.62154	0.16667	0.14314	0.25082	0.45201	0.51844	0.16740
Género	<.0001	<.0001	0.0689	0.1188	0.0057	<.0001	<.0001	0.0676
DPCG7	0.42858	0.43267	0.28307	0.22641	0.36836	0.69457	0.61446	0.37383
Nacionalidad	<.0001	<.0001	0.0017	0.0129	<.0001	<.0001	<.0001	<.0001
DPCG8	0.56834	0.51574	0.26836	0.27351	0.37173	0.60229	0.50818	0.16260
Fecha de nacimiento	<.0001	<.0001	0.0030	0.0025	<.0001	<.0001	<.0001	0.0760

Pearson Correlation Coefficients, N = 120								
Prob > r under H0: Rho=0								
	DPCG9	DPCG10	DPAG2	DPAG3	DPAG6	DPAG8	DPRCCREO1	DPRP7
DPCG9	1.00000	0.46831	0.26690	0.29296	0.25118	0.28917	0.34605	0.02791
Edad o elemento indicativo de edad		<.0001	0.0032	0.0012	0.0057	0.0014	0.0001	0.7622
DPCG10	0.46831	1.00000	0.03284	0.02719	0.12273	0.33853	0.36975	0.07228
Sexo		<.0001	0.7218	0.7682	0.1817	0.0002	<.0001	0.4327
DPAG2	0.26690	0.03284	1.00000	0.95278	0.92566	0.37464	0.16446	0.29952
Ciudad		0.0032	0.7218	<.0001	<.0001	<.0001	0.0727	0.0009
DPAG3	0.29296	0.02719	0.95278	1.00000	0.89711	0.34198	0.11594	0.25749
Municipio		0.0012	0.7682	<.0001	<.0001	0.0001	0.2073	0.0045
DPAG6	0.25118	0.12273	0.92566	0.89711	1.00000	0.51910	0.27869	0.37951
Comunidad autónoma		0.0057	0.1817	<.0001	<.0001	<.0001	0.0021	<.0001
DPAG8	0.28917	0.33853	0.37464	0.34198	0.51910	1.00000	0.62594	0.41972
País de nacimiento		0.0014	0.0002	<.0001	0.0001	<.0001	<.0001	<.0001
DPRCCREO1	0.34605	0.36975	0.16446	0.11594	0.27869	0.62594	1.00000	0.48834
Origen étnico		0.0001	<.0001	0.0727	0.2073	0.0021	<.0001	<.0001
DPRP7	0.02791	0.07228	0.29952	0.25749	0.37951	0.41972	0.48834	1.00000
Profesión		0.7622	0.4327	0.0009	0.0045	<.0001	<.0001	<.0001

Tabla 59 Matriz de correlación para la categoría "datos personales" parte 2 de 2

(2) Datos médicos

A continuación se recoge la matriz de correlación obtenida con el programa SAS para la categoría "datos médicos" y utilizada en la fase del análisis factorial confirmatorio. Debido al tamaño de su representación, se ha recogido en las tablas: Tabla 60, Tabla 61, Tabla 62, Tabla 63, Tabla 64 y Tabla 65.

Pearson Correlation Coefficients, N = 120						
Prob > r under H0: Rho=0						
	DM3	DM5	DM8	DM9	DM10	DM11
DM3 Tratamiento	1.00000	0.53720 <.0001	0.66611 <.0001	0.57827 <.0001	0.58477 <.0001	0.73959 <.0001
DM5 Alergias	0.53720 <.0001	1.00000	0.51759 <.0001	0.56192 <.0001	0.64482 <.0001	0.60878 <.0001
DM8 Diagnosis	0.66611 <.0001	0.51759 <.0001	1.00000	0.91344 <.0001	0.64670 <.0001	0.78968 <.0001
DM9 Enfermedades	0.57827 <.0001	0.56192 <.0001	0.91344 <.0001	1.00000	0.63509 <.0001	0.75713 <.0001
DM10 Estado de embarazo	0.58477 <.0001	0.64482 <.0001	0.64670 <.0001	0.63509 <.0001	1.00000	0.78890 <.0001
DM11 Estado de enfermedad	0.73959 <.0001	0.60878 <.0001	0.78968 <.0001	0.75713 <.0001	0.78890 <.0001	1.00000
DM12 Historial clínico	0.69911 <.0001	0.62818 <.0001	0.73673 <.0001	0.71652 <.0001	0.72204 <.0001	0.90096 <.0001
DM15 Información recolectada durante un servicio sanitario	0.56325 <.0001	0.32802 0.0003	0.61521 <.0001	0.52472 <.0001	0.61440 <.0001	0.72405 <.0001
DM16 Intervenciones quirúrgicas	0.68464 <.0001	0.70745 <.0001	0.66466 <.0001	0.65840 <.0001	0.80706 <.0001	0.80129 <.0001
DM27 Número de historia clínica	0.62084 <.0001	0.49384 <.0001	0.48037 <.0001	0.48115 <.0001	0.42718 <.0001	0.58825 <.0001
DM36 Fecha de alta	0.66258 <.0001	0.48553 <.0001	0.56066 <.0001	0.50296 <.0001	0.64393 <.0001	0.70984 <.0001
DM39 Fecha de tratamiento	0.67897 <.0001	0.50607 <.0001	0.55006 <.0001	0.51957 <.0001	0.60444 <.0001	0.67034 <.0001
DM40 Fecha de visita	0.62625 <.0001	0.48118 <.0001	0.52227 <.0001	0.49888 <.0001	0.54703 <.0001	0.63407 <.0001
DM6 Centros de salud	0.35331 <.0001	0.55623 <.0001	0.24697 0.0065	0.20940 0.0217	0.46929 <.0001	0.34956 <.0001

Pearson Correlation Coefficients, N = 120						
Prob > r under H0: Rho=0						
	DM3	DM5	DM8	DM9	DM10	DM11
DM7 Consumo de sustancias tóxicas	0.59902 <.0001	0.72830 <.0001	0.61656 <.0001	0.59241 <.0001	0.59756 <.0001	0.60833 <.0001
DM13 Información psicológica	0.30547 0.0007	0.39275 <.0001	0.41628 <.0001	0.36940 <.0001	0.45594 <.0001	0.41802 <.0001
DM14 Información psiquiátrica	0.29292 0.0012	0.49103 <.0001	0.43485 <.0001	0.38002 <.0001	0.53999 <.0001	0.47891 <.0001
DM22 Incapacidades médicas	0.42500 <.0001	0.52420 <.0001	0.35138 <.0001	0.26327 0.0037	0.49005 <.0001	0.48560 <.0001
DM24 Estado de discapacidad	0.40785 <.0001	0.43404 <.0001	0.34013 0.0001	0.24020 0.0082	0.42051 <.0001	0.39069 <.0001
DM43 Salud mental (otra información mental no recogida en anteriores campos)	0.44191 <.0001	0.52633 <.0001	0.42474 <.0001	0.36170 <.0001	0.46501 <.0001	0.45928 <.0001
DDI4 Prótesis	0.42810 <.0001	0.39903 <.0001	0.35287 <.0001	0.30916 0.0006	0.40557 <.0001	0.37705 <.0001
DM32 Códigos de diagnóstico	0.24476 0.0071	0.35789 <.0001	0.33528 0.0002	0.25980 0.0042	0.22970 0.0116	0.37561 <.0001
DM34 Códigos de tratamiento	0.27026 0.0028	0.31423 0.0005	0.21538 0.0182	0.13855 0.1313	0.27484 0.0024	0.31975 0.0004
DM35 Fecha de admisión	0.56371 <.0001	0.30302 0.0008	0.49976 <.0001	0.40020 <.0001	0.50518 <.0001	0.62857 <.0001
DM18 Medidas corporales (talla, peso)	0.53983 <.0001	0.41603 <.0001	0.55405 <.0001	0.49364 <.0001	0.44794 <.0001	0.53836 <.0001
DM19 Prognosis	0.43481 <.0001	0.45529 <.0001	0.46339 <.0001	0.38742 <.0001	0.45925 <.0001	0.54698 <.0001
DM20 Residencias de ancianos	0.23832 0.0088	0.26942 0.0029	0.30273 0.0008	0.26369 0.0036	0.29993 0.0009	0.34071 0.0001

