eman ta zabal zazu

Universidad del País Vasco

Euskal Herriko Unibertsitatea

Un estudio del impacto ético y social de las nuevas tecnologías en la práctica de la biomedicina

Iñigo de Miguel Beriain

2021

Contenido

Prefacio

Este documento constituye la plasmación de la tesis doctoral titulada "Un estudio del impacto de las nuevas tecnologías en la práctica de la biomedicina". Ha sido presentada por don Iñigo de Miguel Beriain para la obtención del título de Doctor correspondiente al Programa de Doctorado en Investigación Biomédica en la modalidad de compendio de publicaciones. Esta colección ha sido elaborada siguiendo lo dispuesto en la normativa de Gestión de las Enseñanzas de Doctorado (Acuerdo de 12 de diciembre de 2019, del Consejo de Gobierno de la Universidad del País Vasco/Euskal Herriko Unibertsitatea, por el que se aprueba la Normativa de Gestión de las enseñanzas de doctorado (BOPV de 5 de marzo de 2020). Su estructura sigue lo dispuesto en el artículo 43 de dicha norma, que señala lo siguiente:

> Artículo 43.- Estructura de la tesis por compendio de publicaciones.
>
> La tesis doctoral presentada por compendio de publicaciones estará constituida por un conjunto de trabajos publicados y/o aceptados, justificados por su unidad temática, de acuerdo a la siguiente estructura:
>
> 1.- Una sección inicial de síntesis, con una extensión mínima orientativa de 10.000 palabras, que contenga:
>
> a) Introducción, en la que se realice una presentación de la tesis y se justifique la unidad temática.
>
> b) Marco teórico en el que se inscribe el tema de la tesis y herramientas metodológicas utilizadas.
>
> c) Hipótesis y objetivos generales y específicos a alcanzar, indicando en qué publicación o publicaciones se abordan.
>
> d) Resumen y, en su caso, discusión de los resultados obtenidos.
>
> e) Fuentes referenciadas.
>
> 2.- La segunda sección de la tesis estará formada por las conclusiones de la misma.
>
> 3.- La tercera sección corresponderá al anexo, que debe contener los artículos, libros o capítulos de libro publicados o aceptados, bajo el título de "Trabajos publicados" o, si fuera el caso, "Trabajos Publicados o aceptados". Se incluirá la versión íntegra publicada o aceptada de cada contribución, en el supuesto de disponer la autorización de la revista para utilizar el artículo en la tesis doctoral. En el supuesto de no tenerla, se utilizará una versión accepted manuscript apta para su difusión

pública. En ambos supuestos se incluirán las referencias bibliográficas completas, y se indicará el factor de impacto de la revista en el año de la publicación, su posición relativa en la categoría a la que pertenece, y/u otros indicios de calidad.

Al tratarse de un trabajo que aspira a la mención de "Doctorado Internacional del artículo 25 de la Normativa de Gestión de las enseñanzas de doctorado, he seguido lo más escrupulosamente que ha sido posible lo estipulado en su punto 2 ("al menos el resumen y las conclusiones, se haya redactado y presentado en una de las lenguas habituales para la comunicación científica en su campo de conocimiento, distinta a cualquiera de las lenguas oficial o cooficiales de España"). De ahí que los apartados correspondientes a la presentación de los resultados mi trabajo y su discusión y a las conclusiones se encuentren redactados en inglés.

Me gustaría aprovechar esta ocasión para dar las gracias a las personas que, de un modo u otro, han contribuido a hacer posible este trabajo. En primer lugar, mi directora de tesis, la profesora Begoña Sanz Echeverría, cuya ayuda ha sido imprescindible tanto para la redacción de alguno de los artículos que lo componen, como para su presentación en debida forma. Sin duda, un excelente apoyo que ha ido mucho más allá de lo meramente académico. Este agradecimiento ha de extenderse en general al Departamento de Fisiología y, especialmente, a los profesores Jon Irazusta y Javier Gil, que siempre han estado dispuestos a ayudarme en todo lo necesario.

Junto a ellos, me gustaría agradecer el apoyo de la profesora Maitena Poelemans, que me acogió en su centro de investigación sobre estudios europeos en la Universidad de Pau-Pais del Adour y a la que tendré siempre en mi más alta estima personal y profesional.

A ellas añadiré a todos los miembros del grupo de investigación de la Cátedra de derecho y genoma humano y a todos los investigadores de instituciones como la UNED, el CSIC, la universidad de León, la de Granada, etc. con los que he compartido años de fructífera colaboración que, espero, se prolongarán durante mucho más tiempo.

No puedo, por fin, dejar de mencionar a mi familia: Paulina, Nikolai, Alex y todos los Galka-Mitusinska. Y a mi madre y tíos-padrinos, que tan generosamente me han apoyado en todo durante tantos años.

SECCIÓN 1: SÍNTESIS DE LA TESIS DOCTORAL

Introducción: presentación de la tesis justificación de su unidad temática

La tesis doctoral que ahora presento parte de una evidencia: en los últimos años la práctica de la biomedicina se está viendo afectada por la aparición de nuevas tecnologías que inciden directamente en el enfoque de la actuación biosanitaria y las posibilidades de mejorar considerablemente las intervenciones enfocadas a la prevención y tratamiento de diversas patologías. Entre estas tecnologías destacan dos, que son las que he estudiado de manera particularmente atenta en este trabajo: la edición genética de seres humanos mediante la tecnología CRISPR-Cas y la irrupción de los nuevos mecanismos de inteligencia artificial y otras formas de procesos de control a través del uso de datos, como las "píldoras inteligentes", que combinan el uso de *wereables* con el envío de datos, propiciando las decisiones automatizadas sobre los pacientes. En esta tesis doctoral he analizado el impacto que dichas tecnologías desde la perspectiva de la investigación y la práctica clínica, intentando enfocar los principales retos que presenta cada una de ellas y las iniciativas que habría que adoptar para favorecer un uso aceptable de las mismas desde una perspectiva ética y social.

En el caso de la primera de ambas tecnologías, la edición genética, mi trabajo se ha centrado en las prácticas que afectan a la línea germinal humana, esto es, en la modificación de los genes que se transmiten a nuestra descendencia. Como tal, esta cuestión resulta sumamente polémica a día de hoy, siendo así que el debate sobre su aceptabilidad, así como las consecuencias sociales que podría traer consigo se halla en plena efervescencia (De Miguel, 2019a). Mis aportaciones han intentado hallar puntos de encuentro entre las diferentes posturas que ahora mismo existen, a la par que han incidido en algunas cuestiones poco analizadas, como las que tienen que ver con los problemas de equidad en las sociedades del futuro.

En el caso de la Inteligencia Artificial y otras tecnologías que permiten un seguimiento del paciente y su tratamiento a través del uso de datos, estamos ante un escenario diferente, al menos en tanto que las herramientas tecnológicas, en este caso, despiertan menos animadversión en sí misma. Muchos autores las consideran, en realidad, como unas tecnologías sumamente prometedoras, ya que prometen mejorar

sustancialmente muchos aspectos de nuestra vida, entre ellos, la asistencia sanitaria. El uso de la Inteligencia Artificial en este sector permitirá a buen seguro una mejora de nuestra capacidad de diagnóstico y prognosis. No obstante, también plantea grandes problemas éticos y jurídicos. Entre ellos, la posibilidad o no de autorizar decisiones automatizadas o las necesidades de información a los pacientes; problemas relacionados con el diseño y la validación de estos instrumentos; cuestiones relativas al uso de datos personales para su empleo, o la posibilidad de que acaben generando graves discriminaciones sociales. El desarrollo de *software* inteligente capaz de verificar la adherencia a un tratamiento, por su parte, permitiría, teóricamente, mejorar mucho la práctica de la medicina. Sin embargo, una mirada más atenta muestra que las cosas son más complejas de lo que parece. En mi trabajo, he abordado algunos de estos problemas, fundamentalmente los que tienen que ver con la comprensión del funcionamiento de estos sistemas, la información que ha de darse a los pacientes y el derecho que estos pueden tener de negarse a que se utilice esta tecnología para su atención clínica.

Formalmente, se trata de una tesis que recoge un *corpus* de artículos y correspondencia publicados en revistas científicas de elevado impacto durante los años 2019, 2020 y 2021. Podría decirse que, en su conjunto, este compendio reúne aportaciones que tienen que ver con las diferentes cuestiones que acabo de citar. Todas ellas apuntan hacia la biomedicina del futuro y, más allá, incluso, el tipo de sociedad que podemos construir a consecuencia del uso de estas tecnologías.

Marco teórico y herramientas metodológicas utilizadas

La tesis que ahora presento se circunscribe al marco teórico que representa la aplicación de las nuevas tecnologías en el ámbito de la investigación biomédica y la práctica clínica asociada a esta investigación. Más concretamente, mi análisis se centra en las partes relativas al impacto en la sociedad de dichos desarrollos, con especial atención a los problemas éticos y de justicia social a los que pueden dar lugar.

Las herramientas metodológicas utilizadas para la elaboración de la tesis doctoral, por su parte, han sido diversas. Comprenden, entre otras, las siguientes:

- Análisis de la literatura académica existente. En primer lugar, mi investigación se ha basado en un exhaustivo análisis de la literatura sobre los temas sobre los que versa. Esto incluye tanto libros como, en su mayor parte, artículos científicos. En el quinto apartado de este documento se incluye una relación de los más destacados.
- Participación en proyectos y redes de investigación nacionales e internacionales me ha permitido intercambiar opiniones con investigadores de primer nivel nacionales e internacionales. Entre estas redes y proyectos destacan los siguientes:
    - International Network "Navigating Knowledge Landscapes (NKL)", an interdisciplinary research network focusing on health in digital society, de cuyo Steering Committee tengo el honor de ser miembro (http://knowledge-landscapes.hiim.hr).
    - Proyecto "Participatory Approaches to a New Ethical and Legal Framework for ICT (PANELFIT)" PANELFIT, financiado por la UE, sobre utilización de datos para investigación, del que formo parte como Associated Coordinator (www. Panelfit.eu)
    - Proyecto EU CANIMAGE: A European Cancer Image Platform Linked to Biological and Health Data for Next-Generation Artificial Intelligence and Precision Medicine in Oncology. (https://eucanimage.eu/)
    - Proyecto de Investigación de la Junta de Castilla y León, del programa de Apoyo a Proyectos de Investigación Cofinanciados por el Fondo Europeo de Desarrollo Regional (Fondos FEDER/2020): Medicamentos digitales y Bioderecho: Oportunidades y limitaciones del desarrollo de medicamentos con sensor integrado en conexión con un sistema de recogida y transmisión de datos del paciente (LE043P20)
    - Proyecto «Acción Estratégica en Salud Infraestructura de Medicina de Precisión asociada a la Ciencia y Tecnología (IMPaCT) 2020». Proyecto con número de expediente IMP/00009, financiado por el Instituto de Salud Carlos III.
- Realización de actividades formativas, como la asistencia a congresos y seminarios. Mi participación en varios congresos y seminarios a lo largo de estos años me ha permitido dotarme de un conocimiento mucho más adecuado sobre los temas abordados en la tesis. Con tal fin, fue particularmente importante la defensa de diferentes ponencias sobre estos temas. Entre ellas destacan:

- Ponencia titulada “Algoritmos predictivos y utilización de sistemas de decisión automatizados en la MPP”, en el marco del XXVI Congreso Internacional sobre Derecho y Genoma Humano, Bilbao, Bizkaia Aretoa, UPV/EHU 6-8 de Mayo de 2018
- Ponencia titulada: “Edición Genética: aspectos Éticos”. XXVI Congreso Nacional de la Asociación Española de Derecho Sanitario”, Madrid, 18 de octubre de 2019
- Ponencia en la mesa redonda “Aspectos jurídicos del CRISPR/Cas9”, en el seminario El CRISPR/CAS9 y sus implicaciones para el ser humano, organizado por la Universidad de Comillas, 13 de Noviembre de 2019, Madrid.
- Participación en el seminario internacional “CNIO-Sabadell Foundation Workshop on Philosophy & Biomedical Sciences: Debates on conceptual and social issues”, 18 de noviembre de 2019, Madrid, España
- Ponencia titulada “AI and informed consent: the challenges ahead” in the International Congress CPDP. Data Protection and Artificial Intelligency. 22-24 January 2020. Brussels.
- Ponencia en el Bloque Regulatorio del Congreso ¿Editamos humanos? Congreso Internacional sobre edición genética en seres humanos: aspectos científicos, éticos y jurídicos Viernes, 13 de noviembre de 2020 - Facultat de Dret. Universitat de València.
- Participación en la MESA 2. Inteligencia Artifficial: Retos Éticos y Legales, en la III Jornada Anticipando LA MEDICINA DEL FUTURO, organizada por la Fundación Roche, 16/12/2020
- Participación en la Mesa redonda Inteligencia Artificial: aspectos legales y éticos, II Jornada de protección de datos personales: evolución tecnológica y privacidad a debate: 5g e inteligencia artificial. Organizada por la Universidad Francisco de Vitoria y la International Association of Privacy Professionals. 27 enero 2021, en: https://www.ufv.es/jornada-rgpd-ciber/
- Participación en el Workshop on Artificial Intelligence systems and techniques: addressing ethical implications in Horizon Europe, organized by the EU Commission, Directorate-General For Research & Innovation, on 5 March 2021

- Ponencia titulada Desarrollo de mecanismos de IA por parte de las entidades de salud. Problemas éticos y jurídicos, Congreso "Regulación y explotación de big data para los servicios públicos", martes 2, 9 y 16 de marzo (17 h-20 h), organizado por la Universidad de Valencia, 16 de marzo de 2021.
- Ponencia titulada "Declaraciones y Opiniones sobre edición genética: un análisis crítico", XXVII Congreso Internacional De Derecho Y Genoma Humano 18 y 19 de mayo de 2021.
- Ponencia titulada "Marco legal de la telemedicina y la protección de datos", webinar organizado por Doctoralia, 20/05/2021. https://academy.doctoralia.es/webinar-derecho-sanitario-telemedicina
- Ponencia titulada "AI tool to be used for forensic purposes: the ethical and legal tools needed", in the How to cross the "Valley of Death" between research and the forensic market training course, organized by the Multi-Foresee COST Action, 22-23 June 2021

- Realización de una estancia de investigación en en la Universidad de Pau-Pais del Adour, Francia. 1 de abril-1 de julio de 2019. Esta estancia me permitió tanto conocer la situación en el país vecino como analizar la normativa europea en un centro especializado en estudios de este tipo.

Hipótesis y objetivos generales y específicos a alcanzar, indicando en qué publicación o publicaciones se abordan

Las hipótesis de las que partía en mi investigación figuraban en el Plan de Investigación que presenté al inicio de la misma. Son las siguientes:

- Es posible aceptar la edición genética, incluso la que se realiza sobre la línea germinal una vez que el procedimiento ofrezca las garantías de seguridad y eficiencia suficientes, pero hay que determinar cuáles serán estas
- No hay motivos éticos solventes por los que oponerse a la edición genética
- La normativa al respecto está desfasada y necesita de una actualización. Mientras tanto, es posible realizar interpretaciones abiertas de la misma
- El uso de la inteligencia artificial necesita de una normativa más amplia que recoja requisitos de validación, ausencia de sesgos, transparencia, etc.

- La Inteligencia Artificial puede mejorar sustancialmente la biomedicina y el paciente está obligado a contribuir a que así sea

Obviamente, el principal objetivo de mi investigación ha sido contrastar la veracidad o no de estas hipótesis iniciales. No obstante, a lo largo de su desarrollo he tenido ocasión de abordar otros temas, más allá de los inicialmente previstos, o de incidir más de lo que pensaba en cuestiones específicas. Esto ha permitido obtener unos resultados que probablemente sean más enriquecedores de lo que esperaba inicialmente.

Resumen de los resultados alcanzados y discusión

The main results of my research are the following:

- Germline gene editing (GEE) should not be performed until the risks associated to this technology are acceptable
- Germline gene editing is not against human dignity
- Germline gene editing may be considered as a moral imperative
- Heritable human genome editing may treat, cure or prevent diseases in human beings Heritable human genome editing might be necessary. It is much better from a moral point of view than the Preimplantation Genetic Diagnosis plus embryo selection practices
- Changing the genes of our descendants does not necessarily involve changing the human genome
- GGE should be implemented to all health issues, independent of whether they are serious or not
- GGE should focus on absolute goods instead of positional goods
- AI tools might be a game changer in terms of health care, but they also involve challenging issues
- We need to keep in mind issues related to the lack of transparency
- There should be a right to refuse diagnostics and treatment planning by artificial intelligence
- The use of automated decision making in wearables creates relevant health care issues

In the next pages I expose and discuss all these results in detail.

Germline gene editing (GEE) should not be performed until the risks associated to this technology are acceptable

One of the main issues that hinders the use of GGE technologies in human embryos is the uncertainty related to its outcomes. At the present moment, it is impossible to raise seriously the question of whether GEE can be used as a therapeutic tool. Our poor control of technology and the substantial risks involved make its clinical application fully unethical, since we cannot ensure that the persons created from the modified embryos will enjoy healthy lives (De Miguel Beriain, 2019a). Therefore, the debate must be directed towards a future in which these technical issues have been resolved.

However, if we are considering using GGE for research purposes, the situation is totally different. In these cases, embryos affected by a change in their germline will not be transferred into a woman. This subtle distinction allows researchers to justify genetic modifications performed on in vitro embryos that will subsequently be destroyed: as they will never become adults, the risks of causing someone a catastrophic damage and introducing dramatic changes to the human genome are zero, even if their germ line is manipulated. Therefore, a regulation based on this criterion might allow freely germ line modifications for research purposes while banning its clinical use, even if under strict oversight by Ethics Committees, as proposed by the International Society for Stem Cell Research and the National Academic of Sciences Report (National Academies, 2017).

Germline gene editing is not against human dignity

According to some bioethical trends (Habermas, 2003), GGE would threaten human dignity, since they consider the human genome as the physical representation of human dignity. Therefore, those who wish to engage in the debate on the ethics of gene editing should first focus on the idea of human dignity and its role in this discussion. Quite curiously, even those who disagree with this view have not questioned that GGE might be against human dignity (Sykora, 2017). However, two of the papers I published in the context of my research (de Miguel Beriain, 2018; de Miguel Beriain and Sanz, 2020a) show that respect for human dignity should actually support arguments to continue with genetic editing of embryos.