Pearson Correlation Coefficients, N = 120 Prob > r under H0: Rho=0						
	DM3	DM5	DM8	DM9	DM10	DM11
DM37 Fecha de deceso	0.48713 <.0001	0.12922 0.1595	0.39944 <.0001	0.32315 0.0003	0.32408 0.0003	0.47733 <.0001
DM42 Salud física (otra información física no recogida en anteriores campos)	0.57376 <.0001	0.33227 0.0002	0.53755 <.0001	0.45569 <.0001	0.49569 <.0001	0.60043 <.0001
DM44 Términos médicos que identifican patologías no habituales	0.46446 <.0001	0.42960 <.0001	0.48656 <.0001	0.43152 <.0001	0.48311 <.0001	0.58631 <.0001
DM21 Vacunas	0.34977 <.0001	0.56540 <.0001	0.34421 0.0001	0.38230 <.0001	0.51293 <.0001	0.45063 <.0001
DM28 Número de identificación	0.44860 <.0001	0.42073 <.0001	0.41166 <.0001	0.39628 <.0001	0.45215 <.0001	0.51058 <.0001
DM41 Número de registros de ensayos clínicos	0.36813 <.0001	0.26406 0.0036	0.33340 0.0002	0.32362 0.0003	0.34454 0.0001	0.37251 <.0001
DM2 Servicio prestado	0.39680 <.0001	0.41359 <.0001	0.17449 0.0566	0.10849 0.2382	0.35214 <.0001	0.35460 <.0001
DM17 Laboratorios que analizaron las pruebas médicas	0.28229 0.0018	0.45559 <.0001	0.27549 0.0023	0.22399 0.0139	0.38962 <.0001	0.36974 <.0001
DM38 Fecha de recogida de muestras	0.54109 <.0001	0.50564 <.0001	0.45328 <.0001	0.39221 <.0001	0.50286 <.0001	0.57559 <.0001

Tabla 60 Matriz de correlación para la categoría "datos médicos" parte 1 de 6

Pearson Correlation Coefficients, N = 120						
Prob > r under H0: Rho=0						
	DM12	DM15	DM16	DM27	DM36	DM39
DM3 Tratamiento	0.69911 <.0001	0.56325 <.0001	0.68464 <.0001	0.62084 <.0001	0.66258 <.0001	0.67897 <.0001
DM5 Alergias	0.62818 <.0001	0.32802 0.0003	0.70745 <.0001	0.49384 <.0001	0.48553 <.0001	0.50607 <.0001
DM8 Diagnosis	0.73673 <.0001	0.61521 <.0001	0.66466 <.0001	0.48037 <.0001	0.56066 <.0001	0.55006 <.0001
DM9 Enfermedades	0.71652 <.0001	0.52472 <.0001	0.65840 <.0001	0.48115 <.0001	0.50296 <.0001	0.51957 <.0001
DM10 Estado de embarazo	0.72204 <.0001	0.61440 <.0001	0.80706 <.0001	0.42718 <.0001	0.64393 <.0001	0.60444 <.0001
DM11 Estado de enfermedad	0.90096 <.0001	0.72405 <.0001	0.80129 <.0001	0.58825 <.0001	0.70984 <.0001	0.67034 <.0001
DM12 Historial clínico	1.00000	0.72790 <.0001	0.81212 <.0001	0.66679 <.0001	0.74981 <.0001	0.67499 <.0001
DM15 Información recolectada durante un servicio sanitario	0.72790 <.0001	1.00000	0.65015 <.0001	0.39255 <.0001	0.56218 <.0001	0.50314 <.0001
DM16 Intervenciones quirúrgicas	0.81212 <.0001	0.65015 <.0001	1.00000	0.62983 <.0001	0.73873 <.0001	0.71975 <.0001
DM27 Número de historia clínica	0.66679 <.0001	0.39255 <.0001	0.62983 <.0001	1.00000	0.63562 <.0001	0.58235 <.0001
DM36 Fecha de alta	0.74981 <.0001	0.56218 <.0001	0.73873 <.0001	0.63562 <.0001	1.00000	0.73977 <.0001
DM39 Fecha de tratamiento	0.67499 <.0001	0.50314 <.0001	0.71975 <.0001	0.58235 <.0001	0.73977 <.0001	1.00000
DM40 Fecha de visita	0.59881 <.0001	0.40414 <.0001	0.65014 <.0001	0.49087 <.0001	0.67383 <.0001	0.89813 <.0001
DM6 Centros de salud	0.37740 <.0001	0.33144 0.0002	0.36104 <.0001	0.31362 0.0005	0.36497 <.0001	0.17656 0.0537

Pearson Correlation Coefficients, N = 120						
Prob > r under H0: Rho=0						
	DM12	DM15	DM16	DM27	DM36	DM39
DM7	0.61345	0.47012	0.55569	0.33814	0.45908	0.48033
Consumo de sustancias tóxicas	<.0001	<.0001	<.0001	0.0002	<.0001	<.0001
DM13	0.36561	0.49766	0.35363	0.15286	0.33451	0.21895
Información psicológica	<.0001	<.0001	<.0001	0.0956	0.0002	0.0163
DM14	0.45487	0.60775	0.46334	0.11732	0.39823	0.34313
Información psiquiátrica	<.0001	<.0001	<.0001	0.2019	<.0001	0.0001
DM22	0.44077	0.39967	0.50347	0.18798	0.53545	0.40842
Incapacidades médicas	<.0001	<.0001	<.0001	0.0398	<.0001	<.0001
DM24	0.37291	0.38514	0.37099	0.16426	0.48895	0.34152
Estado de discapacidad	<.0001	<.0001	<.0001	0.0730	<.0001	0.0001
DM43	0.43101	0.47867	0.44763	0.19636	0.38411	0.33560
Salud mental (otra información mental no recogida en anteriores campos)	<.0001	<.0001	<.0001	0.0316	<.0001	0.0002
DDI4	0.35602	0.33989	0.44764	0.23057	0.41921	0.42323
Prótesis	<.0001	0.0001	<.0001	0.0113	<.0001	<.0001
DM32	0.36802	0.21338	0.30775	0.12024	0.38381	0.42532
Códigos de diagnóstico	<.0001	0.0193	0.0006	0.1908	<.0001	<.0001
DM34	0.30846	0.19759	0.34667	0.11021	0.42384	0.43705
Códigos de tratamiento	0.0006	0.0305	0.0001	0.2308	<.0001	<.0001
DM35	0.64054	0.51395	0.50767	0.42541	0.72538	0.66843
Fecha de admisión	<.0001	<.0001	<.0001	<.0001	<.0001	<.0001
DM18	0.49604	0.51179	0.41943	0.43533	0.35917	0.30376
Medidas corporales (talla, peso)	<.0001	<.0001	<.0001	<.0001	<.0001	0.0007
DM19	0.49428	0.56348	0.38565	0.25004	0.37164	0.50392
Prognosis	<.0001	<.0001	<.0001	0.0059	<.0001	<.0001
DM20	0.35128	0.41909	0.27042	0.24152	0.28070	0.30941
Residencias de ancianos	<.0001	<.0001	0.0028	0.0079	0.0019	0.0006

Pearson Correlation Coefficients, N = 120 Prob > r under H0: Rho=0						
	DM12	DM15	DM16	DM27	DM36	DM39
DM37 Fecha de deceso	0.51918 <.0001	0.61792 <.0001	0.37368 <.0001	0.37593 <.0001	0.50703 <.0001	0.53349 <.0001
DM42 Salud física (otra información física no recogida en anteriores campos)	0.54159 <.0001	0.58429 <.0001	0.47721 <.0001	0.34536 0.0001	0.52286 <.0001	0.40377 <.0001
DM44 Términos médicos que identifican patologías no habituales	0.53549 <.0001	0.49093 <.0001	0.49103 <.0001	0.31721 0.0004	0.46132 <.0001	0.52874 <.0001
DM21 Vacunas	0.34088 0.0001	0.32363 0.0003	0.47716 <.0001	0.19172 0.0359	0.31856 0.0004	0.43681 <.0001
DM28 Número de identificación	0.47201 <.0001	0.26228 0.0038	0.48381 <.0001	0.55141 <.0001	0.46019 <.0001	0.59560 <.0001
DM41 Número de registros de ensayos clínicos	0.31995 0.0004	0.36480 <.0001	0.37736 <.0001	0.24880 0.0061	0.31500 0.0005	0.42291 <.0001
DM2 Servicio prestado	0.32721 0.0003	0.33535 0.0002	0.33026 0.0002	0.32027 0.0004	0.31752 0.0004	0.20770 0.0228
DM17 Laboratorios que analizaron las pruebas médicas	0.38875 <.0001	0.21952 0.0160	0.30941 0.0006	0.43450 <.0001	0.29578 0.0010	0.41799 <.0001
DM38 Fecha de recogida de muestras	0.55973 <.0001	0.49272 <.0001	0.51859 <.0001	0.42224 <.0001	0.53457 <.0001	0.70796 <.0001

Tabla 61 Matriz de correlación para la categoría "datos médicos" parte 2 de 6

Pearson Correlation Coefficients, N = 120						
Prob > r under H0: Rho=0						
	DM40	DM6	DM7	DM13	DM14	DM22
DM3 Tratamiento	0.62625 <.0001	0.35331 <.0001	0.59902 <.0001	0.30547 0.0007	0.29292 0.0012	0.42500 <.0001
DM5 Alergias	0.48118 <.0001	0.55623 <.0001	0.72830 <.0001	0.39275 <.0001	0.49103 <.0001	0.52420 <.0001
DM8 Diagnosis	0.52227 <.0001	0.24697 0.0065	0.61656 <.0001	0.41628 <.0001	0.43485 <.0001	0.35138 <.0001
DM9 Enfermedades	0.49888 <.0001	0.20940 0.0217	0.59241 <.0001	0.36940 <.0001	0.38002 <.0001	0.26327 0.0037
DM10 Estado de embarazo	0.54703 <.0001	0.46929 <.0001	0.59756 <.0001	0.45594 <.0001	0.53999 <.0001	0.49005 <.0001
DM11 Estado de enfermedad	0.63407 <.0001	0.34956 <.0001	0.60833 <.0001	0.41802 <.0001	0.47891 <.0001	0.48560 <.0001
DM12 Historial clínico	0.59881 <.0001	0.37740 <.0001	0.61345 <.0001	0.36561 <.0001	0.45487 <.0001	0.44077 <.0001
DM15 Información recolectada durante un servicio sanitario	0.40414 <.0001	0.33144 0.0002	0.47012 <.0001	0.49766 <.0001	0.60775 <.0001	0.39967 <.0001
DM16 Intervenciones quirúrgicas	0.65014 <.0001	0.36104 <.0001	0.55569 <.0001	0.35363 <.0001	0.46334 <.0001	0.50347 <.0001
DM27 Número de historia clínica	0.49087 <.0001	0.31362 0.0005	0.33814 0.0002	0.15286 0.0956	0.11732 0.2019	0.18798 0.0398
DM36 Fecha de alta	0.67383 <.0001	0.36497 <.0001	0.45908 <.0001	0.33451 0.0002	0.39823 <.0001	0.53545 <.0001
DM39 Fecha de tratamiento	0.89813 <.0001	0.17656 0.0537	0.48033 <.0001	0.21895 0.0163	0.34313 0.0001	0.40842 <.0001
DM40 Fecha de visita	1.00000	0.09912 0.2815	0.47374 <.0001	0.16812 0.0664	0.25259 0.0054	0.34082 0.0001
DM6 Centros de salud	0.09912 0.2815	1.00000	0.51467 <.0001	0.58675 <.0001	0.53379 <.0001	0.56553 <.0001

Pearson Correlation Coefficients, N = 120						
Prob > r under H0: Rho=0						
	DM40	DM6	DM7	DM13	DM14	DM22
DM7	0.47374	0.51467	1.00000	0.62365	0.60121	0.55841
Consumo de sustancias tóxicas	<.0001	<.0001		<.0001	<.0001	<.0001
DM13	0.16812	0.58675	0.62365	1.00000	0.81017	0.66710
Información psicológica	0.0664	<.0001	<.0001		<.0001	<.0001
DM14	0.25259	0.53379	0.60121	0.81017	1.00000	0.75097
Información psiquiátrica	0.0054	<.0001	<.0001	<.0001		<.0001
DM22	0.34082	0.56553	0.55841	0.66710	0.75097	1.00000
Incapacidades médicas	0.0001	<.0001	<.0001	<.0001	<.0001	
DM24	0.30624	0.58029	0.61614	0.75664	0.69097	0.88317
Estado de discapacidad	0.0007	<.0001	<.0001	<.0001	<.0001	<.0001
DM43	0.31041	0.46225	0.58590	0.66152	0.75851	0.68837
Salud mental (otra información mental no recogida en anteriores campos)	0.0006	<.0001	<.0001	<.0001	<.0001	<.0001
DDI4	0.44296	0.46094	0.41570	0.42137	0.41048	0.51638
Prótesis	<.0001	<.0001	<.0001	<.0001	<.0001	<.0001
DM32	0.42628	0.18340	0.35862	0.22285	0.36400	0.57323
Códigos de diagnóstico	<.0001	0.0450	<.0001	0.0144	<.0001	<.0001
DM34	0.44073	0.18752	0.38180	0.22609	0.34295	0.57739
Códigos de tratamiento	<.0001	0.0403	<.0001	0.0130	0.0001	<.0001
DM35	0.60385	0.20171	0.29633	0.15944	0.28071	0.39695
Fecha de admisión	<.0001	0.0272	0.0010	0.0820	0.0019	<.0001
DM18	0.22310	0.39520	0.54546	0.54353	0.50171	0.30947
Medidas corporales (talla, peso)	0.0143	<.0001	<.0001	<.0001	<.0001	0.0006
DM19	0.48603	0.36269	0.54285	0.52540	0.60981	0.43950
Prognosis	<.0001	<.0001	<.0001	<.0001	<.0001	<.0001
DM20	0.20820	0.38416	0.38924	0.50657	0.41433	0.35409
Residencias de ancianos	0.0225	<.0001	<.0001	<.0001	<.0001	<.0001