In order to understand why, we should start by knowing that there is a broad tendency in bioethics to directly link the notion of human dignity to the human genome. Since it is the genome what determines who belongs to the human species, and since being part of the human species confers dignity, it seems reasonable to link human dignity to the human genome. Most opponents of germline editing hold a normative conclusion from this belief: since our dignity is embedded in our genome, we must refrain from altering it (Nuffield Council, 2018).

In my papers, I stated that this normative proposal is problematic: under the belief that our dignity can only remain undiminished if the human genome remains unchanged, we cannot tolerate any new mutation in the human genome – regardless of whether it is produced by human action or naturally. Thus, we should indeed proceed to use GGE to erase those changes. As can be seen, this is a totally different logical conclusion to the argument.

Some would argue that my position is a sophism, i.e., that respect for human dignity does not entail preserving the human genome at all costs, but that humans should not arrogate to themselves the power to alter it. Indeed, DB Costam's reply to my paper in EMBO (Costam, 2019) held that human dignity only suffers violation when some individuals are invested with the authority to engineer the genomes of others who are not yet born. Therefore, the relevant question is not *how* heritable information changes, but rather *who* is entitled to change it. The non-metaphysical alternative in support of this argument is the view that "tried and tested" natural processes are more reliable than "human tinkering" (Coyle, 2005).

Nevertheless, such a rejoinder involves some serious issues, as I pointed out in my response to Dr. Costam (De Miguel, 2019g). Firstly, if we share the idea of a non-human force, we might as well do away with the notion of human dignity as autonomous dignity. If we accept that there is a "great unknown" rationality superior to ours, it is no longer possible to hold Kant's claim that "morality, and humanity so far as it is capable of morality, are the only things that have dignity" (Kant, 2008). Thus, we resurrect the idea of *Imago Dei*, so successful in the pre-modern era (Welz, 2016), but at the cost of renouncing to the idea of human autonomy—the capacity to govern our destiny by producing universal moral norms—as the source of our dignity. Secondly, accepting the idea that nature is wiser than human beings implies casting doubt on the legitimacy of science to interfere with evolution. If there is a hidden intelligence in nature, would it be

a symptom of hubris or foolishness to defy its rules? But, in such case, wouldn't we have to—for example—forego plant breeding to increase yield or resistance against disease because it clearly interferes with evolution? Wouldn't we have to ban gene therapy against a rare immune disorder as it corrects a natural mutation? In fact, human intervention is only reasonable if we either think that evolution works blindly and randomly, or if a "great unknown" encourages us to intervene in the world. If not, how can we justify a rebellion against the laws of nature? Last, but not least, we should consider that abstaining from intervention in the human genome when action is possible is already a form of intervention by omission. The emergence of genetic editing means that, whether we like it or not, the human genome will end up being what we decide it to be, by action or by omission. Therefore, the idea that "let it go" would avoid our responsibility is simply false.

In conclusion, even if we were to accept that human dignity and the human genome are linked—which is debatable—this would not necessarily imply an obligation to preserve the human genome. Consequently, there is no reason why we should not proceed to alter the human genome, since it is not a fixed and immutable entity that has to be preserved for the better of humanity. This is precisely the reason why we do not ban interventions for a somatic purpose, which might have unwanted side-effects on the germ cell line, such as certain treatments of cancer by radiotherapy or chemotherapy, which may affect the reproductive system of the person undergoing the treatment: if we believed that any change in the genome caused by human action was a breach of human dignity, we shall refrain from any kind of action that could lead to this consequence whether directly or indirectly – even if this means abolishing radio- or chemotherapy. The fact that this appeal does not arise is the best evidence that we do not believe that human genome is sacred. Thus, GGE is not, it cannot be, against human dignity.

Germline gene editing may be considered as a moral imperative

I have merely stated that modifying the human genome does not violate human dignity, which is a key argument in my thesis. However, I dare go one step further by stating that modification of our germline might even be considered a moral requirement, an exigency of our autonomy, as I did in the paper on human dignity co-authored with Begoña Sanz (De Miguel and Sanz, 2020a). Once understood that both action and

omission are expressions of autonomy, it is easy to conclude that we must make our decision depending on the interests of the affected human being, not the need to preserve the human genome.

This could be better understood if I introduce an example. Let us imagine, for instance, that someone is suffering from a genetic condition and we have to choose not to save his/her life considering that such intervention would alter the human genome. Whatever decision we make, we will be expressing our autonomy. It is simply untrue that an omission will avoid responsibility. What has endowed us with greater autonomy has been the mere fact of having the power to intervene in our genome. Whether we exercise it or not will only affect our moral stature. Whether to save a life will generate an undoubted moral responsibility, regardless of which option is chosen. It is the power given to us by the techniques of genetic modification − not their exercise − that makes us more autonomous, and therefore more responsible. This is why it makes no sense to think that acting as if nothing had happened is the option that respects the idea of human dignity the most. We must react as responsible adult humans, recognise our new capabilities, and begin to use them (or not) according to our moral compass. Our dignity will not be altered by the decisions taken, but our moral status will surely be affected by the reasons involved in them.

This point could, however, be contested by the argument that although such intervention in itself does not pose a direct threat to human dignity, it could constitute an indirect threat—for example, by enabling other interventions that do alter the human gene pool or even redefine fundamental social relationships (between parents and children, individuals and communities, citizens and states), and associated notions of responsibility and care (Saha et al. 2018).

The problem with this kind of argument is that it is disrespectful to what it claims to defend: human dignity. What the argument contends is the following: although we can cure a pathology in an embryo or in a human being, we should refrain from doing so in order to preserve higher social values or goods. However, is this kind of assertion in line with the fundamental normative principle that follows from the idea of dignity? Is this the proper way to treat an embryo, as an end in itself and not as a mere means to preserve a higher good? Or is this suggestion more like a social eugenics exercise based on the assumption that the interests of the group should prevail over those of the individual? The answer is clear: respecting human dignity means putting the individual human and his or

her interests before those of his or her possible descendants, social group or the like. Therefore, respect for human dignity would actually require that we intervene in the human germline in order to preserve the interests of a particular human being, disregarding the collective interest. Thus, GGE might be considered as a moral imperative.

Heritable human genome editing may treat, cure or prevent diseases in human beings

One of the most astonishing statements included in the so-called Geneva Statement on Heritable Human Genome Editing was this: *"Heritable human genome editing would not treat, cure, or prevent disease in any existing person (...)[It] should be understood not as a medical intervention, but as a way to satisfy parental desires for genetically related children or for children with specific genetic traits."* (Andorno et al., 2020). In one of the articles (De Miguel, 2019f) that are part of this thesis, I provide good reasons to think that this assessment in not acceptable. Indeed, I consider that embryo editing can indeed be considered as a medical intervention. This will depend on the purposes of the intervention, of course, but if it is aimed at preventing a disease to happen, this should be considered as such.

There are several reasons to hold this belief. First, there are millions of people who consider human embryos as existing human beings. Furthermore, even if this were not the case, that is, even if we were to accept that the GGE would not treat, cure, or prevent disease in any existing person because embryos cannot be considered as such, this does not mean that it would not produce this result in people who will exist in the future. In fact, if it ever works in the way it is expected, GGE might reduce someone's genetic predisposition to some types of cancer or prevent a person from transmitting hemophilia or suffering from Huntington Disease. Thus, these must be considered as therapeutic interventions. This is crucial, since interventions related to health care can be considered as a moral imperative towards our descendants.

Heritable human genome editing might be necessary. It is much better from a moral point of view than the Preimplantation Genetic Diagnosis plus embryo selection practices

An argument that is linked to the former one considers that GGE is not at all necessary, since preimplantation genetic diagnosis (PGD) can perfectly play the same role and it is much safer. This idea is somehow right, since PGD is a well-know technology, often used to detect genetic abnormalities in human embryos. If these are incompatible with human life or involve a high probability to develop a serious disease, the embryos that show them are discarded. On this basis, understanding why some authors consider that, somehow, this technique already provides the uses that GGE might bring us in the future is easy. But, is it true that GGE is unnecessary? I really do not think so and my papers reflect this position.

In my thesis, I have stated that this is not a fair way to analyze the case, since these arguments forget some essential issues (De Miguel and Ishii, 2019). First, GGE might allow us to obtain results that PGD could never provide (Steffann et al. 2018). Ideally, GGE could allow for correcting multiple genes of an embryo, which would go far beyond preventing the birth of children affected by a monogenic disease. For instance, GGE could give our offspring an expression of genes more suited to reducing their predisposition to cancer or to improving their immune system's performance. While this may not seem easy to implement right now, it cannot be ruled out that the situation will change dramatically in the future. What is undeniable, in any case, is that this kind of substantial improvement will only be possible thanks to the use of GGE techniques. Therefore, it is uncertain whether PGD and GGE possess a similar capacity in purely scientific terms. Indeed, GGE is far more versatile than PGD followed by an embryo selection. Thus, it will be exponentially superior, if we are effectively capable of acquiring sufficient knowledge about the human genome to understand what changes are satisfactory for human beings. To this it must be added that GGE could serve to help some parents to have genetically related descendants, a circumstance that would be impossible by using preimplantation genetic diagnosis, due to the biological constraints at stake (Ranisch, 2020). Thus, GGE can hardly be considered as an unnecessary innovation.

Furthermore, the nature of both types of technologies is strictly different. GGE is intended to safeguard the health of offspring who may be suffering from various

pathologies through genetic modifications. Therefore, it constitutes a therapeutic action, free of any moral suspicion. Embryo selection, instead, can only be considered a therapeutic action for perspective parents who suffer from the impossibility of generating biologically healthy offspring. Indeed, detractors of genetic selection argue that this technology contains an aroma of eugenics (De Miguel and Penasa, 2018). In fact, what the technique involves is not to "cure" (MacKellar, 2014) embryos but simply to choose which embryos will be transferred. Instead, I have argued that GGE involves a therapeutic action on embryos and the people they will become (see former section). Therefore, my analysis concludes, again, that one must consider that GGE as a much preferable technology to PGD.

Changing the genes of our descendants does not necessarily involve changing the human genome

What does a change in the human genome mean? This is a key question, both from a legal and an ethical point of view, that I have explored in depth in one of the papers that are part of this thesis (De Miguel, 2019a). The concept of human genome plays a key role in this debate. However, it is never defined in the academic discussion on gene editing or in the legal framework applicable. Therefore, I focused an essential part of my research in clarifying its meaning. To begin with, I consider that the concept of human genome is not synonymous to the idea of human germline. Indeed, the expression "human genome" could have two different meanings. One the one hand, "human genome" can be defined as "all the genetic information in a person" (MedicineNet 2018), that is, an individual of the human species. However, this kind of definition does not seem to match well either with our common understanding or with the meaning of the idea of "human genome" included in the current regulation. If we were to think that the human genome refers to someone's genetic information, then we would need to conclude that our current laws ban all types of gene editing (even somatic line edition, of course), in so far as all of them involve, by definition, some type of alteration to an individual's genome and, thus, a change in the human genome. However, this is simply absurd.

The conclusion I have arrived into in the former paragraph leads our attention to another definition of human genome (*the* human genome), linked to the idea of the whole pool of genetic information possessed by the human species, that is, a concept completely separated from that of an individual genome (as mentioned in Articles number 1 and 3 of the UNESCO Universal Declaration on the Human Genome). This alternative sense of human genome allows us to trace much subtler distinctions in terms of the ethical and legal debates. For instance, we could hold that editing someone's genome does not necessarily involve a change in the human genome. Suppose, for instance, that we modify someone's genome by changing the expression of a gene producing Huntington's disease for its normal, healthy expression. According to this second definition of human change, such an alteration would certainly change the subject's (and his or her descendants') genome, but not the human genome at all (at least, if we manage to avoid off-target changes), in so far as the modification would not introduce any novelty into the human gene pool. Instead, a modification in someone's genome that creates a new genetic combination, previously non-existent in the human species, should certainly be considered as a human genome modification even if it does not affect his or her germ line

This subtle distinction has been extremely well addressed by Japanese bioethicist Tetsuya Ishii, who has written that: "the functional correction of a small mutation in the embryo via HDR [homology-directed repair] along with a short DNA template appears to be acceptable because this form of genetic modification can leave a wild-type gene, which is in a natural genetic state, and would fall outside of one of the ethical objections against germ line gene modification: transgression of the natural laws. The copying of a naturally occurring variant via HDR, along with a short DNA template, might be considered to be natural" (Ishii, 2015).

Therefore, we must be aware that the concept of changes in an individual genome and changes in the human genome do not necessarily match. It is true that no-one can introduce changes to the human species genome without altering someone's genome (since species are no more than the addition of all their members), but it is also true that someone might change the genome of a single individual or even an extended group of human beings (if his or her germ line is affected, for instance) without changing the human genome. Therefore, changing a human genome is not the same as changing the human genome. For example, if someone changes the pathological expression of the HTT gene that causes Huntington's disease to allow for normal expression, they would be

changing the genome of the affected human being, but not the human genome. Thus, "there could be cases of genetic enhancement when this practice would not alter the human nature, and as such, should not be morally prohibited." (Knoppers et al., 2017)

GGE should be implemented to all health issues, independent of whether they are serious or not

A part of the scientific literature has stated that, in general, we should only implement GGE technologies in those cases that deserve it, namely serious diseases (Kleiderman et al, 2019). My thesis argues that, if there comes a time when the use of genetic editing techniques involves minimal risk, we should not limit to those diseases. Indeed, I hold that we should have to use them for a wide range of pathologies. Moreover, this range would not necessarily include some of the most serious ones. There are several reasons why I think this is true:

1. In the case of some severe pathologies, such as monogenic diseases, we could continue to use preimplantation genetic diagnosis to prevent humans from being born with them. This would eliminate the need for genetic editing.
2. On the other hand, gene editing technologies could be used to modify genes in the case of polygenic diseases that may not be very serious, but that hinder our lives.
3. Finally, editing could be used to modify our genes in a way that decreases the likelihood of developing certain pathologies. Think, for example, of the possibility of altering the BRCAs genes.

Therefore, I conclude in one of my papers that the "serious" factor should not play an essential role in gene editing, provided that the risks involved in these technologies are acceptable. Furthermore, I argue that, if the level of risk were high, it would not be well understood why we would have to edit human embryos instead of continuing to use preimplantation genetic diagnosis to discard embryos suffering severe pathologies. Therefore, my conclusion is that, whatever the level of risk is, the "serious factor" should not play an essential role in the discussion on the applicability of genetic editing techniques (De Miguel, 2019b).

GGE should focus on absolute goods instead of positional goods

One of the most traditional discussions surrounding the acceptability, or otherwise, of GGE, has to do with its use for therapeutic or enhancement purposes. In my articles, I adopt a different perspective. In my opinion, we must concentrate, instead, in the distinction between positional and absolute goods. I define an absolute good as one that does not give its possessor any advantage over third parties, even if it satisfies a self-interest, such as health. On the other hand, a positional good is one that does create such an advantage, such as intelligence. One can be discriminated or feel like being so if he/she understands that there are other humans more intelligent than him/herself. In the case of health, however, this does not happen, or should not happen. Only very mean people are relieved to contemplate another human being with more serious health problems than their own. This is why, in my opinion, using GGE to improve people's absolute goods is perfectly acceptable, while using it to increase positional goods involves issues related to equity.

Keeping this in mind, I partially adhere to the appeal to obsolescence made by Sparrow (Sparrow, 2019), which is a solid argument when it is presented as a rejection of some forms of gene enhancement, rather than an objection to this technology as a whole. In fact, Sparrow's proposal impels us not to improve those traits we associate with purely positional goods, such as intelligence or beauty, for example. On the other hand, it does not pose a serious problem in the case of goods that are absolute, such as health, or that are only partially positional, such as goodness. In this sense, the argument provides good reasons for Savulescu's thesis (Persson and Savulescu, 2016) on the need to paralyse cognitive improvement and promote moral enhancement, while weakening Harris's position, which promotes the opposite (Harris, 2016). Somehow, I adhere to Sparrow's arguments in my paper on this topic, even though I reinforce the need to trace the distinction between positional and absolute goods.

Artificial Intelligence tools might be a game changer in terms of health care, but they also involve challenging issues

Diagnosis and clinical decision-making based on AI technologies are showing significant advances that may change the functioning of our health care systems. They promise more effective and efficient health care at a lower cost. Even though the evidence suggests that all these promises have yet to be demonstrated in clinical practice, it is undeniable that these technologies are already re-signifying the relationships in the health care landscape, particularly in the physician–patient relationship, which we can already redefine as a 'physician–computer–patient relationship' (Pearce, 2017).

In my thesis, I describe some of the most challenging issues at stake, mainly opacity and bias, which hinder the implementation of AI in the health care sector. From different perspectives, a significant number of contributions have exposed how implementing AI systems can pose problems in the flow of doctor–patient information. Inappropriate use of Machine Learning techniques might involve a dramatic loss in patients' rights to informed consent or not to be discriminated by their personal circumstances. Unfortunately, the traditional principles incorporated by medical law are insufficient for facing this challenge.

In one of the papers that are included in my thesis, I argue that the rights to information and explanation and the general prohibition on fully automated individual decision-making proclaimed by the General Data Protection Regulation may serve as an efficient mechanism to protect patient autonomy in this context. However, granting the rights to information and explanation may turn into an impossible task for practitioners. Therefore I hold that the development of new figures, such as the Health Information Counsellors, may be necessary. Finally, the eventuality of automating clinical decision-making poses dilemmas on the prevalence or not of physician judgment over machine output; in either case, patients should maintain control over their personal information and decision capacity, but further research is needed to provide consistent legal and ethical answers to these challenges (De Miguel et al, 2020b).

We need to keep in mind issues related to the lack of transparency

One of the main challenges with AI is that it encompasses a range of techniques that are very different from each other. Some are very simple, so it is easy for the controller to provide all the necessary information. Others, such as deep learning, suffer from serious problems in terms of transparency. This is often referred to as the 'black box' issue, which introduces the topic of opacity in the AI framework, a circumstance that renders transparency difficult to achieve. Indeed, opacity is one of the main threats against fair AI, since it directly defies the need for transparency. There are, at least, three types of opacity that are inherent in AI to a greater or lesser extent: (1) as intentional corporate or state secrecy; (2) as technical illiteracy and (3) epistemic opacity. Sometimes all these types contribute to create an opaque model. However, the GDPR grants the patient the right to be fully aware of the use of personal data collected by the controller if these data are to be used for automated decision-making purposes.