Pearson Correlation Coefficients, N = 120						
Prob > r under H0: Rho=0						
	DM40	DM6	DM7	DM13	DM14	DM22
DM37	0.42205	0.19630	0.34836	0.26107	0.33468	0.24526
Fecha de deceso	<.0001	0.0316	<.0001	0.0040	0.0002	0.0069
DM42	0.45033	0.25108	0.53205	0.54973	0.52832	0.49188
Salud física (otra información física no recogida en anteriores campos)	<.0001	0.0057	<.0001	<.0001	<.0001	<.0001
DM44	0.56535	0.19802	0.44055	0.39474	0.49199	0.46654
Términos médicos que identifican patologías no habituales	<.0001	0.0302	<.0001	<.0001	<.0001	<.0001
DM21	0.48429	0.35019	0.54074	0.46727	0.49947	0.44240
Vacunas	<.0001	<.0001	<.0001	<.0001	<.0001	<.0001
DM28	0.60294	0.10892	0.27000	0.10935	0.17451	0.26696
Número de identificación	<.0001	0.2363	0.0029	0.2345	0.0566	0.0032
DM41	0.53146	0.17144	0.38339	0.34409	0.34601	0.29835
Número de registros de ensayos clínicos	<.0001	0.0612	<.0001	0.0001	0.0001	0.0009
DM2	0.14005	0.50143	0.32283	0.27829	0.26820	0.32927
Servicio prestado	0.1271	<.0001	0.0003	0.0021	0.0031	0.0002
DM17	0.28343	0.29943	0.35284	0.17655	0.20521	0.28064
Laboratorios que analizaron las pruebas médicas	0.0017	0.0009	<.0001	0.0537	0.0246	0.0019
DM38	0.61384	0.38693	0.65357	0.39669	0.46300	0.52781
Fecha de recogida de muestras	<.0001	<.0001	<.0001	<.0001	<.0001	<.0001

Tabla 62 Matriz de correlación para la categoría "datos médicos" parte 3 de 6

Pearson Correlation Coefficients, N = 120 Prob > r under H0: Rho=0						
	DM24	DM43	DDI4	DM32	DM34	DM35
DM3 Tratamiento	0.40785 <.0001	0.44191 <.0001	0.42810 <.0001	0.24476 0.0071	0.27026 0.0028	0.56371 <.0001
DM5 Alergias	0.43404 <.0001	0.52633 <.0001	0.39903 <.0001	0.35789 <.0001	0.31423 0.0005	0.30302 0.0008
DM8 Diagnosis	0.34013 0.0001	0.42474 <.0001	0.35287 <.0001	0.33528 0.0002	0.21538 0.0182	0.49976 <.0001
DM9 Enfermedades	0.24020 0.0082	0.36170 <.0001	0.30916 0.0006	0.25980 0.0042	0.13855 0.1313	0.40020 <.0001
DM10 Estado de embarazo	0.42051 <.0001	0.46501 <.0001	0.40557 <.0001	0.22970 0.0116	0.27484 0.0024	0.50518 <.0001
DM11 Estado de enfermedad	0.39069 <.0001	0.45928 <.0001	0.37705 <.0001	0.37561 <.0001	0.31975 0.0004	0.62857 <.0001
DM12 Historial clínico	0.37291 <.0001	0.43101 <.0001	0.35602 <.0001	0.36802 <.0001	0.30846 0.0006	0.64054 <.0001
DM15 Información recolectada durante un servicio sanitario	0.38514 <.0001	0.47867 <.0001	0.33989 0.0001	0.21338 0.0193	0.19759 0.0305	0.51395 <.0001
DM16 Intervenciones quirúrgicas	0.37099 <.0001	0.44763 <.0001	0.44764 <.0001	0.30775 0.0006	0.34667 0.0001	0.50767 <.0001
DM27 Número de historia clínica	0.16426 0.0730	0.19636 0.0316	0.23057 0.0113	0.12024 0.1908	0.11021 0.2308	0.42541 <.0001
DM36 Fecha de alta	0.48895 <.0001	0.38411 <.0001	0.41921 <.0001	0.38381 <.0001	0.42384 <.0001	0.72538 <.0001
DM39 Fecha de tratamiento	0.34152 0.0001	0.33560 0.0002	0.42323 <.0001	0.42532 <.0001	0.43705 <.0001	0.66843 <.0001
DM40 Fecha de visita	0.30624 0.0007	0.31041 0.0006	0.44296 <.0001	0.42628 <.0001	0.44073 <.0001	0.60385 <.0001
DM6 Centros de salud	0.58029 <.0001	0.46225 <.0001	0.46094 <.0001	0.18340 0.0450	0.18752 0.0403	0.20171 0.0272

Pearson Correlation Coefficients, N = 120						
Prob > r under H0: Rho=0						
	DM24	DM43	DDI4	DM32	DM34	DM35
DM7	0.61614	0.58590	0.41570	0.35862	0.38180	0.29633
Consumo de sustancias tóxicas	<.0001	<.0001	<.0001	<.0001	<.0001	0.0010
DM13	0.75664	0.66152	0.42137	0.22285	0.22609	0.15944
Información psicológica	<.0001	<.0001	<.0001	0.0144	0.0130	0.0820
DM14	0.69097	0.75851	0.41048	0.36400	0.34295	0.28071
Información psiquiátrica	<.0001	<.0001	<.0001	<.0001	0.0001	0.0019
DM22	0.88317	0.68837	0.51638	0.57323	0.57739	0.39695
Incapacidades médicas	<.0001	<.0001	<.0001	<.0001	<.0001	<.0001
DM24	1.00000	0.69294	0.57417	0.52920	0.55776	0.37563
Estado de discapacidad		<.0001	<.0001	<.0001	<.0001	<.0001
DM43	0.69294	1.00000	0.56337	0.43617	0.43632	0.38089
Salud mental (otra información mental no recogida en anteriores campos)	<.0001		<.0001	<.0001	<.0001	<.0001
DDI4	0.57417	0.56337	1.00000	0.51563	0.53900	0.36690
Prótesis	<.0001	<.0001		<.0001	<.0001	<.0001
DM32	0.52920	0.43617	0.51563	1.00000	0.85218	0.54703
Códigos de diagnóstico	<.0001	<.0001	<.0001		<.0001	<.0001
DM34	0.55776	0.43632	0.53900	0.85218	1.00000	0.55774
Códigos de tratamiento	<.0001	<.0001	<.0001	<.0001		<.0001
DM35	0.37563	0.38089	0.36690	0.54703	0.55774	1.00000
Fecha de admisión	<.0001	<.0001	<.0001	<.0001	<.0001	
DM18	0.38493	0.42242	0.30350	0.07897	0.07717	0.30678
Medidas corporales (talla, peso)	<.0001	<.0001	0.0008	0.3913	0.4021	0.0007
DM19	0.40897	0.52827	0.31854	0.35977	0.31944	0.51833
Prognosis	<.0001	<.0001	0.0004	<.0001	0.0004	<.0001
DM20	0.37314	0.37390	0.34996	0.24240	0.24589	0.36675
Residencias de ancianos	<.0001	<.0001	<.0001	0.0076	0.0068	<.0001

Pearson Correlation Coefficients, N = 120 Prob > r under H0: Rho=0						
	DM24	DM43	DDI4	DM32	DM34	DM35
DM37 Fecha de deceso	0.27120 0.0027	0.33886 0.0002	0.45012 <.0001	0.37361 <.0001	0.38997 <.0001	0.61934 <.0001
DM42 Salud física (otra información física no recogida en anteriores campos)	0.54640 <.0001	0.76662 <.0001	0.47824 <.0001	0.36310 <.0001	0.39768 <.0001	0.48987 <.0001
DM44 Términos médicos que identifican patologías no habituales	0.39646 <.0001	0.64894 <.0001	0.45522 <.0001	0.47080 <.0001	0.46328 <.0001	0.61025 <.0001
DM21 Vacunas	0.38202 <.0001	0.51793 <.0001	0.46812 <.0001	0.35293 <.0001	0.33191 0.0002	0.28282 0.0017
DM28 Número de identificación	0.22941 0.0117	0.32534 0.0003	0.44720 <.0001	0.48441 <.0001	0.38303 <.0001	0.49350 <.0001
DM41 Número de registros de ensayos clínicos	0.27454 0.0024	0.46512 <.0001	0.53912 <.0001	0.32947 0.0002	0.38264 <.0001	0.34038 0.0001
DM2 Servicio prestado	0.34577 0.0001	0.30739 0.0006	0.29662 0.0010	-0.03518 0.7029	-0.01974 0.8305	0.18674 0.0411
DM17 Laboratorios que analizaron las pruebas médicas	0.27171 0.0027	0.10961 0.2334	0.02568 0.7807	0.17217 0.0601	0.20057 0.0281	0.38258 <.0001
DM38 Fecha de recogida de muestras	0.53126 <.0001	0.39867 <.0001	0.40276 <.0001	0.43827 <.0001	0.45748 <.0001	0.50361 <.0001

Tabla 63 Matriz de correlación para la categoría "datos médicos" parte 4 de 6

Pearson Correlation Coefficients, N = 120						
Prob > r under H0: Rho=0						
	DM18	DM19	DM20	DM37	DM42	DM44
DM3 Tratamiento	0.53983 <.0001	0.43481 <.0001	0.23832 0.0088	0.48713 <.0001	0.57376 <.0001	0.46446 <.0001
DM5 Alergias	0.41603 <.0001	0.45529 <.0001	0.26942 0.0029	0.12922 0.1595	0.33227 0.0002	0.42960 <.0001
DM8 Diagnosis	0.55405 <.0001	0.46339 <.0001	0.30273 0.0008	0.39944 <.0001	0.53755 <.0001	0.48656 <.0001
DM9 Enfermedades	0.49364 <.0001	0.38742 <.0001	0.26369 0.0036	0.32315 0.0003	0.45569 <.0001	0.43152 <.0001
DM10 Estado de embarazo	0.44794 <.0001	0.45925 <.0001	0.29993 0.0009	0.32408 0.0003	0.49569 <.0001	0.48311 <.0001
DM11 Estado de enfermedad	0.53836 <.0001	0.54698 <.0001	0.34071 0.0001	0.47733 <.0001	0.60043 <.0001	0.58631 <.0001
DM12 Historial clínico	0.49604 <.0001	0.49428 <.0001	0.35128 <.0001	0.51918 <.0001	0.54159 <.0001	0.53549 <.0001
DM15 Información recolectada durante un servicio sanitario	0.51179 <.0001	0.56348 <.0001	0.41909 <.0001	0.61792 <.0001	0.58429 <.0001	0.49093 <.0001
DM16 Intervenciones quirúrgicas	0.41943 <.0001	0.38565 <.0001	0.27042 0.0028	0.37368 <.0001	0.47721 <.0001	0.49103 <.0001
DM27 Número de historia clínica	0.43533 <.0001	0.25004 0.0059	0.24152 0.0079	0.37593 <.0001	0.34536 0.0001	0.31721 0.0004
DM36 Fecha de alta	0.35917 <.0001	0.37164 <.0001	0.28070 0.0019	0.50703 <.0001	0.52286 <.0001	0.46132 <.0001
DM39 Fecha de tratamiento	0.30376 0.0007	0.50392 <.0001	0.30941 0.0006	0.53349 <.0001	0.40377 <.0001	0.52874 <.0001
DM40 Fecha de visita	0.22310 0.0143	0.48603 <.0001	0.20820 0.0225	0.42205 <.0001	0.45033 <.0001	0.56535 <.0001
DM6 Centros de salud	0.39520 <.0001	0.36269 <.0001	0.38416 <.0001	0.19630 0.0316	0.25108 0.0057	0.19802 0.0302

Pearson Correlation Coefficients, N = 120						
Prob > r under H0: Rho=0						
	DM18	DM19	DM20	DM37	DM42	DM44
DM7 Consumo de sustancias tóxicas	0.54546 <.0001	0.54285 <.0001	0.38924 <.0001	0.34836 <.0001	0.53205 <.0001	0.44055 <.0001
DM13 Información psicológica	0.54353 <.0001	0.52540 <.0001	0.50657 <.0001	0.26107 0.0040	0.54973 <.0001	0.39474 <.0001
DM14 Información psiquiátrica	0.50171 <.0001	0.60981 <.0001	0.41433 <.0001	0.33468 0.0002	0.52832 <.0001	0.49199 <.0001
DM22 Incapacidades médicas	0.30947 0.0006	0.43950 <.0001	0.35409 <.0001	0.24526 0.0069	0.49188 <.0001	0.46654 <.0001
DM24 Estado de discapacidad	0.38493 <.0001	0.40897 <.0001	0.37314 <.0001	0.27120 0.0027	0.54640 <.0001	0.39646 <.0001
DM43 Salud mental (otra información mental no recogida en anteriores campos)	0.42242 <.0001	0.52827 <.0001	0.37390 <.0001	0.33886 0.0002	0.76662 <.0001	0.64894 <.0001
DDI4 Prótesis	0.30350 0.0008	0.31854 0.0004	0.34996 <.0001	0.45012 <.0001	0.47824 <.0001	0.45522 <.0001
DM32 Códigos de diagnóstico	0.07897 0.3913	0.35977 <.0001	0.24240 0.0076	0.37361 <.0001	0.36310 <.0001	0.47080 <.0001
DM34 Códigos de tratamiento	0.07717 0.4021	0.31944 0.0004	0.24589 0.0068	0.38997 <.0001	0.39768 <.0001	0.46328 <.0001
DM35 Fecha de admisión	0.30678 0.0007	0.51833 <.0001	0.36675 <.0001	0.61934 <.0001	0.48987 <.0001	0.61025 <.0001
DM18 Medidas corporales (talla, peso)	1.00000	0.60805 <.0001	0.49769 <.0001	0.35444 <.0001	0.51814 <.0001	0.38533 <.0001
DM19 Prognosis	0.60805 <.0001	1.00000	0.66145 <.0001	0.44488 <.0001	0.54657 <.0001	0.67350 <.0001
DM20 Residencias de ancianos	0.49769 <.0001	0.66145 <.0001	1.00000	0.38618 <.0001	0.42245 <.0001	0.50323 <.0001