Of course, this means that we need to define what the right to information means in practice. In my thesis, I have supported the idea that the right to an explanation by no means implies that the data subject is empowered to have access to the algorithm as such. This would clearly render industrial secrecy impossible and would deprive the developer of the algorithm of any way to exploit the result of his/her investment commercially. However, this does not mean that patients cannot be provided with any form of relevant information. Indeed, there are some fruitful ways to guarantee that the explanation is sufficient to facilitate the exercise by patients of the rights granted to them by the GDPR and human rights law. To begin with, it is perfectly possible to provide a layperson with general information about how an algorithm has been constructed or what type of data categories it uses. Similarly, patients must be made aware of the importance of the contribution made by the AI system in the final decision, including receiving all available information on the main factors in the decision, whether changing a certain factor would or would not have changed the decision, and why different decisions are reached in similar-looking cases, or the same decision in different-looking cases (Doshi-Velez et al., 2017).

In one of the articles included in my thesis I hold that, on the one hand, this also means that, from the very first moment, patients should know about the use that might be made of their data and the foreseeable consequences of the data processing for this

purpose, as, indeed, is required by the Regulation; however, this requirement could be very limited in an actual scenario of big data analytics, where new data are created from inferred and derived data (De Miguel et al, 2020b). Looking at how the automated processing of data and profiling works, it is undeniably true that the GDPR focuses primarily on mechanisms to manage the input side of the processing, and that the legal mechanisms that address the outputs of the processing, including inferred and derived data, profiles, and decisions, are far weaker (Wachter & Mittelstadt, 2019). On the other hand, it also means that physicians and/or health care providers must explain to their patients the weight that automated decision-making and profiling represented in their final decision and provide understandable explanations of why the automated decision-makers suggestions were or were not followed. For this purpose, a flexible, functional approach will be most appropriate for understanding the term 'meaningful information' that is included in the right to an explanation (Selbst & Powles, 2017).

There should be a right to refuse diagnostics and treatment planning by artificial intelligence

Recently, some authors have considered that the use of artificially intelligent (AI) systems for medical diagnostics and treatment planning might cause rejection to some patients, due to their medical and non-medical preferences. For instance, as long as AI systems do not have this ability to engage in open-ended, meaningful conversation with patients about their preferences, some patients may have a good reason to prefer the involvement of physicians (Ploug and Holm, 2019)

In one of the articles that are included in my thesis, I consider that respect for health care principles and values is a strong enough reason to conclude that we need to protect the patients' right to refuse diagnostics and treatment planning by artificial intelligence (AI). However, I believe that there are exceptions to this general rule. If a treatment is particularly costly, for example, it should not be administered without first having recourse to the advice of the AI if the efficiency of the corresponding predictive algorithm had been demonstrated. Before applying this right, one should also consider possible harm to third parties, who would be deprived of adequate care as a consequence. In my opinion, if AI were able to suggest an efficient form of allocation, we should not allow the right to refuse treatment planned by AI to deny scarce health resources to patients

who are able to benefit from it. Thus, the scope of the right should be limited, at least, under three circumstances: (1) if it is against a physician's obligation to not cause unnecessary harm to a patient or to not provide futile treatment, (2) in cases where the costs of implementing this right are too high, or (3) if recognizing the right would deprive other patients of their own rights to adequate health care.

The use of automated decision making in wearables creates relevant health care issues

Some new technological tools, such as wearables, are providing health care givers with innovative ways to control and monitor their patients' health. Digital pills constitute an excellent example to this purpose. These are complex systems that include a drug and an electronic tracker that is activated when the patient takes the pill. Therefore, they can be appropriate to measure adherence to a treatment. However, they involve a number of challenging issues.

In my thesis I hold that the use of digital pills for monitoring patient adherence constitutes a radical turn in the way we focus this theme. Our current healthcare system is built on a mentality in which trust between clinicians, caregivers, or social workers and the patient is a fundamental piece. The introduction of digital pills replaces this framework with a new policy in which monitoring and control play a key role. It is no longer the patients who reveal data to the physician on a voluntary basis. Instead, the physician becomes a kind of 'Big Brother' who knows everything about the patient, even if they are unwilling to share such information. Furthermore, this causes erosion of the personal relationship between patients and their physicians. By now, patients usually discuss with their doctor the problems arising from the follow-up of the prescribed treatment. Nevertheless, in the new scenario, patients somehow become the objects of inspection by the health system, which watches closely for any deviation from the correct administration of the treatment.

It is very important that patients have sufficient confidence in their doctors to discuss with them the reasons why they are reluctant to take the medication prescribed. It is also essential that the system provides with the possibility of building that relationship through adequate means. Thus, with the use of such smart devices, trust would be compromised both from the professional's and patients' perspectives. First, data generated by the device may cast doubt on the truthfulness of the patient's self-report.

Conversely, patients may distrust physicians and their therapeutic recommendations if they receive a different diagnosis from that suggested by the device on which they rely (Ho and Quick, 2018).

On the other hand, digital pills open a major gateway to distant and mediated interaction between doctors and patients, thereby decreasing the need for face-to-face communication. Finally, the use of these new technologies may over-technify the monitoring of treatment or decision-making about a patient. This new scenario, which constitutes a serious challenge in the health care arena, is by no means inevitable, but it requires the adoption of an appropriate mentality and measures capable of preventing it. It is essential to keep in mind that technologies should serve to enhance the physician–patient relationship, rather than to replace it. For the sake of maximizing the usefulness of these cutting-edge medical technologies in the way we conceive medicine of even the integral care of patients, we must make a proper use of them in terms of both safety and confidence, otherwise, a key aspect of the patient–physician relationship would be broken: trust. The question, in short, is whether the possible increase in adherence to treatment would compensate for the decrease in this fundamental value, confidence, if patients were forced to use this new technology. In my thesis I held that this is not the case. That is why I advocate for a system that is respectful with patients' autonomy and that only allows the use of digital pills in cases in which the patient encourages it, unless the defense of a public good, such as health or safety, makes it essential (De Miguel and Morla, 2020c).

Fuentes referenciadas (más allá de los artículos que componen esta tesis)

Andorno, Roberto, Françoise Baylis, Marcy Darnovsky, Donna Dickenson, Hille Haker, Katie Hasson, Leah Lowthorp et al. "Geneva statement on heritable human genome editing: the need for course correction." *Trends in biotechnology* 38, no. 4 (2020): 351-354, available at: https://www.cell.com/trends/biotechnology/fulltext/S0167-7799(19)30317-8#articleInformation

Constam, Daniel B. "Comment on “Human dignity and gene editing”." EMBO reports 20, no. 1 (2019): e47220.

Coyle, Fiona, and John Fairweather. "Space, time and nature: exploring the public reception of biotechnology in New Zealand." Public Understanding of Science 14, no. 2 (2005): 143-161.

D Miguel Beriain I, Penasa S. The embryo survival criterion: a moral obligation or a eugenic practice? Rom J Leg Med. 2018; 26(2): 212- 217.

Doshi-Velez, Finale, Mason Kortz, Ryan Budish, Chris Bavitz, Sam Gershman, David O'Brien, Kate Scott et al. "Accountability of AI under the law: The role of explanation." arXiv preprint arXiv:1711.01134 (2017), at: https://dash.harvard.edu/bitstream/handle/1/34372584/2017-11_aiexplainability-1.pdf

Harris, John. How to be good: The possibility of moral enhancement. Oxford University Press, 2016.

Ho, Anita, and Oliver Quick. "Leaving patients to their own devices? Smart technology, safety and therapeutic relationships." BMC medical ethics 19, no. 1 (2018): 1-6. https://doi.org/10.1186/s12910-018-0255-8

Ishii, Tetsuya. "Germ line genome editing in clinics: the approaches, objectives and global society." Briefings in functional genomics 16, no. 1 (2017): 46-56.

Kant I (1785) Groundwork for the metaphysic of morals, in the version by Jonathan Bennett presented at www.earlymoderntexts.com (2008). Available at: https://www.earlymoderntexts.com/assets/pdfs/kant1785.pdf

Kleiderman, Erika, Vardit Ravitsky, and Bartha Maria Knoppers. "The ‘serious’ factor in germline modification." Journal of medical ethics 45, no. 8 (2019): 508-513.

Knoppers, Bartha Maria, Rosario Isasi, Timothy Caulfield, Erika Kleiderman, Patrick Bedford, Judy Illes, Ubaka Ogbogu, Vardit Ravitsky, and Michael Rudnicki. "Human gene editing: revisiting Canadian policy." (2017): 1-2.

MacKellar, Calum, and Christopher Bechtel, eds. The ethics of the new eugenics. Berghahn Books, 2014.

MedicineNet.com (2016). Definition of the human genome. Retrieved May 2, 2016 from http://www.medicinenet.com/script/main/art.asp?articlekey=3818

of Sciences, National Academy, and National Academies of Sciences, Engineering, and Medicine. "Human Genome Editing: Science, Ethics, and Governance." (2017).Bioethics, NCo. "Genome Editing and Human Reproduction: social and ethical issues." London: Nuffield Council on Bioethics (2018).

O'neil, Cathy. Weapons of math destruction: How big data increases inequality and threatens democracy. Crown, 2016.

Pearce, Christopher. "Computers, patients, and doctors—theoretical and practical perspectives." In Health Professionals' Education in the Age of Clinical Information Systems, Mobile Computing and Social Networks, pp. 5-22. Academic Press, 2017.

Persson, Ingmar, and Julian Savulescu. "Moral bioenhancement, freedom and reason." Neuroethics 9, no. 3 (2016): 263-268.

Ranisch, Robert. "Germline genome editing versus preimplantation genetic diagnosis: Is there a case in favour of germline interventions?." Bioethics 34, no. 1 (2020): 60-69.

Saha, Krishanu, J. Benjamin Hurlbut, Sheila Jasanoff, Aziza Ahmed, Anthony Appiah, Elizabeth Bartholet, Françoise Baylis et al. "Building capacity for a global genome editing observatory: institutional design." Trends in biotechnology 36, no. 8 (2018): 741-743.

Selbst, Andrew, and Julia Powles. "“Meaningful Information” and the Right to Explanation." In Conference on Fairness, Accountability and Transparency, pp. 48-48. PMLR, 2018.

Sparrow, Robert. "Yesterday’s child: How gene editing for enhancement will produce obsolescence—and why it matters." The American Journal of Bioethics 19, no. 7 (2019): 6-15.

Wachter, Sandra, Brent Mittelstadt, and Luciano Floridi. "Why a right to explanation of automated decision-making does not exist in the general data protection regulation." International Data Privacy Law 7, no. 2 (2017): 76-99.

Welz, Claudia. Humanity in god's image: an interdisciplinary exploration. Oxford University Press, 2016.

SECTION II: CONCLUSIONS (English version)

1. **GGE should not be performed until the risks associated to this technology are acceptable.** At the present moment, these are too high, but this might change in the future. However, if we are considering using GGE for research purposes, the situation is totally different, since there are no persons affected by the consequences of the research.

2. **If GGE is ever applicable, it should be performed in all cases, not only in serious cases.** This is due to the fact that once the technology is safe, there are no good reasons not to benefit a child if possible be (at least, if only absolute goods are affected, see conclusion number 7.

3. **GGE is not against human dignity.** We are not obliged to preserve the human genome as it is. The human genome is not sacred. Trying to fix it makes no sense from a biological perspective. Otherwise, we should indeed proceed with germline gene editing so as to undo the changes caused by natural selection.

4. **Human dignity does indeed support GGE.** Dignity is directly related to autonomy. Once we are able to edit genes, action or omission are not key issues. The most important factor is whether we treat someone as a mere means. If we can improve someone's health through this technology and we refuse to do it on the basis of the defense of dignity, we would be treating this human being as a mere means. Thus, human dignity supports GGE.

5. **One must consider that changing an embryo's genes does not necessarily mean changing the human genome.** This would only coincide if we introduce new gene expressions in our genome. If we are only substituting a pathological variant for a non-pathological one, this would not change the human genome as such.

6. **Heritable human genome editing is much more advisable from a moral point of view than the Preimplantation Genetic Diagnosis plus embryo selection practices.** Contrary to PGD, GGE serves to avoid discarding embryos. Instead, it

"cures" such offspring by changing the pathologic biological material. Furthermore, GGE is much appropriate in terms of capability to improve someone's health.

7. **GGE should be limited to absolute goods, such as health or morality.** We should not use it if positional goods are involved. Otherwise, we could be creating an impressive social issue in terms of equity.

8. Diagnosis and clinical decision-making based on AI technologies are already re-signifying the relationships in the health care landscape, particularly in the physician–patient relationship, which we can already redefine as a **'physician–computer–patient relationship'. This might be extremely challenging in terms of patients' and health care givers' rights and freedoms.**

9. The eventuality of automating clinical decision-making poses some dilemmas on the prevalence of physicians' judgment over machine outputs; in either case, **patients should maintain control over their personal information and decision capacity**. However, this is hard to implement in practice. Further research is needed to provide consistent legal and ethical answers to these challenges.

10. Essential information should be provided to the patient if an AI tool is used. This should include **information on how the algorithm has been constructed, what type of data categories it uses, as well as the weight of the contribution made by the AI system in the final decision.**

11. Granting the rights to information and explanation may turn into an impossible task for practitioners. T**he development of new figures, such as the Health Information Counsellors, might be very helpful to avoid these issues and improve the performance by the system.**

12. We should recognize patients their right to r**efuse diagnostics and treatment by artificial intelligence** by virtue of values such as social pluralism or individual autonomy.

13. **The scope of such a right should be limited, at least, under three circumstances**: (1) if it is against physicians' obligation not to cause unnecessary harm to patients or to not provide futile treatments, (2) in cases where the costs of implementing this

right are too high, or (3) if recognizing the right would deprive other patients of their own rights to adequate health care.

14. N**ew tools such as wearables might cause a dramatic erosion of the personal relationship between patients and their physicians**. Nowadays, patients usually discuss with their doctor the problems arising from the follow-up of the treatments prescribed. Nevertheless, in the new scenario, patients somehow would become the objects of inspection by the health system, which watches closely for any deviation from the correct administration of the treatments.

CONCLUSIONES (versión en castellano)

1. La edición de la línea germinal (GGE) no debería realizarse hasta que los riesgos asociados a esta tecnología sean aceptables. En la actualidad, éstos son demasiado elevados, pero esto podría cambiar en el futuro. Sin embargo, si nos planteamos utilizar la GGE con fines de investigación, la situación es totalmente diferente, ya que no habría personas afectadas por las consecuencias de la investigación.

2. Si alguna vez se aplica la GGE, debería realizarse en todos los casos, no sólo en los casos graves. Esto se debe al hecho de que una vez que la tecnología es segura, no hay buenas razones para no beneficiar a un niño si es posible (al menos, si sólo se ven afectados los bienes absolutos, véase la conclusión número 7).

3. La GGE no va en contra de la dignidad humana. No estamos obligados a preservar el genoma humano tal y como es. El genoma humano no es sagrado. Intentar arreglarlo no tiene sentido desde una perspectiva biológica. De lo contrario, deberíamos proceder a la edición de genes en la línea germinal para deshacer los cambios causados por la selección natural.

4. La dignidad humana sí es compatible con la edición génica germinal. La dignidad está directamente relacionada con la autonomía. Una vez que seamos capaces de editar genes, la acción o la omisión no son cuestiones clave. El factor más importante es si tratamos a alguien como un mero medio. Si podemos mejorar la salud de alguien mediante esta tecnología y nos negamos a hacerlo basándonos en la defensa de la dignidad,

estaríamos tratando a este ser humano como un mero medio. Por lo tanto, la dignidad humana apoya la GGE.

5. Hay que tener en cuenta que cambiar los genes de un embrión no significa necesariamente cambiar el genoma humano. Esto sólo coincidiría si introducimos nuevas expresiones genéticas en nuestro genoma. Si sólo sustituimos una variante patológica por otra no patológica, esto no cambiaría el genoma humano como tal.

6. La edición hereditaria del genoma humano es mucho más recomendable desde un punto de vista moral que las prácticas de Diagnóstico Genético Preimplantatorio más selección de embriones. Al contrario que el DGP, el GGE sirve para evitar el descarte de embriones. En su lugar, "cura" esa descendencia cambiando el material biológico patológico. Además, el GGE es mucho más apropiado en términos de capacidad para mejorar la salud de alguien.

7. La GGE debe limitarse a los bienes absolutos, como la salud o la moralidad. No deberíamos utilizarla si se trata de bienes posicionales. De lo contrario, podríamos estar creando un problema social impresionante en términos de equidad.

8. El diagnóstico y la toma de decisiones clínicas basados en tecnologías de IA ya están resignificando las relaciones en el panorama de la atención sanitaria, especialmente en la relación médico-paciente, que ya podemos redefinir como una "relación médico-ordenador-paciente". Esto podría suponer un gran reto en cuanto a los derechos y libertades de los pacientes y de los profesionales sanitarios.

9. La posibilidad de automatizar la toma de decisiones clínicas plantea algunos dilemas sobre la prevalencia del juicio de los médicos sobre los resultados de las máquinas; en cualquier caso, los pacientes deberían mantener el control sobre su información personal y su capacidad de decisión. Sin embargo, esto es difícil de aplicar en la práctica. Es necesario seguir investigando para dar respuestas legales y éticas consistentes a estos desafíos.

10. Si se utiliza una herramienta de IA, debe proporcionarse al paciente información esencial. Esto debería incluir información sobre cómo se ha construido el algoritmo, qué tipo de categorías de datos utiliza, así como el peso de la contribución del sistema de IA en la decisión final.

11. La concesión de los derechos de información y explicación puede convertirse en una tarea imposible para los profesionales. El desarrollo de nuevas figuras, como los Consejeros de Información Sanitaria, podría ser muy útil para evitar estos problemas y mejorar el rendimiento del sistema.

12. Deberíamos reconocer a los pacientes su derecho a rechazar diagnósticos y tratamientos mediante inteligencia artificial en virtud de valores como el pluralismo social o la autonomía individual.

13. El alcance de ese derecho debería limitarse, al menos, en tres circunstancias (1) si va en contra de la obligación de los médicos de no causar daños innecesarios a los pacientes o de no proporcionar tratamientos fútiles, (2) en los casos en que los costes de aplicación de este derecho sean demasiado elevados, o (3) si el reconocimiento del derecho privara a otros pacientes de sus propios derechos a una atención sanitaria adecuada.

14. Las nuevas herramientas, como los wearables, podrían causar una dramática erosión de la relación personal entre los pacientes y sus médicos. Hoy en día, los pacientes suelen discutir con su médico los problemas derivados del seguimiento de los tratamientos prescritos. Sin embargo, en el nuevo escenario, los pacientes se convertirían de alguna manera en objeto de inspección por parte del sistema sanitario, que vigila de cerca cualquier desviación de la correcta administración de los tratamientos.

SECCION III: ANEXO. Artículos publicados en las diferentes revistas-

Índice, con los impactos correspondientes

2018

de Miguel Beriain, I., "Human dignity and gene editing: Using human dignity as an argument against modifying the human genome and germline is a logical fallacy", *EMBO Reports*, October 2018 (EMBO reports (2018) 19, e46789) Indizado en JCR Thomson. Impact factor (2018): 8.383. Five years impact: 9.214. Q1/D1 (25/299) in Biochemistry and Molecular Biology. Q1 (26/193) in Cell Biology. ISSN: 1469-221X.
DOI: 10.15252/embr.201846789
Páginas: 1-4 compilación

2019

1. de Miguel Beriain, Iñigo. "Should human germ line editing be allowed? Some suggestions on the basis of the existing regulatory framework." *Bioethics* 33, no. 1 (2019a): 105-111. 5 year IF: 1.787. IF (2019): 1.799. Q2 in Medical Ethics (5/16). DOI: 10.1111/bioe.12492.

Páginas: 5-26 compilación

2. de Miguel Beriain, Iñigo. "Obsolescence is not a good reason to oppose all types of enhancement." *American Journal of Bioethics* 19, no. 7 (2019). IF (2019b): 7.647. Q1 in Ethics; Q1 in Social Issues; Q1 in Social Sciences, Biomedical.

DOI: 10.1080/15265161.2019.1618950).

Páginas: 27-32 compilación

3. de Miguel Beriain, Iñigo, and Tetsuya Ishii. "Comment on" Should gene editing replace embryo selection following PGD? Some comments on the debate held by the International Society for Prenatal Diagnosis"." *Prenatal diagnosis* 39, no. 12 (2019c): 1170-1172. JIF (2019): 2.425. Q2 in Obstetrics & Gynecology (29/82)

DOI: 10.1002/pd.5542

Páginas: 33-37 compilación

4. Beriain, Iñigo De Miguel. "Is the 'serious' factor in germline modification really relevant? A response to Kleiderman, Ravitsky and Knoppers." *Journal of medical*

ethics 46, no. 2 (2020): 151-152. doi: 10.1136/medethics-2019-105744. Pub Date : 2019-10-17 (2019d), IF (2019): 2.021. Q1 in Medical Ethics (4/16).

DOI: 10.1136/medethics-2019-105744

Páginas: 38-43 compilación

5. Ishii, Tetsuya, and Iñigo de Miguel Beriain. "Safety of germline genome editing for genetically related "future" children as perceived by parents." *The CRISPR journal* 2, no. 6 (2019): 370-375. IF (2019e): 5.343. Q1 in Genetics and Heredity (25/178).

DOI: https://www.liebertpub.com/doi/full/10.1089/crispr.2019.0010.

Páginas: 44-50 compilación

6. de Miguel Beriain, Íñigo. "The Geneva Statement on Heritable Human Genome Editing: A Criticism." Trends in Bioetchnology, VOLUME 39, ISSUE 3, P219-220, MARCH 01, 2021 IF (2019e): 14.343. Q1 in Biotechnology and Applied Microbiology.

DOI: 10.1016/j.tibtech.2020.11.004.

Páginas: 50-54 compilación

2020

7. de Miguel Beriain, Iñigo, and Begoña Sanz. "Human dignity and gene editing: Additional support for Raposo's arguments." *Journal of bioethical inquiry* 17, no. 2 (2020a): 165-168. IF (2019): 1.425. Q3 in Medical Ethics (10/16).

DOI: 10.1007/s11673-020-09969-8.

Páginas: 55-61 compilación

8. de Miguel, Iñigo, Begoña Sanz, and Guillermo Lazcoz. "Machine learning in the EU health care context: exploring the ethical, legal and social issues." *Information, Communication & Society* 23, no. 8 (2020b): 1139-1153. 5 year IF: 4.793. Q1 in Sociology. Q1 in Communication.

DOI: https://doi.org/10.1080/1369118X.2020.1719185

Páginas: 62-88 compilación

9. de Miguel Beriain, Iñigo. "Should we have a right to refuse diagnostics and treatment planning by artificial intelligence?." *Medicine, Health Care and Philosophy* (2020c): 1-6. IF (2019): 1.708. Q1 in History and Philosophy of Science (8/48). Q2 in Ethics (15/55).

DOI: https://doi.org/10.1007/s11019-020-09939-2

Páginas: 89-99 compilación

10. de Miguel Beriain, Iñigo, and Marina Morla González. "'Digital pills' for mental diseases: an ethical and social analysis of the issues behind the concept." *Journal of Law and the Biosciences* 7, no. 1 (2020d): lsaa040. Q1 in Medicine Legal 2/16, IF (2020): 3.583. 5 year IF: 2.568. Q1 in Medical Ethics (3/16). Q1 in Medicine, Legal (2/16).

DOI: https://doi.org/10.1093/jlb/lsaa040

Páginas: 100-120 compilación

Science & Society

EMBO reports

Human dignity and gene editing

Using human dignity as an argument against modifying the human genome and germline is a logical fallacy

Iñigo de Miguel Beriain[1,2]

Human germline editing raises a number of essential ethical issues that have spawned intense debate. According to some bioethical arguments and legal documents, germline modification would threaten human dignity, since they consider the human genome as the physical representation [of human dignity]. This article will highlight the inherent contradictions of these arguments and argue that claims that germline editing would violate human dignity are invalid.

> *"... respect for human dignity should actually support arguments to continue with genetic editing of embryos."*

The emergence of gene editing technology—particularly CRISPR-Cas9—and the first experiments to modify the genome of human embryos have given rise to an intense ethical debate. Such an in-depth discussion of the potential ethical, societal and medical implications is indeed highly relevant as modifications of the germline would not only affect individual patients or humans but the human species as a whole. In fact, the debate began even before the availability of the CRISPS/Cas system when the first gene therapies were developed to cure a few select pathologies. Time has not brought any general agreement, and a universal consensus on whether or not to allow human germline editing is still remote.

The debate

While some commentators call for a total ban on any form of gene editing that affects the human germline [1], others advocate for a moratorium until the risks have been sufficiently addressed and resolved [2]. Other authors even claim that germline editing should be considered a moral imperative to improve the human species [3]. Similar disagreements exist about the ends to which these techniques should be applied. Some consider that only therapeutic purposes are acceptable; others support their use for human enhancement, a view that many bioethicists and most international declarations and conventions consider as a form of eugenics.

The aim of this article was not to resolve this debate, but to focus on the idea of human dignity and its role in this discussion. Most authors simply assume that respect for human dignity is a crucial argument for opposing germline editing [4]. Moreover, even those who disagree with this view have not questioned whether we should accept this postulate as a general moral rule [5].

This paper attempts to show that, in contrast to what is generally accepted, respect for human dignity should actually support arguments to continue with genetic editing of embryos. This article will justify this statement by demonstrating, first, that the arguments in favour of prohibiting genetic manipulation of the germline on the basis of human dignity are not solid. Second, it will explain how the opposite conclusion seems far more reasonable, whatever interpretation of the idea of human dignity we adopt.

Establishing the terms of the debate

A clear definition of human dignity remains elusive, because the concept of dignity is, in itself, extremely problematic to the extent that some people consider that it should be completely omitted from bioethics [6]. For the sake of the argument that human dignity should not be an obstacle to genome editing, it is not necessary to delve too deep into this controversy. It is enough to summarize that those who believe in the importance of human dignity and its relation to the human genome generally share three main ideas.

> *"... there is a broad tendency in bioethics to directly link the notion of human dignity to the human genome"*

The first is that dignity—which is linked to the concept of autonomy—is an intrinsic and non-negotiable value shared by all human beings, simply by virtue of belonging to the human species. This idea has deep roots in our collective imagination. For instance, it is included in the Universal Declaration of Human Rights of 10 December 1948 whose article 1 states "All human beings are born free and equal in dignity and rights".

Second, advocates of the idea of human dignity believe that since every human being has the same intrinsic moral value, every person should be treated as though he or she was an end in his or herself and not a mere means. This fundamental idea is based on Kant's moral imperative. Even those who oppose the concept of dignity generally agree—perhaps with the exception of pure utilitarianists—that we cannot instrumentalize human beings. By way of example, the Council of Europe Convention on Action against Trafficking in Human Beings of 16 May 2005 (Warsaw Convention) states in its

1 Chair in Law and the Human Genome RG, Department of Public Law, University of the Basque Country (UPV/EHU), Leioa, Spain. E-mail: inigo.demiguelb@ehu.eus
2 IKERBASQUE, Basque Foundation for Science, Bilbao, Spain
DOI 10.15252/embr.201846789 | EMBO Reports (2018) 19: e46789 | Published online 20 September 2018

Preamble "trafficking in human beings constitutes a violation of human rights and an offence to the dignity and the integrity of the human being".

Third, there is a broad tendency in bioethics to directly link the notion of human dignity to the human genome. Since it is the genome that determines who belongs to the human species and since being part of the human species confers dignity, it seems reasonable to link human dignity to the human genome. For example, the Universal Declaration on the Human Genome and Human Rights adopted by the General Conference of UNESCO on 11 November 1997 states "The human genome underlies the fundamental unity of all members of the human family, as well as the recognition of their inherent dignity and diversity. In a symbolic sense, it is the heritage of humanity" (Article 1).

> *"A major fallacy of arguments based on preserving the human genome is ignoring the fact that the human genome is not a fixed entity…"*

Having established these fundamental premises, we can now explore the role that the idea of dignity should play with respect to human germline editing. The argument will focus first on one of the main demands by those who wish to ban the application of gene editing in humans: the need to preserve the human genome. The article will then analyse whether a distinction between direct and indirect actions that cause germline changes is reasonable. Finally, it will discuss the implications of banning therapeutic applications in an embryo from the perspective of human dignity.

The need to preserve the human genome

As previously noted, various authors believe that human dignity is intrinsically linked to the human genome. The question then is as follows: What are the normative implications of this belief? The general opinion among opponents of germline editing is that since our dignity is embedded in our genome, we must refrain from altering [7]. This idea has received support from the Declaration on the Human Genome and Human Rights (11 November 1997), whose article 24 reads: "The International Bioethics Committee of UNESCO should contribute to the dissemination of the principles set out in this Declaration and to the further examination of issues raised by their applications and by the evolution of the technologies in question. It should organize appropriate consultations with parties concerned, such as vulnerable groups. It should make recommendations, in accordance with UNESCO's statutory procedures, addressed to the General Conference and give advice concerning the follow-up of this Declaration, in particular regarding the identification of practices that could be contrary to human dignity, such as germ-line interventions".

In my opinion, this normative proposal is problematic, because it contradicts the axiom on which it is based: if one believes that our dignity can only remain undiminished if the human genome remains unchanged, we cannot tolerate any new mutation in the human genome—regardless of whether it is produced by human action or naturally. A major fallacy of arguments based on preserving the human genome is ignoring the fact that the human genome is not a fixed entity, an immutable biological substrate, but subject to mutations like any other genome: otherwise, evolution would come to a halt. Even more so, if we claim that respect for human dignity means preserving the human genome in its current state, the logical conclusion would be to not only to renounce sexual reproduction in favour of cloning—as it generates mutations—but even to use gene editing to remove any natural, random mutations because it might violate human dignity and autonomy.

Some would argue that this argument is a sophism: that respect for human dignity does not entail preserving the human genome at all costs; but that humans should not arrogate to themselves the power to alter it [4]. But such a rejoinder is at least suspect of a religious or naturalistic component as it implies that only nature or a creator should have dominion over the human genome. The non-metaphysical alternative in support of this argument is the view that "tried and tested" natural processes are more reliable than "human tinkering" [8]. Both assumptions do not hold up to scrutiny.

There is nothing to demonstrate a presumed natural "goodness" or "intelligence". On the contrary, it seems reasonable to think that editing the human genome to prevent a lethal or devastating disease poses less risk to the future of human beings than changes introduced by the vagaries of random genetic crosses or mutation. Moreover, abstaining from intervention in the human genome when action is possible is already a form of intervention by omission. The emergence of genetic editing means that, whether we like it or not, the human genome will end up being what we decide it to be, by action or by omission. Finally, the idea of dignity is linked to autonomy, that is, the ability to set rules. Trying to renounce this power once it is within our reach is not only almost impossible—knowledge cannot be erased or suppressed—but also implies a refusal to assume the very principle that guarantees human dignity: the ability to impose reasonable human norms on irrational, random nature: a renunciation of autonomy, so to speak.

> *"… even if we were to accept that human dignity and the human genome are linked […] this would not necessarily imply an obligation to preserve the human genome."*

In conclusion, even if we were to accept that human dignity and the human genome are linked—which is debatable—this would not necessarily imply an obligation to preserve the human genome. Consequently, there is no reason why we should not proceed to alter the human genome, since it is not a fixed and immutable entity that has to be preserved for the better of humanity. But if we believe the opposite, that we need to conserve the momentary version of the human gene pool, we would have a moral obligation to use genetic editing techniques to reverse the changes brought about by nature. Whatever choice, the belief in a link between the human genome and dignity is not helpful to discuss whether we should allow or ban editing of the human genome.

Changes in the germline

Consider again that we accept the premise that respect for human dignity implies that we must refrain from altering the human

genome. If this is the case, should we not then advocate the abolition of any action that would have this effect? Those who argue for the need to preserve the human genome draw a subtle distinction between direct actions that deliberately introduce changes such as gene therapy or therapeutic gene editing and actions that introduce mutations indirectly as a side effect, such as radio- or chemotherapy that carry a risk of generating mutations of germline cells.

Indeed, a sizeable number of ethicists and some of our main regulatory instruments argue that direct actions to change the human genome are morally reprehensible and must therefore be illegal. However, interventions that tolerate such changes may be morally acceptable for the sake of defending a greater good. The best example of this principle is the Explanatory Report on the Oviedo Convention, which states: "92. On the other hand the article does not rule out interventions for a somatic purpose, which might have unwanted side-effects on the germ cell line. Such may be the case, for example, for certain treatments of cancer by radiotherapy or chemotherapy, which may affect the reproductive system of the person undergoing the treatment".

However, if we accept that our dignity is linked to the human genome and that any change in the latter affects the former, it is not clear why the type of action—direct or indirect—and not its result should be relevant. Is it really pivotal whether the genome has been altered directly or as an inevitable consequence of the action? If human dignity is compromised by the alteration of the human genome, then it clearly does not make a difference. But if so, why should we accept one type of alteration and reject the other?

The fundamental problem is that such conclusions are based on theories of double effect but do not interpret them properly and end up contradicting the importance that human dignity should have. Defenders of involuntary alteration of the human genome argue that if an action is directed at a good end, that is, to save the life of a cancer patient, and that if negative consequence—alteration of the genome or the germline—occurs unintentionally, it is legitimate. However, this notion forgets that the double effect argument includes, as a condition, that the negative consequence must be of minor importance, or at least of proportionally less important than the intended positive effect (literally: "The good effect must be sufficiently desirable to compensate for the allowing of the bad effect" [9]). But when the action causes a serious harm to human dignity, it is impossible to apply the principle of double effect, at least if we accept that respect for dignity is an absolute requirement and not a principle that can be waived if circumstances advise so.

Therefore, if we believe that any change in the genome caused by human action is a breach of human dignity, we must refrain from any kind of action that could lead to this consequence whether directly or indirectly—even if this means abolishing radio- or chemotherapy. The fact that this appeal does not arise often is a clear inconsistency that can be justified only on pragmatic grounds. But this justification is an unacceptable trick when human dignity is at stake. Last but not least, a consciously caused alteration of the human genome is much closer to the idea of autonomy than is random change brought about by chemotherapy. Again, inconsistencies in the more conservative discourse appear to give a poor answer to this question. Therefore, if we are to accept actions that alter the human genome, even if indirectly, then we have plenty of reasons to consider that those intended to directly produce such a result should also be considered morally acceptable on the basis of human dignity.

Dignity of the embryo

If, in spite of this conclusion, we should persist to veto any alteration of the human genome, it would not necessarily mean refraining from changing the genome of an embryo. This is for one simple reason: there is a big difference between changing the genome of a human being and changing the human genome. This is not dependent on whether it is the genome of one or several human beings, or whether the change occurs in the somatic or germinal line. It depends on the type of alteration. If the final result of the intervention—for instance, replacing a mutated gene to restore its original function—does not introduce any novelty into the human gene pool, then it is inaccurate to speak of an alteration in the human genome.

Imagine, for example, a human embryo with mutations of the Huntington gene that will inevitably lead to Huntington's disease if the embryo grows into a human being. If we edit its germline to replace the gene with a normal variant, we will modify the embryo's genome but not the human genome. The ultimate result of the intervention—a human being with a genome that does not show the specific pathological variant that triggers Huntington's disease—will not introduce any novelty into the human gene pool. Therefore, any appeal not to carry out such intervention on the basis of the defence of the human genome and, hence, human dignity, would be manifestly illogical.

This point could, however, be contested by the argument that although such intervention in itself does not pose a direct threat to human dignity, it could constitute an indirect threat—for example, by enabling other interventions that do alter the human gene pool or even redefine fundamental social relationships (between parents and children, individuals and communities, citizens and states), and associated notions of responsibility and care [10].

> *"... respecting human dignity means putting the individual human and his or her interests before those of his or her possible descendants, social group or the like."*

The problem with this kind of argument is that it is disrespectful to what it claims to defend: human dignity. What the argument contends is the following: although we can cure a serious pathology in an embryo or in a human being, we should refrain from doing so in order to preserve higher social values or goods. But is this kind of assertion in line with the fundamental normative principle that follows from the idea of dignity? Is this the proper way to treat an embryo, as an end in itself and not as a mere means to preserve a higher good? Or is this suggestion more like a social eugenics exercise based on the assumption that the interests of the group should prevail over those of the individual? The answer is clear: respecting human dignity means putting the individual human and his or her interests before those of his or her possible descendants, social group or the like. This is laid down in the Universal Declaration on Bioethics and Human Rights of 19 October 2005, whose article 3.2 reads: "The interests and welfare of the individual should

have priority over the sole interest of science or society", which reproduces article 2 of the Oviedo Convention: "The interests and welfare of the human being shall prevail over the sole interest of society or science". Therefore, respect for human dignity would actually require that we intervene in the human germline in order to preserve the interests of a particular human being, disregarding the collective interest.