Pearson Correlation Coefficients, N = 120 Prob > r under H0: Rho=0						
	DM18	DM19	DM20	DM37	DM42	DM44
DM37 Fecha de deceso	0.35444 <.0001	0.44488 <.0001	0.38618 <.0001	1.00000	0.46511 <.0001	0.53682 <.0001
DM42 Salud física (otra información física no recogida en anteriores campos)	0.51814 <.0001	0.54657 <.0001	0.42245 <.0001	0.46511 <.0001	1.00000	0.75608 <.0001
DM44 Términos médicos que identifican patologías no habituales	0.38533 <.0001	0.67350 <.0001	0.50323 <.0001	0.53682 <.0001	0.75608 <.0001	1.00000
DM21 Vacunas	0.41187 <.0001	0.60158 <.0001	0.51562 <.0001	0.33206 0.0002	0.45010 <.0001	0.57843 <.0001
DM28 Número de identificación	0.20402 0.0254	0.25897 0.0043	0.21749 0.0170	0.45172 <.0001	0.34384 0.0001	0.47116 <.0001
DM41 Número de registros de ensayos clínicos	0.21773 0.0169	0.36712 <.0001	0.34806 <.0001	0.45851 <.0001	0.48051 <.0001	0.61293 <.0001
DM2 Servicio prestado	0.34434 0.0001	0.30514 0.0007	0.11445 0.2132	0.09316 0.3115	0.22108 0.0152	0.18931 0.0384
DM17 Laboratorios que analizaron las pruebas médicas	0.29531 0.0011	0.41459 <.0001	0.25147 0.0056	0.20981 0.0214	0.01141 0.9016	0.19513 0.0327
DM38 Fecha de recogida de muestras	0.35994 <.0001	0.47720 <.0001	0.40159 <.0001	0.55919 <.0001	0.36927 <.0001	0.46820 <.0001

Tabla 64 Matriz de correlación para la categoría "datos médicos" parte 5 de 6

Pearson Correlation Coefficients, N = 120 Prob > r under H0: Rho=0						
	DM21	DM28	DM41	DM2	DM17	DM38
DM3 Tratamiento	0.34977 <.0001	0.44860 <.0001	0.36813 <.0001	0.39680 <.0001	0.28229 0.0018	0.54109 <.0001
DM5 Alergias	0.56540 <.0001	0.42073 <.0001	0.26406 0.0036	0.41359 <.0001	0.45559 <.0001	0.50564 <.0001
DM8 Diagnosis	0.34421 0.0001	0.41166 <.0001	0.33340 0.0002	0.17449 0.0566	0.27549 0.0023	0.45328 <.0001
DM9 Enfermedades	0.38230 <.0001	0.39628 <.0001	0.32362 0.0003	0.10849 0.2382	0.22399 0.0139	0.39221 <.0001
DM10 Estado de embarazo	0.51293 <.0001	0.45215 <.0001	0.34454 0.0001	0.35214 <.0001	0.38962 <.0001	0.50286 <.0001
DM11 Estado de enfermedad	0.45063 <.0001	0.51058 <.0001	0.37251 <.0001	0.35460 <.0001	0.36974 <.0001	0.57559 <.0001
DM12 Historial clínico	0.34088 0.0001	0.47201 <.0001	0.31995 0.0004	0.32721 0.0003	0.38875 <.0001	0.55973 <.0001
DM15 Información recolectada durante un servicio sanitario	0.32363 0.0003	0.26228 0.0038	0.36480 <.0001	0.33535 0.0002	0.21952 0.0160	0.49272 <.0001
DM16 Intervenciones quirúrgicas	0.47716 <.0001	0.48381 <.0001	0.37736 <.0001	0.33026 0.0002	0.30941 0.0006	0.51859 <.0001
DM27 Número de historia clínica	0.19172 0.0359	0.55141 <.0001	0.24880 0.0061	0.32027 0.0004	0.43450 <.0001	0.42224 <.0001
DM36 Fecha de alta	0.31856 0.0004	0.46019 <.0001	0.31500 0.0005	0.31752 0.0004	0.29578 0.0010	0.53457 <.0001
DM39 Fecha de tratamiento	0.43681 <.0001	0.59560 <.0001	0.42291 <.0001	0.20770 0.0228	0.41799 <.0001	0.70796 <.0001

Pearson Correlation Coefficients, N = 120 Prob > r under H0: Rho=0						
	DM21	DM28	DM41	DM2	DM17	DM38
DM40	0.48429	0.60294	0.53146	0.14005	0.28343	0.61384
Fecha de visita	<.0001	<.0001	<.0001	0.1271	0.0017	<.0001
DM6	0.35019	0.10892	0.17144	0.50143	0.29943	0.38693
Centros de salud	<.0001	0.2363	0.0612	<.0001	0.0009	<.0001
DM7	0.54074	0.27000	0.38339	0.32283	0.35284	0.65357
Consumo de sustancias tóxicas	<.0001	0.0029	<.0001	0.0003	<.0001	<.0001
DM13	0.46727	0.10935	0.34409	0.27829	0.17655	0.39669
Información psicológica	<.0001	0.2345	0.0001	0.0021	0.0537	<.0001
DM14	0.49947	0.17451	0.34601	0.26820	0.20521	0.46300
Información psiquiátrica	<.0001	0.0566	0.0001	0.0031	0.0246	<.0001
DM22	0.44240	0.26696	0.29835	0.32927	0.28064	0.52781
Incapacidades médicas	<.0001	0.0032	0.0009	0.0002	0.0019	<.0001
DM24	0.38202	0.22941	0.27454	0.34577	0.27171	0.53126
Estado de discapacidad	<.0001	0.0117	0.0024	0.0001	0.0027	<.0001
DM43	0.51793	0.32534	0.46512	0.30739	0.10961	0.39867
Salud mental (otra información mental no recogida en anteriores campos)	<.0001	0.0003	<.0001	0.0006	0.2334	<.0001
DDI4	0.46812	0.44720	0.53912	0.29662	0.02568	0.40276
Prótesis	<.0001	<.0001	<.0001	0.0010	0.7807	<.0001
DM32	0.35293	0.48441	0.32947	-0.03518	0.17217	0.43827
Códigos de diagnosis	<.0001	<.0001	0.0002	0.7029	0.0601	<.0001