"Given the logical fallacies, it seems about time to give up on a notion of human dignity and autonomy that is closely linked to the human genome…"

The overall conclusion is that there are no good reasons to justify a general ban on genetic editing of the human germline on the basis of human dignity. On the contrary, it is precisely this factor that should prompt us to use genome editing in the interests of the individual human being. Given the logical fallacies, it seems about time to give up on a notion of human dignity and autonomy that is closely linked to the human genome and consider each of these separately when discussing the ethical implications of human genome and germline editing.

Acknowledgements

This work was supported by a Grant from the Government of the Basque Country, Reference IT-1066-16, and two Grants from the Spanish Ministry of Economy, Industry and Competitiveness (MINECO), Projects: Functional and Clinical Impact of the Genome Analysis in CLL and DECADE. Discapacidad, Enfermedad Crónica y Accesibilidad a los Derechos, DER2016-80138-R.

References

1. Annas GJ, Andrews LB, Isasi RM (2002) Protecting the endangered human: toward an international treaty prohibiting cloning and inheritable alterations. *Am J Law Med* 28: 151 – 178
2. Lanphier E, Urnov F, Haecker SE, Werner M, Smolenski J (2015) Don't edit the human germ line. *Nature* 519: 410 – 411
3. Savulescu J, Pugh J, Douglas T, Gyngell C (2015) The moral imperative to continue gene editing research on human embryos. *Protein Cell* 6: 476 – 479
4. Habermas J (2003) *The future of human nature*. Malden, MA: Polity
5. Sykora P, Caplan A (2017) The Council of Europe should not reaffirm the ban on germline genome editing in humans'. *EMBO Rep* 18: 1871 – 1872
6. Macklin R (2003) Dignity is a useless concept. *BMJ* 327: 1419 – 1420
7. Nuffield Council on Bioethics (2018) Genome editing and human reproduction: social and ethical issues. August 2018 http://nuffieldbioethics.org/project/genome-editing-human-reproduction
8. Coyle F, Fairweather J (2005) Space, time and nature: exploring the public reception of biotechnology in New Zealand. *Public Underst Sci* 14: 143 – 161
9. McIntyre A (2014) Doctrine of double effect. In *The Stanford encyclopedia of philosophy*, Zalta EN (ed.). Stanford, CA: Metaphysics Research Lab, Stanford University. https://plato.stanford.edu/
10. Saha K, Hurlbut JB, Jasanoff S, Ahmed A, Appiah A, Bartholet E, Baylis F, Bennett G, Church G, Cohen IG *et al* (2018) Building capacity for a global genome editing observatory: institutional design. *Trends Biotechnol* 36: 741 – 743

Title page

Title of the paper:
Should human germ line editing be allowed? Some suggestions on the basis of the existing regulatory framework

Author:
Iñigo de Miguel Beriain
University of the Basque Country. UPV/EHU
Contact information:
Inigo.demiguelb@ehu.es
Phone number: 0034 670015746

Word count: 6232 words.

Key words: **Gene editing ethics, gene editing regulation, germ line protection, human genome ethics, germ line therapy ethics**

1. Introduction

In 1993, a Spanish researcher, Francisco Mojica[1], characterised for the first time what is now called clustered regularly interspaced short palindromic repeats (CRISPR). Not even twenty years later, in 2012, an article published in Science[2] showed that CRISPR combined with an enzyme called Cas9 (CRISPR-Cas9) could be used as a genome-editing tool in human cell culture. Shortly after, it was clear that CRISPR technology represented a game changer in the field of genetic engineering[3]. It is economical and relatively easy to use, and thus its results are extremely accessible. At present, it could not only open new therapeutic alternatives for those suffering from diseases related to genetic anomalies, but also provide us with a much better knowledge of diseases and new, less expensive drugs, especially if combined with other tools such as stem cells or synthetic biology technologies. In synthesis, CRISPR-Cas9 offers a new and exciting paradigm in terms of human health care.

Unfortunately, these promises are not easily realised due to a number of factors that converge to paralyse the progress of gene editing. The concerns are twofold: the first is connected with the safety of the technology, and the second relates to ethical issues. Regarding the first, one must keep in mind that CRISPR manipulation involves a number of technological issues that require urgent answers, including both off-target and on-target, but unwanted, effects that might cause severe harm to actual and future

[1] F. J. Mojica, G. Juez and F. Rodriguez-Valera, F. Transcription at different salinities of *Haloferax mediterranei* sequences adjacent to partially modified PstI sites. *Molecular Microbiology*. 1993; 9: 613–621.

[2] M. Jinek, K.Chylinski, I. Fonfara et al. Programmable dual-RNA-guided DNA endonuclease in adaptive bacterial immunity. *Science*. 2012; 337: 816–821.

[3] H. Ledford. CRISPR, the disruptor. *Nature*. 2015; 522: 20–24.

individuals[4]. Far more worrisome are those concerns directly linked to our common ethical beliefs due to the simple fact that technical and safety issues have the potential to be resolved over time by further research and advances, while moral considerations will likely continue to be the focus of public debate[5]. These are not minor concerns. It is clear that optimisation of the benefits of gene editing technologies might introduce changes in the germ line of the subjects involved, a consequence that raises deep ethical objections on the basis of the transgression of natural[6] and divine laws[7], the irremediable risks to any offspring and future generations[8], and the possibility of opening the door to eugenics[9].

---

[4] B. D. Baltimore, P. Berg et al. Biotechnology. A prudent path forward for genomic engineering and germ line gene modification. *Science*; 2015; 348: 36–38.

[5] The Hinxton Group. *Statement on Genome Editing Technologies and Human Germ line Genetic Modification*. 2015 Retrieved May 02, 2016, from http://www.hinxtongroup.org/hinxton2015_statement.pdf.

[6] B. Leiter. 2004. *The Future for Philosophy*. Oxford: Oxford University Press.

[7] N. Messer. Human cloning and genetic manipulation: some theological and ethical issues. *Studies in Christian Ethics.* 1999; 12: 1–16. J. Glover. 1984. *What Sort of People Should There Be?* London: Penguin Book.

[8] M.S. Frankel, A. S. Chapman. 2000. *Human Inheritable Genetic Modifications. Assessing Scientific, Ethical, Religious and Policy Issues*. Washington, DC: American Association for the Advancement of Sciences.

Billings, P.R., Hubbard, R., Newman, S.A. (1999). Human germ line gene modification: a dissent. *Lancet*. 353, 1873–75.

[9] M.S. Frankel, A. S. Chapman. 2000. *Human Inheritable Genetic Modifications. Assessing Scientific, Ethical, Religious and Policy Issues*. Washington, DC: American Association for the Advancement of Sciences; F. Baylis, J.S. Robert. The inevitability of genetic enhancement technologies. *Bioethics*. 2004; 18: 1–26; M. Sandel. The case against perfection. *Atl Mon*. 2004; 293: 51–62.

These considerations have already produced several preliminary calls for a cautionary approach to gene editing, specifically when the human germ line is involved. For instance, the IGI Forum on Bioethics held in Napa, California in January 2015, recommended that steps be taken to 'strongly discourage, even in those countries with lax jurisdictions where it might be permitted, any attempts at germ line genome modification for clinical application in humans, while societal, environmental, and ethical implications of such activity are discussed among scientific and governmental organisations'[10]. Similarly, the European Group of Ethics (EGE) stated that 'the EGE is of the view that there should be a moratorium on gene editing of human embryos or gametes which would result in the modification of the human genome'[11]. However, none of these statements introduced any objection to gene editing in somatic cells.

These attitudes illustrate that bioethics is currently willing to draw a legal paradigm based on the somatic/germ line distinction. According to this criterion, we should impose a moratorium on germ line editing while pursuing the application of CRIPR-Cas9 in somatic cells; 'because proposed clinical uses are intended to affect only the individual who receives them, they can be appropriately and rigorously evaluated within existing and evolving regulatory frameworks for gene therapy, and regulators can weigh risks and potential benefits in approving clinical trials and therapies'[12]. In this way, we could ensure

[10] D. Baltimore, P. Berg, M. Botchan et al. A prudent path forward for genomic engineering and germ line gene modification. *Science*. 2015; 348(6230): 36–38.

[11] European Group On Ethics In Science And New Technologies. Statement on Gene Editing. 2015. Retrieved May 02, 2016, from https://ec.europa.eu/research/ege/pdf/gene_editing_ege_statement.pdf.

[12] Organizing Committee for the International Summit on Human Gene Editing. 2015. On Human Gene Editing: International Summit Statement. 3 December 2015. Retrieved May 02, 2016, from http://www8.nationalacademies.org/onpinews/newsitem.aspx?RecordID=12032015a.

that 'legitimate concerns regarding the safety and ethical impacts of germ line editing must not impede the significant progress being made in the clinical development of approaches to potentially cure serious debilitating diseases'[13]. These statements are usually reinforced by recalling the currently existing legal framework, which—allegedly—is based on this same distinction (somatic/germ line).

The main purpose of this paper is not to perform a general analysis of the ethics of germ line editing[14]. Instead, it is primarily meant to denounce most of the distinctions drawn by our legal framework, even if they are commonly shared[15]. Indeed, I will show that, in reality, it is not any modification of the germ line that is banned by regulation in most countries (as the legal framework on mitochondrial replacement currently in place in the UK[16] or Mexico[17], for instance, demonstrates) but only the voluntary alteration of our descendants' germ line. Moreover, in the second part of this paper I will argue that

---

[13] E. Lanphier, F. Urnov, S.E. Haecker, M. Werner, J. Smolenski. Don't edit the human germ line. *Nature*. 2015; 519: 410–411.

[14] This is why issues such as the eugenic use of germ line editing or the distributive justice issues it might involve will not be addressed by this paper. Indeed, the ban on eugenics is shared by all regulations focusing on germ line issues.

[15] P. Tebas et al. Gene Editing of CCR5 in Autologous CD4 T Cells of Persons Infected with HIV. *N. Engl. J. Med.* 2014; 370(10): 901–910; M. Araki, T. Ishii, T. International regulatory landscape and integration of corrective genome editing into *in vitro* fertilization. *Reproductive Biology and Endocrinology: RB&E*. 2014; 12: 108.

[16] Castro, Rosa J. Mitochondrial replacement therapy: the UK and US regulatory landscapes, *Journal of Law and the Biosciences*, 1 December 2016; Volume 3, Issue 3: 726–735.

[17] C. Palacios-González, & M de J. Medina-Arellano. Mitochondrial replacement techniques and Mexico's rule of law: on the legality of the first maternal spindle transfer case. *Journal of Law and the Biosciences*, 2017; *4*(1): 50–69.

even this alternative criterion, voluntary alteration, which is apparently consistent according to our generally accepted moral intuitions, might become obsolete if a significant improvement in the safety of the technique happens in the future. Therefore, new legal criteria will be needed. To this purpose, I will provide a step-by-step analysis of our legal framework, showing the strengths and weaknesses of each of the currently existing normative approaches to this issue and I will conclude with a proposal for a new legal framework, that allows for both research and clinical application of germ line modification, provided that circumstances make it recommendable.

2. A preliminary approach: the absurdity of a general protection of the germ line *as such*

A quite simple way to regulate germ line editing technologies consists of establishing a general ban on each and every one of its uses[18]. This was, for instance, the strategy adopted by UNESCO in the Universal Declaration on the Human Genome and Human Rights of 1997, which states (art. 24) that '[International Bioethics Committee of UNESCO] should make recommendations, in accordance with UNESCO's statutory procedures, addressed to the General Conference and give advice concerning the follow-up of this Declaration, in particular regarding the identification of practices that could be contrary to human dignity, such as germ-line interventions.' Following these guidelines somewhat, the USA imposed a temporary moratorium on the germ line gene modification under FDA vigilance and the NIH guidelines on the basis of protection against risk[19].

[18] C. M. Romeo, Casabona. 2002. *Los genes y sus leyes*. Bilbao-Granada: Comares.

[19] M. Araki, T. Ishii. International regulatory landscape and integration of corrective genome editing into *in vitro* fertilization. *Reproductive Biology and Endocrinology: RB&E*. 2014; 12: 108.

Afterwards, the USA approved the Consolidated Appropriations Act (2016)[20] and the Consolidated Appropriations Act (2017)[21], which both ruled that no funding would be made available for a research project in which a human embryo is intentionally created or modified to include a heritable genetic modification. In practice, the effects of these provisions make it impossible for U.S. authorities to review proposals for clinical trials of heritable genome editing. As a consequence, the development of this technology might take place under alternative jurisdictions, some regulated and others not[22]. In my opinion, this attitude could hardly contribute to the development of a global strategy that serves to control and monitor germ line gene editing. On the contrary, it might worsen the current situation.

However, setting aside these examples, it is difficult to find a national regulation forbidding genome editing. This is quite an understandable situation if we keep in mind the ethical issues that a general ban on this technology involves. Indeed, the philosophy that supports this first option is unclear, the strong belief in the idea that the germ line is valuable *as such,* that is, that the germ line is *intrinsically* valuable. This is, in general, the position held by all those who adhere to the so-called defence of the sanctity of the human genome or the 'playing God' argument. According to them, we should never introduce any change to the human genome, in so far as it is 'sacred'[23], or the main tenet

[20] Available at: https://www.congress.gov/bill/114th-congress/house-bill/2029/text. See section 749.

[21] Available at: https://www.congress.gov/bill/115th-congress/house-bill/244/text. See section 736.

[22] National Academies of Sciences, Engineering, and Medicine (2017) Human Genome Editing: Science, Ethics, and Governance. Washington, DC: The National Academies Press. doi: https://doi.org/10.17226/24623. (accessed December 5, 2017), p. 191.

[23] J. Rifkin. *Algeny*. 1983. New York: Viking.

of human dignity[24], as the UNESCO Universal Declaration on the Human Genome and Human Rights stated. Therefore, if we alter it, we would be affecting the inherent nature of the human being involved[25]. Nevertheless, there are a number of reasons to oppose this assumption.

First of all, it must be highlighted that this vision is quite essentialist[26] or, to be more precise, can be considered as a kind of "genomic metaphysics" (to borrow an expression from Alex Mauron)[27]: it holds that there is something like an inner human nature that is grounded in our genes, that is, in the human genome. This belief incorporates a number of significant issues both from a biological and a moral point of view. To start with, its biological basis is feeble in so far as a common heritage of humanity is a concept that has never been universally understood. It is grounded on the basis of what all humans share like the sea and space – the commons. But it has mistakenly been interpreted as a property concept as illustrated by Okasha, who maintains, […] 'it is simply not true that there is some common genetic property which all members of a given species share, and which all members of the other species lack'[28]. Therefore, if we build our essence, nature, or worse, dignity, on the basis of this biological approach, we could be building castles on sand. This is because not only does biology not provide a sufficient

[24] L. R. Kass. 2004. *Life, Liberty and the Defense of Dignity*. San Francisco, CA: Encounter Books.

[25] J. Habermas. 2003. *The future of human nature*. Malden: Polity; G. Annas. 2005. *American bioethics: Crossing human rights and health law boundaries*. Oxford: Oxford University Press.

[26] M. Ereshefsky. What's wrong with the new biological essentialism. *Philosophy of Science*. 2010: 77(5): 674–685.

[27] A. Mauron. Genomic metaphysics. *J Mol Biol* 2002; 319: 957–962.

[28] S. Okasha, S. Darwinian metaphysics: Species and the question of essentialism. *Synthese*. 2002; 131(2): 191–213.

basis to justify human dignity, but furthermore, if we build a strict linkage between both concepts, we could feel defenceless against threats to our identity that do not form part of, or cause changes to, our genetic endowment, such as the introduction of biochips in our mind to reorientate our thoughts, memories, or feelings.

Moreover, even if we dismiss this evidence and accept the biological nature of our essence, we would not avoid all issues rising from such a vision. Even if we are to believe in a biology designed by an intelligent system—call it God or not—we would have to recognise that this design is variable because the human genome is constantly changing. Therefore, it is hard to think about it as a finished work that needs to be preserved. As John Harris wrote, 'human nature in its present form, and the genome that so dramatically influences that nature, must be considered to be both a work in progress and a mixed blessing'[29]. Indeed, each and every new human embryo carries gene mutations that affect his/her germ line. This evidence that has been recognised by article 3 of the Universal Declaration on the Human Genome and Human Rights, which states that the human genome is not static in nature, but evolves: 'The human genome, which by its nature evolves, is subject to mutations. It contains potentialities that are expressed differently according to each individual's natural and social environment, including the individual's state of health, living conditions, nutrition and education'. Therefore, it is unclear why we should refrain from introducing changes into something that is constantly changing on its own. A defender of the argument would likely refuse this criticism by stating that natural change does not justify artificial modification. Indeed, he/she might state that

[29] J. Harris, J. Gene editing technology can be harnessed for our benefit. *Financial Times*. Retrieved May 02, 2016, from https://next.ft.com/content/9fd0529e-bb6a-11e5-b151-8e15c9a029fb.

variation is not in and of itself incompatible with some form of essentialism[30], while human intervention clearly defies the sacredness of nature. However, in this case, he/she should tell us why we should be more confident in natural evolution than in human-intervention evolution. Or, if a kind of wisdom in nature is supposed, he/she should tell us why we should be more confident in nature's wisdom than in our own wisdom. Here, a metaphysical approach seems necessary, but not sufficient[31]. Were the answer that God is behind nature, then we could perfectly claim that a religious answer can hardly be assumed by those who do not share this faith, and even could be refused by believers who hold that the sacrality of the Creator could never be extended to His creation and therefore oppose to the idea of the 'sacrality of human DNA'[32].

On the other hand, it seems absurd from a moral point of view to consider that a general avoidance of human germ line editing would free us from all types of responsibility to our descendants. Would anyone hold that we would take no moral responsibility for allowing someone to be born with Huntington's disease if we had a tool to cure him/her? It seems clear that the answer to this question must be a resounding 'no'[33]. But if this is the case, then we must assume that we do not play God when

[30] N. Morar. An Empirically Informed Critique of Habermas' Argument from Human Nature. *Science and Engineering Ethics*. 2015; 21(1): 95–113.

[31] A. Buchanan. 2000. *From Chance to choice. Genetics and Justice*. Cambridge: Cambridge University Press.

[32] M. Bratton. 2009. *God, Ethics and the Human Genome: Theological, Legal and Scientific Perspectives*. Church House Publishing; T. Peters, 2010. Is the Human Genome Sacred? In *GenEthics and Religion*. G. Pfleiderer, G. Brahier, K. Lindpaintner. Eds. Basel: Karger: 108–117.

[33] B. Barnes, J. Dupré, 2008. *Genomes and What to Make of Them*. Chicago and London: University of Chicago Press; P. Singer. 1993. *Practical ethics*. Cambridge: Cambridge University Press; J. Savulescu, J.

modifying the human genome, but did when we developed the mechanisms able to change our germ line. From that moment on, we became God in so far as it was the power to make a change (or not) that conferred the role on us. Thus, we would be playing God both if we edit the human genome and if we do not, as in both cases we would be determining our descendants' future[34]. Tracing a rigid moral distinction between action and omission seems here a feeble argument that needs to be revisited.

Moreover, we must always keep in mind that even if we share the idea of the sacrality of the human genome, this belief would hardly justify a general ban on germ line gene editing, in so far as both concepts (human genome and germ line) are not synonymous. To understand why, let me start by clarifying that the expression 'human genome' could have two different meanings. One the one hand, 'human genome' can be defined as 'all the genetic information in a person'[35], that is, an individual of the human species. However, this kind of definition does not seem to match well either with our common understanding or with the meaning of the idea of 'human genome' included in the current regulation. If we were to think that the human genome refers to someone's genetic information, then we would need to conclude that our current laws ban all types of gene editing (even somatic line edition, of course), in so far as all of them involve, by

---

Pugh, T. Douglas, C. Gyngell. The moral imperative to continue gene editing research on human embryos. *Protein & Cell*. 2015; 6(7): 476–479.

[34] J. McFadden, J. Genetic editing is like playing God – and what's wrong with that? *The Guardian*. Retrieved May 02, 2016, from http://www.theguardian.com/commentisfree/2016/feb/02/genetic-editing-playing-god-children-british-scientists-embryos-dna-diseases.

[35] Definition of the Human genome. *Medicine.net.* Retrieved May 02, 2016, from http://www.medicinenet.com/script/main/art.asp?articlekey=3818.

definition, some type of alteration to an individual's genome and, thus, a change in the human genome. However, this is simply absurd.

This conclusion leads our attention to another definition of human genome (*the* human genome), linked to the idea of the whole pool of genetic information possessed by the human species, that is, a concept completely separated from that of *an* individual genome (as mentioned in articles number 1 and 3 of the UNESCO Universal Declaration on the Human Genome). This alternative sense of human genome allows us to trace much subtler distinctions in terms of the ethical and legal debate. For instance, we could hold that editing someone's genome does not necessarily involve a change in the human genome. Suppose, for instance, that we modify someone's genome by changing the expression of a gene producing Huntington's disease for its normal, healthy expression. According to this second definition of human change, such an alteration would certainly change the subject's (and his/her descendants') genome, but not the human genome at all (at least, if we manage to avoid off-target changes), in so far as the modification would not introduce any novelty into the human gene pool. Instead, a modification in someone's genome that creates a new genetic combination, previously non-existent in the human species, should certainly be considered as a human genome modification even if it does not affect his/her germ line. This subtle distinction has been extremely well addressed by Japanese bioethicist Tetsuya Ishii, who has written that 'the functional correction of a small mutation in the embryo via HDR [homology-directed repair] along with a short DNA template appears to be acceptable because this form of genetic modification can leave a wild-type gene, which is in a natural genetic state, and would fall outside of one of the ethical objections against germ line gene modification: transgression of the natural

laws. The copying of a naturally occurring variant via HDR along with a short DNA template might be considered to be natural'[36].

Therefore, we must be aware that the concept of changes in *an* individual genome and changes in *the* human genome do not necessarily match. It is true that no one can introduce changes to the human species genome without altering someone's genome (since species are no more than the addition of all their members), but it is also true that someone might change the genome of a single individual or even an extended group of human beings (if his/her germ line is affected, for instance) without changing *the* human genome. Therefore, changing *a* human genome is not the same as changing *the* human genome. For example, if someone changes the pathological expression of the HTT gene that causes Huntington's disease to allow for normal expression, they would be changing the genome of the affected human being, but not the human genome. Thus, 'there could be cases of genetic enhancement when this practice would not alter human nature, and as such, should not be morally prohibited'[37]. To summarise, we must bear in mind that a change in the human genome does not come from a mere germ line modification, but necessarily involves the introduction of a genuine novelty to the human gene pool. Therefore, a general ban on germ line editing on the grounds of the defence of the human genome is clearly unjustified, even if we believe in the sanctity of the idea.

Finally, we must keep in mind that, even if we were not convinced by previous arguments, we should at least concede that basic and pre-clinical research on human germ

---

[36] T. Ishii. Germ line genome editing in clinics: the approaches, objectives and global society. *Briefings in Functional Genomics*. 2015; 27: 4–5.

[37] N. Morar. An Empirically Informed Critique of Habermas' Argument from Human Nature. *Science and Engineering Ethics*. 2015; 21(1): 95–113.

cells and embryos in the earliest stages of development could hardly introduce a modification in the human germ line, if they are never transferred into a woman's uterus[38]. Therefore, such research application should be accepted even if we hold the idea of the sacrality of the human genome and the moral imperative not to modify it, unless we believe that an *in vitro* human embryo is part of humankind and any alteration of its genome involves a change in the human genome. But this is not a widely shared belief.

As a final recapitulation, I consider that there are no good reasons to support a general ban on germ line interventions for the sake of the germ line *as such.* In practice, it is extremely difficult to accept the idea of the sacrality of the human genome, due to a number of biological, ontological, and ethical reasons, but even if we were to accept it, this would not necessarily support a general ban on human germ line editing, but only on those edits which involve an alteration to the human genome according to the distinction traced. Moreover, it could hardly extend to interventions on early embryos that would never be transferred into a woman's womb. Thus, we should consider that a general ban on germ line editing *as such* would never make real sense both from a legal and/or ethical point of view. Therefore, a reasonable legal framework can hardly be built on this feeble basis. But, if this is the case, then we should follow a different path. Let us now examine the case of the regulations that do not ban germ line modification, but the alteration of our descendants' genome.

3. Germ line and risk

[38] B. M. Knoppers et al. Human gene editing: revisiting Canadian policy. npj Regenerative Medicine 2017; 2.1: 3.

I previously stated that a general ban on germ line modification is unreasonable. Nevertheless, this does not preclude a limited opposition to certain modifications on the basis of a convincing reason, such as the risks they might involve to the patients and, especially, to their descendants or society as a whole. This is indeed the approach that has been adopted by a number of national regulations that only ban germ line modifications when they might affect our descendants' genomes. I am referring, for instance, to the German law which states that 'anyone who artificially alters the genetic information of a human germ line cell will be punished with imprisonment up to five years or a fine', but leaves unpunished 'an artificial alteration of the genetic information of a germ cell situated outside the body, if any use of it for fertilisation has been ruled out'[39].

This approach is particularly interesting because it allows us to distinguish between germ line modification and the transmission of modifications to human lineage in legal terms. This is a crucial point that creates an enormous difference in terms of what is, what is not, what should be, and what should not be banned due to one simple fact: these two alternatives—that is, altering the human germ line and creating a modified offspring—do not directly match. Indeed, altering the human germ line is a *necessary* but not *sufficient* condition to produce an inherited change in the human genome. Although it seems obvious, it is worth remembering that to transmit a change to our descendants, it is necessary, first of all, to create that lineage. Otherwise, it is difficult to think about a possible way to produce this result[40].

---

[39] Federal Law Gazette, Part I, No. 69, issued in Bonn, 19th December 1990, page 2746. Act for Protection of Embryos (The Embryo Protection Act) Gesetz zum Schutz von Embryonen (Embryonenschutzgesetz – ESchG) Of 13th December 1990.

[40] J. Savulescu, J. Pugh, T. Douglas, C. Gyngell. The moral imperative to continue gene editing research on human embryos. *Protein & Cell*. 2015; 6(7): 476–479.

Therefore, the distinction between the germ line/somatic line and the descendant-affecting/non-affecting gene modification criterion not only makes perfect sense in that it provides us with an essential criterion to solve some of the most challenging ethical issues that gene editing technologies pose. As an example, this subtle distinction allows researchers to justify genetic modifications performed on *in vitro* embryos that will subsequently be destroyed: as they will never become adult people, the risk of introducing dramatic changes to the human genome is zero, even if their germ line is manipulated. Therefore, a regulation based on this criterion might freely allow germ line modification for research purposes while banning its clinical use, even if under strict oversight by Ethics Committees, as proposed by the International Society for Stem Cell Research and the National Academic of Sciences Report (2017)[41].

4. Introducing researchers' will into the equation

[41] The International Society for Stem Cell Research suggested that "The proposed committees would assess research goals "within an ethical framework to ensure that research proceeds in a transparent and responsible manner. The project proposal should include a discussion of alternative methods and provide a rationale for employing the requested human materials, including justification for the numbers of preimplantation embryos to be used, the proposed methodology, and for performing the experiments in a human rather than animal model system" (See: ISSCR (2016) Guidelines for stem cell research and clinical translation; at: http://www.isscr.org/docs/default-source/guidelines/isscr-guidelines-for-stem-cell-research-and-clinicaltranslation.pdf?sfvrsn=2 (accessed December 5, 2017), p. 6). This statement was endorsed by the US National Academic of Sciences (See: National Academies of Sciences, Engineering, and Medicine (2017) Human Genome Editing: Science, Ethics, and Governance. Washington, DC: The National Academies Press. doi: https://doi.org/10.17226/24623. (accessed December 5, 2017), p. 100.

Even if disjunction descendant-affecting/non-affecting gene editing makes more sense than the somatic line/germ line distinction, it still remains clear that it does not match perfectly well with the scientific evidence with which we have to deal. According to our current scientific knowledge, some medical treatments might introduce changes to the germ line that are transmitted to the patient's descendants inadvertently. An alteration of the germ line can occur quite easily even if it is not at all intended, as happens with chemotherapy[42]. Moreover, with the current state of the technology it is impossible to know if or how gene editing might only alter somatic DNA. Indeed, the danger of dramatic changes in the germ line due to an unexpected effect of a treatment targeted at somatic cells is one of the most challenging issues in terms of risk prevention. Finally, we must not forget that environmental interventions, such as modified social interaction, have epigenetic effects, modify brain development, and can be passed on to the next generation[43]. Therefore, it seems difficult to establish a regulation on the basis of such feeble boundaries, if we are not willing to place an imposing burden on our physicians' shoulders. This is why I believe that the only route worth following is the prosecution of conduct that deliberately seeks to spread a change in the human genome or, at least, conduct that does not take the necessary measures to avoid or minimise the chances of change happening. To sum up, we should not ban changes affecting the subject's

[42] E. Lunshof, 2015. Human germ line editing—roles and responsibilities. Protein & Cell. 2015; 7 (1): 7–10.

[43] C. Gingell. Editing the germ line – a time for reason, not emotion. *Practical Ethics*. 2015. http://blog.practicalethics.ox.ac.uk/2015/03/editing-the-germ line-a-time-for-reason-not-emotion/. Accessed 02 May 2016.

descendants *as such*, but only those behaviours that provoke such modification deliberately or due to gross negligence[44].

My point, thus, is that we should definitively substitute the somatic/germ line disjunction with a regulation based on the idea of willingness or gross negligence capable of provoking a change in the genome of the descendants—a criterion already included in some relevant legal tools. For instance, the Oviedo Convention (Convention for the protection of Human Rights and Dignity of the Human Being with regard to the Application of Biology and Medicine: Convention on Human Rights and Biomedicine, 1997) states that 'an intervention seeking to modify the human genome may only be undertaken for preventive, diagnostic or therapeutic purposes and only if its aim is not to introduce any modification in the genome of any descendants'. Therefore, it seems clear from its wording that an intervention made on the basis of preventive, diagnostic, or therapeutic purposes whose aim is not to introduce modifications in the genome of the descendants should be allowed. In its negative sense, this clause also encompasses the notion that we should only try to avoid practices that voluntarily introduce a modification to the genome of descendants. This criterion has been adopted by a number of national regulations in countries such as France, Australia, and Spain[45]. For instance, in the case of Spain, article 74.2.c of *Law 14/2007*, of 3 July, on Biomedical Research only considers

[44] This would make a dramatic difference in terms of Criminal Law. In the case of other jurisdictions, such as Civil Law, the situation might remain the same, since causing harm, even if unintendedly, creates a responsibility. But researchers and practitioners are usually much more concerned about Criminal Law (as no insurance policies cover the risk of a prison sentence). Therefore, the inclusion of this clause is indeed relevant in terms of research policy.

[45] M.S. Frankel, A. R. Chapman. Facing Inheritable Genetic Modifications. *Science Magazine*. 2016. at: http://science.sciencemag.org/content/suppl/2001/05/18/292.5520.1303.DC1. Accessed 02 May 2017.

'the carrying on of any intervention *aimed* at the introduction of a modification in the genome of the descent' a serious infraction, not an intervention aimed at changing the germ line.

To sum up, this final normative option allows germ line modification if it is not intended to cause (and indeed, it cannot possibly cause) any alteration in the genome of descendants. This seems to be a particularly timely regulative option in so far as it conciliates the need to continue research on human genome editing with the protection that the human genome deserves. On the other hand, it allows the clinical use of genetic modification both when it is applied in the somatic line and when it alters the germ line of infertile people. Finally, it permits medical practices that might cause unintended modifications in the germ line, provided they are not caused by a serious negligence. Therefore, there are good reasons to support its general acceptance as a minimal common regulation that might serve to prevent the appearance of any kind of 'biotechnological paradise' somewhere in the world.

5. A final remark: the need to keep an open spirit in the future

The 'apparently' final conclusion reached in the former paragraph synthesises quite well the general acknowledgement that some kind of conciliation between research needs and the protection of human beings against the risks that gene modification involves is urgently needed right now[46]. A temporal ban on actions causing inheritable modifications

[46] See: Council of Europe Recommendation 2115 (2017). The use of new genetic technologies in human beings. Text adopted by the Assembly on 12 October 2017 (35th Sitting). At: http://assembly.coe.int/nw/xml/XRef/Xref-XML2HTML-EN.asp?fileid=24228&lang=en. Accessed 16 April 2018.

to our descendants seems reasonable at present. Needless to say, if the future development of this technology provides us with good reasons to think that the weight of the risk factor is no longer strong in this discussion, we should consider a severe modification to these restrictions, unless we base the restrictions on different legal reasons such as eugenics or justice and fairness – which makes perfect sense, of course, but is beyond the scope of this paper.

But, if this is not the case—and I do not believe it is—then it seems meaningful to incorporate a final word to our currently existing ban that links the prohibition to the risk involved in the practice, and to make it clear that if this risk is acceptable (zero risk, indeed, does not exist and good practices and safety should not be confused with acceptability of risks—those judgements are ultimately personal, and perceptions also differ among researchers[47]), the reasons for a general prohibition will also vanish. On this basis, we could at least reconsider the possibility to authorise a modification in our descendants' genome if it will provide relevant benefits to their health, while banning any intervention devoted to eugenic purposes[48].

Therefore, as a final conclusion, I propose to only prohibit the carrying out of any deliberate and voluntary intervention aimed at the introduction of a modification in the descendants' genome insofar as it involves a disproportionate risk to the descendants. Of course, one might think that in doing so, we would be introducing the possibility of uncertainty, because 'disproportionate risk' is a concept that entails a high level of discretion, but, is this not the criterion that we actually follow in the case of somatic gene

[47] J. E. Lunshof. Human germ line editing—roles and responsibilities. *Protein & Cell.*2016; 7(1): 7–10.

[48] R. Isasi, B. M. Knoppers. Oversight of human inheritable genome modification. *Nature Biotechnology*. 2015; 33: 454–455.

therapy clinical trials?[49] Then, why should we not do the same in this case? I do consider that the methods and level of monitoring used in the case of clinical trials related to gene editing in the somatic line are sufficient to protect us from their most worrisome consequences. But, if this is true, then I can find no reason to think that this criterion is not also applicable to germ line therapies[50]. Of course, we could add extra guarantees, such as the intervention of a highly skilled ethics committee in the case of germ line modification, an option already implemented in Israel[51] and suggested by the International Society for Stem Cell Research and the NAS Report (2017)[52]. Similarly, professional self-regulatory approaches are a possible solution; if well developed and validated over time, they could gain significant legitimacy and have a measureable positive effect[53]. In any case, the implementation of any of these self-regulatory approaches would not change our arguments concerning germ line modification regulation: in the case of a technique that might evolve quickly and whose benefits might be considerable, we would be wise to leave some doors open and perhaps such a formula is our best chance to do so.

---

[49] J. Kimmelman. Recent developments in gene transfer: risk and ethics. *British Medical Journal.* 2005; 330(7482): 79–82.

[50] The Hinxton Group (2015). Statement on Genome Editing Technologies and Human Germ Line Genetic Modification. http://www.hinxtongroup.org/hinxton2015_statement.pdf. Accessed 02 May 2016.

[51] M. Araki, T. Ishii. International regulatory landscape and integration of corrective genome editing into *in vitro* fertilization. *Reproductive Biology and Endocrinology : RB&E*. 2014; 12: 108.

[52] See note 41.

[53] R. Isasi, E. Kleiderman, B. M. Knoppers. Editing policy to fit the genome?. *Science*. 2016; 351(6271): 337–339.

Acknowledgements

I would like to take this opportunity to acknowledge the support from the Spanish Ministry of Economy, Industry and Competitiveness (MINECO), through the Health Institute Carlos III. Project: Functional and Clinical Impact of the Genome Analysis in CLL, and the Government of the Basque Country, grant award Reference IT-1066-16.