Pearson Correlation Coefficients, N = 120 Prob > r under H0: Rho=0						
	DM21	DM28	DM41	DM2	DM17	DM38
DM34 Códigos de tratamiento	0.33191 0.0002	0.38303 <.0001	0.38264 <.0001	-0.01974 0.8305	0.20057 0.0281	0.45748 <.0001
DM35 Fecha de admisión	0.28282 0.0017	0.49350 <.0001	0.34038 0.0001	0.18674 0.0411	0.38258 <.0001	0.50361 <.0001
DM18 Medidas corporales (talla, peso)	0.41187 <.0001	0.20402 0.0254	0.21773 0.0169	0.34434 0.0001	0.29531 0.0011	0.35994 <.0001
DM19 Prognosis	0.60158 <.0001	0.25897 0.0043	0.36712 <.0001	0.30514 0.0007	0.41459 <.0001	0.47720 <.0001
DM20 Residencias de ancianos	0.51562 <.0001	0.21749 0.0170	0.34806 <.0001	0.11445 0.2132	0.25147 0.0056	0.40159 <.0001
DM37 Fecha de deceso	0.33206 0.0002	0.45172 <.0001	0.45851 <.0001	0.09316 0.3115	0.20981 0.0214	0.55919 <.0001
DM42 Salud física (otra información física no recogida en anteriores campos)	0.45010 <.0001	0.34384 0.0001	0.48051 <.0001	0.22108 0.0152	0.01141 0.9016	0.36927 <.0001
DM44 Términos médicos que identifican patologías no habituales	0.57843 <.0001	0.47116 <.0001	0.61293 <.0001	0.18931 0.0384	0.19513 0.0327	0.46820 <.0001
DM21 Vacunas	1.00000 0.55716	<.0001 0.58752	<.0001 0.08448	0.3589 0.12838	0.1623 0.48139	<.0001
DM28 Número de identificación	<.0001 0.55716	1.00000 0.48895	<.0001 0.02106	0.8194 0.29835	0.0009 0.49604	<.0001

Pearson Correlation Coefficients, N = 120 Prob > r under H0: Rho=0						
	DM21	DM28	DM41	DM2	DM17	DM38
DM41 Número de registros de ensayos clínicos	0.58752 <.0001	0.48895 <.0001	1.00000	0.03724 0.6864	0.01154 0.9004	0.42597 <.0001
DM2 Servicio prestado	0.08448 0.3589	0.02106 0.8194	0.03724 0.6864	1.00000	0.50316 <.0001	0.31584 0.0004
DM17 Laboratorios que analizaron las pruebas médicas	0.12838 0.1623	0.29835 0.0009	0.01154 0.9004	0.50316 <.0001	1.00000	0.54236 <.0001
DM38 Fecha de recogida de muestras	0.48139 <.0001	0.49604 <.0001	0.42597 <.0001	0.31584 0.0004	0.54236 <.0001	1.00000

Tabla 65 Matriz de correlación para la categoría "datos médicos" parte 6 de 6

ANEXO 7 Resultado de las entrevistas

En este anexo se recoge la Tabla 66 con los resultados de las entrevistas para la revisión interna del conjunto de datos médicos:

R.D. 1093/2010	Dato médico	INV1	INV2	MUE	MP	PI	MF	MU
Sí	Nombre	No	No	Sí	Sí	Sí	Sí	Sí
Sí	Primer apellido	No	No	Sí	Sí	Sí	Sí	Sí
Sí	Segundo apellido	No	No	Sí	Sí	Sí	Sí	Sí
No	Apodo	No	No	Sí	No	No	No	No
Sí	Persona de contacto del paciente	No	No	Sí	Sí	Sí	Sí	Sí
Sí	Fecha de nacimiento	Sí	Sí	Sí	Sí	Sí	Sí	Sí
No	Estado civil	No	Sí	Sí	Sí	Sí	Sí	Sí
Sí	Género	Sí	Sí	Sí	Sí	Sí	Sí	Sí
Sí	Nacionalidad	Sí	Sí	Sí	Sí	Sí	Sí	Sí
No	Idioma	No	Sí	Sí	No	No	No	No
Sí	Teléfono	No	No	Sí	Sí	No	Sí	Sí
No	Fax	No	No	Sí	No	No	Sí	No
No	Correo electrónico	No	No	Sí	No	No	Sí	No
Sí	Calle	No	No	Sí	Sí	No	Sí	Sí
No	Ciudad	No	Sí	Sí	Sí	No	Sí	Sí
Sí	Municipio	No	Sí	Sí	Sí	No	Sí	Sí
No	Distrito	No	Sí	Sí	No	No	No	Sí
Sí	Dirección	No	No	Sí	Sí	No	Sí	Sí
Sí	Comunidad autónoma	Sí	Sí	Sí	Sí	No	Sí	Sí
Sí	País de nacimiento	Sí	Sí	Sí	Sí	No	Sí	Sí
Sí	Código postal	Sí	No	Sí	Sí	No	Sí	Sí
No	Otros códigos de localización	No	No	Sí	No	Sí	No	No
No	Matrícula del vehículo	No	No	No	No	Sí	No	No
No	Bastidor del vehículo	No	No	No	No	No	No	No
No	Número de licencia de conducir	No	No	No	No	No	No	No
No	Números certificado/licencia	No	No	No	No	No	No	No
No	Otros identificadores de vehículo	No	Sí	No	No	No	No	No
No	Origen étnico	Sí	Sí	Sí	Sí	Sí	Sí	No
No	Religión	No	Sí	Sí	No	No	No	No
No	Creencias	No	Sí	Sí	No	No	No	No

R.D. 1093/2010	Dato médico	INV1	INV2	MUE	MP	PI	MF	MU
No	Opiniones personales	No	No	Sí	No	No	No	No
No	Color	Sí	Sí	Sí	No	No	No	No
No	Condición de minoría visible	Sí	Sí	Sí	No	No	No	No
Sí	Cultura	No	Sí	Sí	No	No	No	No
No	Necesidades personales	Sí	Sí	Sí	No	Sí	No	No
No	Opinión política	No	Sí	Sí	No	No	No	No
No	Opiniones, comentarios o preferencias	No	Sí	Sí	No	No	No	No
No	Orientación sexual	Sí	Sí	Sí	No	Sí	No	No
No	Origen filosófico	No	Sí	Sí	No	No	No	No
No	Origen religioso	No	Sí	Sí	No	No	No	No
Sí	Origen social	Sí	Sí	Sí	No	No	No	No
No	Pluralidad de nacimiento	Sí	Sí	Sí	No	Sí	No	No
No	Historial criminal	No	Sí	Sí	No	No	No	No
No	Símbolos (tatuajes, piercings, escarificaciones, ...)	No	Sí	Sí	No	No	No	No
No	Identidad aborigen	No	Sí	Sí	No	No	No	No
No	Dirección IP (Dirección de red o Internet)	Sí	No	Sí	No	No	No	No
No	Enlaces a páginas o usuarios en Internet	No	No	Sí	No	No	No	No
No	Afiliación sindical	No	No	Sí	No	No	No	No
No	Años de escolaridad	Sí	Sí	Sí	Sí	Sí	No	No
No	Historial de educación	No	Sí	Sí	Sí	Sí	No	No
Sí	Historial de empleo	Sí	Sí	Sí	Sí	Sí	Sí	No
No	Iniciales	No	No	Sí	No	No	No	No
No	Posición dentro de la familia (padre, hermano mayor, etc)	No	No	Sí	Sí	Sí	Sí	No
No	Nombre del último empleo	No	Sí	Sí	Sí	No	No	No
No	Organización en la que trabaja	No	Sí	Sí	No	No	Sí	No
Sí	Profesión	Sí	Sí	Sí	Sí	Sí	Sí	No
No	Donación fluidos corporales	No	Sí	Sí	No	Sí	No	Sí
No	Donación sustancia corporal	No	Sí	Sí	No	Sí	No	Sí