Obsolescence is not a good reason to oppose all types of enhancement

Journal:	*American Journal of Bioethics*
Manuscript ID	UAJB-2019-0188.R1
Manuscript Type:	Open Peer Commentary

Obsolescence is not a good reason to oppose all types of enhancement

Iñigo de Miguel Beriain[1,2]

The introduction of the concept of obsolescence into the debate on human enhancement is a laudable initiative. For too long, our discussion on this topic seems to have stalled around concepts such as autonomy, dignity or social justice (De Miguel Beriain 2018; De Miguel Beriain 2019). In this sense, a novel and intelligent argument such as the one proposed by Sparrow is something to congratulate the author on. However, his reasoning includes some flaws that diminish the strength of his main conclusions. In this commentary, I focus specifically on the difficulty of extrapolating his objection to human enhancement to all the different forms of enhancement. Indeed, I hold that the obsolescence argument works well with some *positional goods* (Bostrom and Roache 2008), such as intelligence or artistic capacities, but it becomes weak when applied to absolute goods, such as health, or even when it refers to goods that are partly positional and partly absolute, such as goodness.

To this end, I start by addressing the issues related to health. This makes sense since the narrow definition of therapy (and consequently the wide definition of enhancement) adopted by the author and some other scholars he quotes has important consequences in terms of determining the conducts capable of causing obsolescence. Indeed, the author states that "*[g]enetic therapy ends when normal functioning is achieved and so effective genetic therapies will not generate obsolescence in the fashion that interests me here*". If we accept this point of view, we have to include within the concept of enhancement both those interventions that

[1] Chair in Law and the Human Genome Research Group, Department of Public Law, University of the Basque Country, UPV/EHU, Barrio Sarriena s/n, Leioa, Bizkaia, Spain. E-mail: inigo.demiguelb@ehu.eus.
[2] IKERBASQUE, Basque Foundation for Science, Bilbao, Spain.

are aimed at improving our therapeutic response capacity over a human range or even beyond the *maximum* of the species range and those that seek to reduce the genetic predisposition to suffer any pathology to a range below a human range or even beyond the minimum of the species range.

However, such types of intervention do not work well with the reasoning developed by Sparrow in his article. To begin with, they can hardly be associated with the concept of obsolescence. If we manage to inactivate a gene that predisposes an individual to a disease, it is complex to think of an improvement beyond this point. If we are able to alter a gene to successfully combat another pathology, it will be difficult to think that the future will bring an improvement. Furthermore, if gene editing succeeds as fast as Sparrow expects it to do (and this is a crucial part of his reasoning), all possible enhancements related to human health will be achieved very soon. Therefore, there might be a gap between some generations, but from the moment when all improvements are made, there will no longer exist a distinction between the generations to come.

Besides, even if this were not the case, the appeal to obsolescence as a reason to abstain from a genetic intervention would remain very feeble. Indeed, it is hard to understand why the obsolescence argument should not appeal to gene therapy (as Sparrow claims) and yet should be applicable to interventions intended for the same purpose: to optimize our health. I cannot really see a difference between them. The main problem facing Sparrow's argument here is that once we accept this initial premise (that improving our defences against pathology or lowering the probability of developing one is aimed at improving our health), the moral concerns about obsolescence disappear, since health is an absolute good, not a positional good.

That is why his idea that enhancement can make some people feel obsolete or be seen as obsolete by others doesn't work properly in this case. What fails is a very simple part: the relevance or not of the comparison between some human beings. In the case of pure positional goods, such as intelligence, this is perfectly relevant. Indeed, one can feel bad if one understands that there are other humans

more intelligent than oneself. In the case of health, however, this does not happen, or should not happen. Only very mean people are relieved to contemplate another human being with more serious health problems than their own. Most of us usually focus on our own health exclusively. Hence, the comparison loses relevance in the debate. But if this is the case, then it is unlikely that the effect of discrimination between human beings based on their obsolescence would actually occur.

Moreover, even if we do not share this rationale – that is, if we consider that obsolescence could yet be an issue – we should definitively recognize that it would not be relevant enough to refrain from enhancing the health of the new generations. There is a consensus on the idea that health is the most relevant factor when deciding on our descendants' happiness, that is, something that is beyond the discussion on the different ways in which "flourishing", "freedom" or "well-being" may be realized. Therefore, arguing that we should refrain from implementing preventive or ameliorative measures because of the obsolescence argument does not seem to be so consistent. Furthermore, even if adopted a selfish perspective, it would be easy to find good reasons why a generation would want the next ones to be as healthy as possible. This would lead to both savings in health resources that could be applied to their own pathologies and an increase in the population available to meet their needs. Therefore, rationality rules that the obsolescence argument should not impede enhancement even thought we do not concede that humans are generally worried about the happiness of their descendants.

To sum up, we must conclude that the obsolescence argument is not applicable to the discussion about absolute goods, such as health. Let us now focus on the case of goods that are partly positional, such as moral enhancement. Indeed, in a world of saints, many of us who consider ourselves to be minimally decent beings would probably have to endorse our place in the team of evil people Hence, if our descendants are successfully subjected to moral

enhancement techniques, the old generations will feel like they are worse. Thus, it looks like the obsolescence argument and its consequences have a point here.

However, there are some very good reasons to think that such a comparison would never be the case. After all, Harris (Harris, 2012) is right when he says that goodness is not just about doing good, but about doing good even though our impulses take us in another direction. It would be easy, therefore, for the old non-enhanced generations to believe that, in fact, genetic improvement would not necessarily make better people, but people somehow forced by their genes to do better deeds. Therefore, they could hold that any comparison between the two types of human beings (enhanced/non-enhanced) would be somehow futile, as it would obviate the different autonomy (and the need to use it in a morally appropriate sense) they would hold.

On the other hand, it does not seem that the possibility of being seen as obsolete could really happen in the case of moral enhancement. Indeed, there are very good reasons to believe that old generations would not have much to fear from young people undergoing moral enhancement techniques. If they really made their descendants behave in a more benevolent way, then it is likely that the enhancement would end up providing benefits to the old generations. It is very likely that their pious progeny would feel compassion rather than resentment towards them. Therefore, I can find no reasons to conclude that the obsolescence argument might apply to moral enhancement.

The appeal to obsolescence is, in sum, only a solid argument when it is presented as a rejection of some forms of gene enhancement, rather than an objection to this technology as a whole. In fact, what Sparrow should conclude is that the idea of obsolescence should impel us not to improve those traits we associate with purely positional goods, such as intelligence or beauty, for example. On the other hand, it would not pose a serious problem in the case of goods that are absolute, such as health, or that are only partially positional, such as goodness. In this sense, the argument provides good reasons for the thesis of Savulescu (Persson and Savulescu, 2016) on the need to paralyse cognitive

improvement and promote moral enhancement (Savulescu, 2016), while weakening Harris's position, which promotes the opposite (Harris, 2016)

References

Bostrom, N. and Roache, R. Ethical Issues in Human Enhancement (2008), in New Waves in Applied Ethics, edited by. Jesper Ryberg, Thomas Petersen & Clark Wolf. London: Palgrave Macmillan. 2008: 120-152]

De Miguel Beriain, I. 2018. Human dignity and gene editing: Using human dignity as an argument against modifying the human genome and germline is a logical fallacy, EMBO Reports 19, e46789.

De Miguel Beriain, I. 2019. Should human germ line editing be allowed? Some suggestions on the basis of the existing regulatory framework. *Bioethics* 33: 105–111

Harris, J. 2012. *What It's Like to Be Good. Cambridge Quarterly of Healthcare Ethics,* 21(03): 293–305.

Harris, John. 2016. *How to be good*. Oxford: Oxford University Press.

Persson, I. & Savulescu, J. 2016. Moral Bioenhancement, Freedom and Reason, Neuroethics 9 (3): 263-268.

Comment on "Should gene editing replace embryo selection following PGD? Some comments on the debate held by the International Society for Prenatal Diagnosis"

Article *in* Prenatal Diagnosis · November 2019

DOI: 10.1002/pd.5542

CITATIONS
2

READS
47

2 authors:

Inigo DE MIGUEL BERIAIN
Universidad del País Vasco / Euskal Herriko Unibertsitatea
119 PUBLICATIONS **286** CITATIONS
SEE PROFILE

Tetsuya Ishii
Hokkaido University
50 PUBLICATIONS **1,493** CITATIONS
SEE PROFILE

Photo by Stéphan Poulin

16-19 June 2020
Montreal, Canada

Save the dates and join ISPD for another great International Conference on Prenatal Diagnosis and Therapy!

Learn and network at the Palais des congrès de Montréal in the city that hosts the most international conventions in the Americas!

www.ispdhome.org/2020

ispd International Society for Prenatal Diagnosis

Building Global Partnerships in Genetics and Fetal Care

A global society with a membership of more than 400 professionals from multiple disciplines such as prenatal and perinatal care, medical genetics, basic and translational science, genetic counseling, laboratory management, fetal therapy, ultrasonography, and ethics

Membership is complimentary for students and trainees.

www.ispdhome.org
info@ispdhome.org
+1 434.979.4773

Received: 6 July 2019 | Accepted: 28 July 2019
DOI: 10.1002/pd.5542

Comment on "Should gene editing replace embryo selection following PGD? Some comments on the debate held by the International Society for Prenatal Diagnosis"

We have recently read the article entitled "Gene Editing Should Replace Embryo Selection Following PGD," recently published (26 April) in *Prenatal Diagnosis*.[1] It corresponds to the debate held in the 22nd annual meeting of the International Society for Prenatal Diagnosis held in 2018. However, our final thoughts about the discussion and the subsequent article are somewhat mixed, since they offered appreciable arguments on the issues at stake, but also substantial issues and gaps that deserve comments. This paper is aimed at introducing an alternative and four deep nuances to what the experts stated.

The alternative, nevertheless, can be summarized quite easily: it might happen that the proposed crossroad is not inevitable. Indeed, it might perfectly happen that both techniques coexist in the near future. PGD might continue to be the standard response to all those couples who are willing to prevent monogenic disease, at least for the time being. Instead, gene editing (GE) could be used to reach those aims that can hardly be attained via embryo selection following PGD, such as a general improvement in the genes of the embryos, which substantially reduces the risks of postnatally suffering from serious pathologies.[2]

The first nuance has to do with the question of risk. As the moderator rightly asserted, it is impossible to raise seriously the question of whether GE can replace PGD at the present moment. Our poor control of GE and the substantial risks involved make the negative answer obvious.[3] Therefore, the debate must be directed towards a future in which these technical issues have been resolved. However, if this is the starting hypothesis, then the allusions made by J. R. Vermeesch to the risk inherent in the technique lose all sense. The same applies to his claims that we will never be able to avoid using PGD because GE will never be safe enough. If risk factors are to be included in the debate, then it should be underlined that we are not certain at all about the safety of PGD. To begin with, PGD entails embryonic cell biopsy, a circumstance that sometimes causes the loss of the embryo. Furthermore, we are more or less sure that PGD is clinically safe at birth, but this is still unclear in older individuals born via PGD. Indeed, according to animal studies, this might not be the case.[4] Therefore, if we are to ban a technology on the basis of its possible risk, then PGD should also be banned. So, one must conclude that allusions to risk should have been avoided both because they depart from the agreed hypothesis—in the future, GE techniques will be improved and will be acceptably safe—and because they assume the impossible: to be able to guess what the course of events will be.

Our second concern has to do with an issue that somehow went unnoticed. It is true that the first participant (Dagan Wells), who defends the FOR position, shows correctly that GE allows for a reduction in the number of surplus embryos in assisted reproduction techniques. However, to our astonishment, what he does not say is that this is not the only point at which GE is far more morally acceptable than PGD followed by embryo selection. In fact, Wells overlooks a crucial conceptual issue that differentiates the two techniques. In fact, GE is intended to safeguard the health of offspring who may be suffering from various pathologies through genetic modifications. Therefore, it constitutes a therapeutic action, free of any moral suspicion. Embryo selection, instead, can only be considered a therapeutic action for perspective parents who suffer from the impossibility of generating biologically healthy offspring. Indeed, detractors of genetic selection argue that this technology contains an aroma of eugenics..[5]In fact, what the technique involves is not to "cure"[6] embryos but simply to choose which embryos will be transferred. Therefore, considering that both techniques—GE and PGD—are similar is a blurred statement for an ethical discussion.

Thirdly, it is quite striking that the participants in the discussion accept that the scope of GE is limited to a few concrete circumstances similar to those that justify the use of PGD. This statement completely dismisses the possibilities that GE offers in practice. PGD followed by embryo selection and GE share a common use: they can both be employed to efficiently prevent monogenic diseases prenatally. However, GE promises much wider applications.[7] Ideally, GE could allow for correcting multiple genes of an embryo, which would go far beyond preventing the birth of children affected by a monogenic disease. For instance, GE could give our offspring an expression of genes more suited to reducing their predisposition to cancer or to improving their immune system's performance. While this may not seem easy to implement right now, it cannot be ruled out that the situation will change dramatically in the future. What is undeniable in any case is that this kind of substantial improvement will only be possible, thanks to the use of GE techniques. Therefore, it is uncertain whether PGD and GE possess a similar capacity in purely scientific terms. Indeed, GE is far more versatile than PGD followed by embryo selection. Thus, it will be exponentially superior, if we are effectively capable of acquiring sufficient knowledge about

the human genome to understand what changes are satisfactory for human beings.

J. R. Vermeesch could reply to our comment by saying that we are talking about enhancement, not therapy. Indeed, this is quite probable, since he apparently assumes that only those interventions aimed at curing monogenic diseases in embryos can be considered as therapeutic. However, medicine is increasingly seen as a global intervention that is aimed not only at curing but also at preventing diseases. Moreover, we must remember that GE that alters a gene that triggers a monogenic disease is not curative since that embryo does not suffer from the disease. It is therefore clearly preventative not curative. Therefore, if this type of GE is morally acceptable, then GE for preventive purposes should also be acceptable. Both behaviours are equally therapeutic.

What would happen in cases where the intervention is aimed at improving the immune system? In our opinion, we would also have to think of these activities as therapeutic actions.[8] This is due to the evidence drawn from some of the interventions aimed at purposes that have little to do with the cure of illnesses and which, nevertheless, are described as therapeutic. The best example of this is vaccines. Vaccines do not cure any disease. They do not even diminish our predisposition to suffer from them: they improve our response to them. However, this improvement is not usually considered a form of enhancement but a form of therapy. So, why should not we think in the same way about GE?[9] If this is the case, we must conclude that the therapeutic use of GE extends far beyond the cure of a disease. But then it is entirely possible to maintain the therapy/enhancement distinction in GE and thereby avoid the slippery slope effect and thereby avoid the slippery slope effect that the rapporteur describes as almost inevitable.

Moreover, it is important to emphasize that even if we do not accept our main argument—that is, if we consider the use of GE for preventive purposes or to improve the immune system as enhancement—it would still be possible to draw distinctions between enhancement that affects absolute goods, such as health and enhancement that affects positional goods, such as intelligence. From this distinction, it would also be possible to draw a barrier between what is permissible and what is not. Of course, this does not necessarily mean that this barrier would not go unchallenged on a regular basis. However, this also happens in the context of PGD, where evidence shows some questionable uses of this technique.[10]

The reader should not think that the conclusion of everything we have argued in this text is necessarily that GE should replace PGD followed by embryo selection. Our purpose has not been to answer this question. What we have tried to do is to clarify that the original discussion suffered from some issues and gaps, which our contribution may have helped to clarify. Obviously, there are a lot of moral arguments against GE that we have not dealt with here. The readers may well consider them when deciding his or her answer to the question posed. But at least now they can do so knowing that(a) the risk argument should not be seriously taken into account if we think about the future, (b) PGD followed by embryo selection and GE are conceptually different (eugenics/therapy), (c) GE can potentially give us far superior options to the alternative, and (d) the use of GE to prevent or improve our response to specific pathologies does not definitely constitute enhancement or, even if it were to, this would not mean that it could not be distinguished from other forms of enhancement.

FUNDING STATEMENT

Iñigo de Miguel Beriain research is funded by the Government of the Basque Country, Grant Award Ayudas Grupos de Investigación IT-1066-16 and the Instituto de Salud Carlos III, Grant Functional and Clinical Impact of Genomic Analysis in CLL. Tetsuya Ishii' work was supported by the JSPS KAKENHI Project Innovative Ethics (Graduate School of Humanities, Kobe University, Kobe, Japan).

DATA AVAILABILITY STATEMENT

Not applicable.

CONFLICT OF INTEREST

Both authors declare no conflict of interest at all.

ORCID

Iñigo de Miguel Beriain https://orcid.org/0000-0002-2650-5280

Iñigo de Miguel Beriain[1,2]
Tetsuya Ishii[3]
[1]*University of the Basque Country, UPV/EHU, Leioa, Spain*
[2]*IKERBASQUE, Basque Foundation for Science, Bilbao, Spain*
[3]*Office of Health and Safety, Hokkaido University, Sapporo, Japan*

Correspondence
Iñigo de Miguel Beriain, Chair in Law and the Human Genome Research Group, Library Building. Leioa Campus, University of the Basque Country, Leioa 48000, Bizkaia, Spain.
Email: inigo.demiguelb@ehu.eus

REFERENCES

1. Wells D, Vermeesch JR, Simpson JL. Current controversies in prenatal diagnosis 3: gene editing should replace embryo selection following PGD. *Prenat Diagn*. 2019;39(5):344-350.
2. Cavaliere G. Genome editing and assisted reproduction: curing embryos, society or prospective parents? *Med Health Care Philos*. 2018;21(2):215-225. https://doi.org/10.1007/s11019-017-9793-y
3. de Miguel Beriain I. Should human germ line editing be allowed? Some suggestions on the basis of the existing regulatory framework. *Bioethics*. 2019;33(1):105-111.
4. Zacchini F, Arena R, Abramik A, Ptak GE. Embryo biopsy and development: the known and the unknown. *Reproduction*. 2017;154(5):R143-R148.
5. de Miguel Beriain I, Penasa S. The embryo survival criterion: a moral obligation or a eugenic practice? *Rom J Leg Med*. 2018;26(2):212-217.
6. MacKellar C, Bechtel C. *The ethics of the new eugenics*. New York: Berghahn Books; 2014.

7. Steffann J, Jouannet P, Bonnefont JP, Chneiweiss H, Frydman N. Could failure in preimplantation genetic diagnosis justify editing the human embryo genome? *Cell Stem Cell*. 2018;22(4):481-482.
8. Braun M, Meacham D. The trust game: CRISPR for human germline editing unsettles scientists and society. *EMBO Rep*. 2019;20(2): e47583. https://doi.org/10.15252/embr.201847583
9. Savulescu J, Singer P. An ethical pathway for gene editing. *Bioethics*. 2019;33(2):221-222.
10. Baruch S, Kaufman D, Hudson KL. Genetic testing of embryos: practices and perspectives of US in vitro fertilization clinics. *Fertil Steril*. 2008;89(5):1053-1058.