R.D. 1093/2010	Dato médico	INV1	INV2	MUE	MP	PI	MF	MU
No	Donación de cuerpo	No	Sí	Sí	No	Sí	No	Sí
No	Donación parte corporal	No	Sí	Sí	No	Sí	No	Sí
No	Ingresos totales	Sí	Sí	Sí	Sí	Sí	No	No
Sí	Proveedor de atención médica	No	Sí	Sí	No	Sí	No	Sí
No	Transacciones financieras	No	No	Sí	No	No	No	No
No	Números de los beneficiarios del plan de salud	No	No	Sí	No	Sí	No	No
Sí	Número de la Seguridad Social	No	No	Sí	Sí	Sí	Sí	Sí
No	Datos financieros (No clínicos)	No	No	Sí	No	Sí	No	No
Sí	Representante del paciente	No	No	Sí	Sí	No	No	No
No	Número del beneficiario del seguro médico	No	No	Sí	No	Sí	No	No
No	Números de cuenta	No	No	Sí	No	Sí	No	No
Sí	Número de tarjeta sanitaria	No	No	Sí	Sí	Sí	No	Sí
No	Datos biométricos de voz	No	No	Sí	No	No	No	No
No	Otros datos biométricos (iris, firma, ...)	No	No	Sí	No	No	No	No
Sí	Uso de aparatos auditivos	Sí	Sí	Sí	Sí	Sí	Sí	No
Sí	Uso de aparatos oftalmológicos	Sí	Sí	Sí	Sí	Sí	Sí	No
Sí	Uso de aparatos ortopédicos	Sí	Sí	Sí	Sí	Sí	Sí	No
Sí	Prótesis	Sí	Sí	Sí	Sí	Sí	Sí	No
No	Número identificativo de los dispositivos implantados	No	No	Sí	No	Sí	No	Sí
No	Tipo sanguíneo	Sí	Sí	Sí	Sí	Sí	Sí	Sí
No	Salud física	Sí	Sí	Sí	Sí	Sí	Sí	No
No	Salud mental	Sí	Sí	Sí	Sí	Sí	Sí	No
Sí	Servicio prestado	No	Sí	Sí	Sí	Sí	No	Sí
Sí	Tratamiento	Sí	Sí	Sí	Sí	Sí	Sí	Sí
No	Ácido desoxirribonucleico	Sí	Sí	Sí	No	Sí	No	No

R.D. 1093/2010	Dato médico	INV1	INV2	MUE	MP	PI	MF	MU
Sí	Alergias	Sí	Sí	Sí	Sí	Sí	Sí	Sí
Sí	Centros de salud	Sí	Sí	Sí	Sí	Sí	Sí	Sí
Sí	Consumo de sustancias tóxicas	Sí	Sí	Sí	Sí	Sí	Sí	Sí
Sí	Diagnos	Sí	Sí	Sí	Sí	Sí	Sí	Sí
No	Elemento indicativo de edad	Sí	Sí	Sí	No	Sí	No	Sí
Sí	Enfermedades	Sí	Sí	Sí	Sí	Sí	Sí	Sí
No	Estado embarazo	Sí	Sí	Sí	Sí	Sí	Sí	Sí
No	Estado de enfermedad	Sí	Sí	Sí	Sí	Sí	Sí	Sí
Sí	Historial clínico	Sí	Sí	Sí	Sí	Sí	Sí	Sí
No	Información psicológica	Sí	Sí	Sí	Sí	Sí	Sí	No
No	Información psiquiátrica	Sí	Sí	Sí	Sí	Sí	Sí	No
Sí	Información recolectada durante un servicio sanitario	Sí	Sí	Sí	Sí	Sí	No	Sí
Sí	Intervenciones quirúrgicas	Sí	Sí	Sí	Sí	Sí	Sí	Sí
Sí	Laboratorios que analizaron las pruebas médicas	Sí	Sí	Sí	No	Sí	No	No
No	Medidas corporales (talla, peso)	No	Sí	Sí	Sí	Sí	Sí	No
No	Prognosis	Sí	Sí	Sí	Sí	Sí	No	No
No	Residencias de ancianos	No	Sí	Sí	No	Sí	No	No
No	Términos médicos	Sí	Sí	Sí	Sí	Sí	No	No
Sí	Vacunas	Sí	Sí	Sí	Sí	Sí	Sí	No
	Incapacidades médicas	Sí	Sí	Sí	Sí	Sí	Sí	Sí
Sí	Imágenes fotográficas identificables	Sí	Sí	Sí	No	Sí	Sí	No
No	Estado de discapacidad	Sí	Sí	Sí	Sí	Sí	Sí	Sí
Sí	Genética familiar cercano	Sí	Sí	Sí	No	Sí	No	No
No	Genética propia	Sí	Sí	Sí	No	Sí	No	No
No	Número de historia clínica	No	No	Sí	Sí	Sí	Sí	Sí
Sí	Número de identificación	No	No	Sí	Sí	Sí	Sí	Sí
No	Número de placa	No	No	Sí	No	Sí	No	Sí

R.D. 1093/2010	Dato médico	INV1	INV2	MUE	MP	PI	MF	MU
	médica							
Sí	Número de registro médico	No	No	Sí	No	Sí	No	Sí
Sí	Números de serie de dispositivos identificadores	No	No	Sí	No	Sí	No	Sí
Sí	Códigos de diagnóstico	No	Sí	Sí	Sí	Sí	Sí	Sí
No	Códigos de eventos adversos	Sí	Sí	Sí	Sí	Sí	Sí	Sí
Sí	Códigos de tratamiento	Sí	Sí	Sí	Sí	Sí	Sí	Sí
Sí	Fecha de admisión	Sí	Sí	Sí	Sí	Sí	Sí	Sí
Sí	Fecha de alta	Sí	Sí	Sí	Sí	Sí	Sí	Sí
No	Fecha de deceso	Sí	Sí	Sí	Sí	Sí	Sí	Sí
Sí	Fecha de recogida de muestras	Sí	Sí	Sí	Sí	Sí	Sí	Sí
No	Fecha de tratamiento	Sí	Sí	Sí	Sí	Sí	Sí	Sí
Sí	Fecha de visita	Sí	Sí	Sí	Sí	Sí	Sí	Sí
<ul style="list-style-type: none"> • 2 Investigadores expertos en el área de informática médica (INV1, INV2). • Médica, profesora e investigadora de la Universidad Europea (MUE). • 3 médicos de las especialidades de: urología (MU), medicina forense (MF) y psicólogo clínico (MP). • Técnico informático (TI). 								

Tabla 66 Datos médicos resultantes de la revisión interna