Journal of
Medical Ethics

Is the "serious" factor in germline modification really relevant? A response to Kleiderman, Ravitsky and Knoppers

Journal:	*Journal of Medical Ethics*
Manuscript ID	Draft
Article Type:	Response
Keywords:	Gene Therapy/Transfer, Genetic Selection, Genetic Engineering, Ethics

Is the "serious" factor in germline modification really relevant? A response to Kleiderman, Ravitsky and Knoppers

Abstract

Should we use human germline genome modification only when serious diseases are involved? This belief is the underlying factor in the article written by Kleiderman, Ravitsky and Knoppers to which I now respond. In my opinion, the answer to this question should be negative. In this paper, I attempt to show that there are no good reasons to think that this technology should be limited to serious diseases once it is sufficiently proven to be safe and efficient. In fact, opting otherwise would negatively harm human beings' right to the highest standard of health that unmodified embryos could promote. Therefore, the issue should not be so much to define adequately what a serious disease is, but rather to elucidate whether this concept should play any role beyond the context of preimplantation genetic testing. This paper argues that we should not accept the similarity between technologies such as preimplantation genetic testing and human germline genome modification, because they face different challenges and offer totally different possibilities. Therefore, we are in urgent need to build a completely new ethical architecture that covers the application of germline editing in human embryos. As a part of that process, a much deeper debate on the necessity of distinguishing different disease types is required.

Main body

The paper by Kleiderman, Ravitsky and Knoppers[1] is an excellent piece; it adds important reflections to the debate on assisted reproductive technologies. First of all, they are quite right to underline the importance that the documents produced by important institutions, including Quebec's Commission on Ethics in Science and Technology[1], the

US National Academies of Sciences, Engineering and Medicine[3] and the German Ethics Council[1], provide to the concept of "serious diseases". Moreover, their efforts to build a new approach to this concept that overcomes the obstacles that have traditionally hindered its development are particularly praiseworthy.

In my opinion, however, the article also has an underlying error that should be highlighted. Kleiderman, Ravitsky and Knoppers accept as a fact that human germline genome modification (HGGM) should only be performed in cases of serious diseases. I believe that this supposition is not necessarily true or, rather, that it may be clearly mistaken. Moreover, except in those cases in which HGGM may be the best or only option for couples to have a healthy, genetically related child, the success of this innovative technique will be measured precisely by its ability to go beyond the fight against serious diseases through the modification of human embryos.

In my opinion, the error of these authors is based on the parallel they draw between preimplantation genetic testing (PGT) and HGGM. Indeed, they state that "past experience with the normative analysis and governance of PGT and prenatal testing can serve as a model to guide similar debates surrounding the acceptability of HGGM"[1]. I believe that accepting this parallelism is a mistake with serious consequences. The point to keep in mind is that in the context of PGT—but only in this specific context—is where the distinction between serious diseases and other diseases that cannot be considered as such makes sense for various reasons[2].

The first reason is that PGT is an invasive procedure that injures the embryo and can lead to its loss[4]. Moreover, it is possible that PGT provokes long-term consequences on the human being who suffered it in the embryonic state[5]. Therefore, from the point of view of the welfare of the embryo and the person it will produce, it makes sense to limit the circumstances in which this technique should be applied. Furthermore, it is important to remember that PGT does not in any case improve the health of the embryo and/or the person it generates. PGT is a technique that only provides us with the capacity to discriminate between embryos, and thus it clearly operates as a negative selection mechanism[6]. Therefore, it is only worth using when we suspect the presence of factors with enough weight to justify this screening. This last assessment is particularly important if we bear in mind that its very selective nature makes PGT often accused of being a refined form of eugenics[6]. Moreover, and as the authors express, the fact that a pathology is included in the catalogue of serious diseases incompatible with a reasonably good life

"could lead to further stigmatisation of people with disabilities" (p. 4)[1]. For all these reasons, it makes sense to limit the possible the use of PGT as much as possible, so that it is only used in cases where there seems to be no other reasonable option. Thus, PGT makes perfect sense only in the context of serious diseases.

The question then: is the same true for HGGM? In my opinion, the answer is clearly negative, for multiple reasons. The fundamental one is that this technique, unlike PGT, is clearly therapeutic, in the sense that it allows us to improve the health of human beings[7]. Therefore, if the technique is safe and effective—the hypothesis on which the whole reasoning of the article I am now criticising is built—it is hard to understand why it should apply exclusively to serious diseases and not to all diseases. If we accept that there is a "right to the highest attainable standard of health" and that HGGM "could be perceived as a form of preventive personalised medicine and a tool to foster the realisation of the right to health" (p. 3)[1], then why should we limit its use exclusively to dealing with serious diseases? I do not find any good reason for such a restriction, especially if we bear in mind that, precisely for those serious diseases, there is already a more or less functional tool (with all the issues exposed-), namely PGT. Consequently, it seems reasonable to conclude that the value of HGGM comes mainly from its ability to go much further than what PGT is and will be able to accomplish. However, if this possibility is the case, limiting its applicability to serious diseases is depriving the technique of its raison d'être, which, in turn, implies renouncing to facilitate the "right to the highest attainable standard of health" mentioned above.

There are, of course, some possible objections to this argument. For instance, one might reply that I am forgetting the relevance of the risk/benefit criterion[3]. Following this principle, we should use HGGM only when the potential benefit far exceeds the risk inherent in the use of this technique. Evidently, the more serious the disease to be faced, the lower the ratio and, therefore, the more advisable the use of HGGM. However, this objection is based on a contradiction of what the authors assume in their text, namely that HGGM will at some point be "safe and efficient". If we do not arrive at such a scenario, its use will be unethical for all diseases. In other words, if the relevant safety conditions are not met, the distinction between one type of disease and another will be completely irrelevant.

A second objection—which may be more substantial in my view—is that the monitoring of all human beings to whom HGGM has been applied for many years will

only be possible if the number of cases to follow is low. Hence, it seems appropriate to apply the technique only to the most serious diseases. The question here, however, is whether it is necessary to perform this control on each and every modified human being[8]. This debate is complex, because while it is true that risk is inherent in science and that we can hardly know the long-term consequences of HGGM, it is also true that it does not seem necessary to extend this type of control to all cases, but rather to a significant sample. Moreover, it is worth remembering that we were not aware of the long-term effects of assisted human reproduction techniques when we started to use them, but this fact never provoked a general veto for their use. Why should we opt for a different approach in the case of HGGM?

Finally, it is objectionable to my argument that "the notion of serious may be useful in determining who has the most urgent claim to HGGM (e.g. families suffering from serious genetic diseases) and therefore should be assisted or favoured to enable equitable access" (p. 5)[1,9]. However, this objection has at least two weaknesses. First, it should only apply to publicly funded interventions: there seems to be no reason to prohibit a person or couple from using these techniques to improve the health of their offspring if they are willing to finance them from their own pocket. Second, we can accept that this criterion—opting first to deal with serious diseases—may be reasonable within the framework of a public health service, but I very much doubt that it is necessary to externalise this relevance by producing a standard or recommendation of the type cited in the article. It is indeed a general criterion of efficiency in the use of public resources that they are allocated to cases in which the cost/benefit ratio is optimal. Therefore, it is not necessary to generate new specific regulations for HGGM to achieve this goal.

In short, my conclusion is that, although the article I am now criticising makes commendable contributions, it is worth seriously considering whether its underlying assumption—that HGGM should be applied at least preferably to serious diseases—is reasonable. I am inclined to think that it is not, although it is true that there are some factors that operate for the other side of the argument. In any case, I believe that a more in-depth discussion on this subject is advisable, at least if we want to reach a broad consensus on it[10].

References

1 Kleiderman E, Ravitsky V, Knoppers BM. The 'serious' factor in germline modification, *J Med Ethics,* Epub ahead of print, 28 June 2019, doi: 10.1136/medethics-2019-105436

2 Knoppers, B, Bordet, S & Isasi, R., Preimplantation Genetic Diagnosis: An Overview of Socio-Ethical and Legal Considerations. *Annual review of genomics and human genetics.* 2006, 7. 201-21.

3 National Academies of Sciences, Engineering, and Medicine. *Human genome editing: science, ethics, and governance.* Washington DC: The National Academies Press, 2017

4 Zacchini, F, et al. Embryo biopsy and development: the known and the unknown. Reproduction. 2017; 154.5: R143-R148.

5.- De Miguel Beriain I, and Ishii T, Should Gene Editing Replace Embryo Selection Following PGD? Some comments on the debate held by the International Society for Prenatal Diagnosis, *Prenatal Diagnosis*, forthcoming issue.

6 De Miguel Beriain I, Penasa S. The embryo survival criterion: a moral obligation or a eugenic practice?, *Rom J Leg Med.* 2018; 26(2): 212-217

7 Savulescu J, Pugh J, Douglas T, Gyngell C, The moral imperative to continue gene editing research on human embryos. *Protein & Cell.* 2015; 6(7): 476–479

8 Howard HC, van El CG, Forzano F, et al. One small edit for humans, one giant edit for humankind? Points and questions to consider for a responsible way forward for gene editing in humans. *Eur J Hum Genet*. 2018;26(1):1–11. doi:10.1038/s41431-017-0024-z

9 Farrelly C, Genes and equality, *Journal of Medical Ethics*, 2004;30:587-592.

10 Baylis F. 'Broad societal consensus' on human germline editing. *Harvard Health Policy Review*, 2016;15(2):19–22;

The CRISPR Journal
Volume 2, Number 6, 2019
Mary Ann Liebert, Inc.
DOI: 10.1089/crispr.2019.0010

The CRISPR Journal

PERSPECTIVE

Safety of Germline Genome Editing for Genetically Related "Future" Children as Perceived by Parents

Tetsuya Ishii[1,*] and Iñigo de Miguel Beriain[2,3]

Abstract

The social acceptability of germline genome editing (GGE) depends on its perceived safety, as well as respect for reproductive autonomy. However, it is doubtful that prospective parents sufficiently understand the risks of GGE. In the future, the use of GGE in specific situations seems plausible, as it offers couples potential means to safeguard genetically related future children from a serious disease and overcome infertility due to a gene mutation. Should GGE fail, however, some couples may be obliged to abort affected fetuses, or give birth to adversely affected children, which would be a tragedy. Some children might develop diseases later in life due to overlooked off-target mutations. Compounding this, some parents are unlikely to inform their offspring about the details of conception, hampering necessary follow-up. Prospective parents, scientists and policy makers should carefully discuss the safety implications of GGE for genetically related future children.

Introduction

In theory, genetically modifying human germ cells, which include the egg cells, sperm cells, and zygotes (collectively referred to as the germline), can enhance the developmental potential of embryos and result in children with an intended trait. However, due to the inheritance of genetic modification among future generations, germline genetic modification has been tremendously controversial, raising concerns over the safety and welfare of future generations, potential changes to the nature of human reproduction and parent–child relationships, exacerbation of prejudice against people with disabilities, and potential misuse for genetic enhancement.[1]

In the past two decades, some clinics have attempted to use germline genetic modifications primarily to treat intractable infertility.[2,3] Since 1996, several reproductive techniques involving cytoplasmic or nuclear transfer have been developed to modify the composition of mitochondrial genome (mtDNA) of human eggs or zygotes using donor eggs.[3] Although some of these cases led to live births, others resulted in miscarriages, chromosomally abnormal pregnancies, and the development of disorders in offspring.[3,4] Although some countries have prohibited such germline genetic modification, the United Kingdom became the first nation explicitly to permit two types of cytoplasmic replacement using nuclear transfer to exclude most mutated mitochondria in the eggs or zygotes (mitochondrial donation) in 2015 in order to allow prospective parents to have genetically related children free from serious mitochondrial disease.[5]

Recently, genetic modification technologies using programmable bacterial nucleases (DNA-cutting enzymes), collectively called ''genome editing,'' have spread worldwide as efficient, versatile, and cheap genetic engineering tools. One of them, CRISPR-Cas9, uses nucleases and programmable guide RNA molecules to modify specific genes in the nuclear and mitochondrial genomes of various species.[6] Since 2015, basic research on human germline genome editing (GGE) has proceeded toward clinical applications to prevent genetic disease prenatally.[7–9] In contrast to germline genetic modifications through cytoplasmic or nuclear transfer, GGE technically does not depend on gamete donation from third parties. It requires only the introduction of programmed nucleases into the germline. Therefore, it is likely that more prospective parents will consider GGE more favorably. Indeed, a recent survey found that approximately 60% of respondents

[1]*Office of Health and Safety, Hokkaido University, Sapporo, Japan;* [2]*Law and the Human Genome RG, University of the Basque Country UPV/EHU, Faculty of Law, Leioa, Spain;* [3]*IKERBASQUE, Basque Foundation for Science, Bilbao, Spain.*

*Address correspondence to: *Tetsuya Ishii, PhD, Office of Health and Safety, Hokkaido University, Sapporo 0600808, Japan,* E-mail: tishii@general.hokudai.ac.jp

accept GGE for medical purposes, whereas only 27% accept it for nonmedical purposes.[10]

Nonetheless, reports on basic research into GGE technology have still stirred a fierce global debate. In recent years, international societies and communities have issued more than 60 ethics statements regarding GGE.[11] Notably, a 2017 report by the U.S. National Academies of Sciences, Engineering, and Medicine (NASEM) concluded that GGE trials might be permitted only after further preclinical research clarifies the potential ''risks and benefits,'' and only for ''compelling medical reasons … [in the] … absence of reasonable alternatives.''[12] Additionally, a 2018 report by the Nuffield Council on Bioethics (NCB) in the United Kingdom concluded that GGE could be acceptable if it is intended to secure, and be consistent with, the ''welfare …[of the] … future person'' and should not increase disadvantage, discrimination, or division in society.[13]

At the end of 2018, a Chinese researcher, who asserted that his research conformed with the guidelines of the NASEM report, claimed that twin girls had been born safely via GGE using CRISPR-Cas9.[14] Subsequently, a Chinese regional government confirmed their births but found serious compliance violations.[15] In this use of CRISPR-Cas9, the researcher attempted to introduce a CCR5 gene deletion that confers resistance to human immunodeficiency virus (HIV) infection (naturally found in approximately 10% of Northern European populations). Importantly, the father of the twins was HIV-positive, and both parents considered that providing HIV resistance for their future children would enhance the welfare of their offspring. However, in this and other cases, GGE might unintentionally cause off-target mutations in genes that are important for health, which can potentially affect the resultant embryos, fetuses, and children systemically, infringing on their human rights.

Therefore, we must be aware that a world where GGE is available either legally or illicitly is imminent. Before prospective parents with autonomy widely pursue experimental GGE, it is imperative that we consider the safety of GGE and its social implications. To facilitate much-needed discussion, the present article will examine the safety of GGE for medical purposes.

Differences Between Somatic Editing and Germline Editing

Previously, in two trials of conventional somatic gene therapy for a severe combined immunodeficiency (SCID), 5/20 subjects developed leukemia several years after the administration of CD34+ progenitor cells in which a retroviral vector encoding *IL2RG* was introduced.[16,17] The side effect of leukemia was due to the activation of protooncogenes caused by the genomic insertion of retroviral vectors in an unintended manner. Physicians treated the five affected subjects, and while four recovered, one ultimately died. However, the safety of viral vectors has recently been improved.[18] Currently, clinical trials using somatic cell genome editing (SGE) are ongoing to develop novel therapies for patients with cancers and genetic diseases.[6] In SGE, programmed nucleases are introduced into somatic cells, and then a target DNA sequence is efficiently cut and modified for a therapeutic effect. However, there remains a risk of unintentional large-scale rearrangements or small insertion and deletion mutations at off-target sites,[19] which could result in serious side effects, including tumor formation through the activation of protooncogenes or disruption of tumor suppressor genes.[20] Despite the different mechanisms between SGE and more traditional gene therapy, such unintended genetic modifications are irretrievable and persistent in some cells of subjects, which drastically differentiates SGE and gene therapy from chemical agents that are soon metabolized and excreted from the body.[20] Before obtaining consent from volunteers, the potential risks and burdens, as well as the benefits of SGE, must be carefully explained to them.

The use of germline genetic modification in the context of contemporary reproductive medicine introduces more complex ethical issues into the debate. In this case, the individuals providing consent will be prospective parents. However, the actual subjects being directly affected will be eggs or sperm cells, embryos, and ultimately children via reproduction. Of course, unborn children cannot be informed of the risks and cannot give consent. Again, we consider the case of germline genetic modifications through cytoplasmic or nuclear transfer, including mitochondrial donation. These processes involve the transfer of mitochondria (containing their mitochondrial genome) from eggs donated from a third party, and thus are controversial due to their involvement of a third ''parent'' in the *in vitro* fertilization (IVF) process.[21,22]

The direct use of donor eggs can help prospective parents have children, but at the same time, egg donation has raised ethical issues concerning female exploitation and the commodification of eggs in addition to potential harms to female donors.[23] Moreover, the intended mother is not genetically related to the donor-conceived children, which can also lead to concerns, such as emotional conflict over whether to disclose the fact of donor conception to the children,[24] and problems of resemblance due to a lack of physical similarity between the mother and such children.[25] Mitochondrial donation could alleviate these concerns by helping prospective mothers who consent to it to have genetically related children (genetically related children free from mitochondrial disease when the prospective mothers have a pathogenic mtDNA).

The fact that GGE does not technically require gamete donation will make it more appealing to prospective parents, including parents who seek genetically related children free from disease, as well as parents wishing to overcome infertility due to a gene mutation.[26,27] However, there is no medicine without risks. In the above-mentioned Chinese GGE case, several embryos that underwent genome editing were tested for the presence of intended and unintended genetic modifications. Then, selected embryos were transferred to a prospective mother. In this context, we note that genetic testing of modified embryos may overlook unintended, small genetic modifications that result from the use of insufficiently programmed nucleases.[26] Although the author of the Chinese report claimed that the GGE process ended in healthy births, it is unlikely that all GGE cases will proceed without risks to the resultant embryos, fetuses, and children. As such, no ethics committees are likely to approve large-scale GGE studies enrolling hundreds or thousands of couples, even though some authors have asserted that large studies are needed.[28] Indeed, GGE studies are likely to be limited to small, open-labeled, uncontrolled case series, if performed at all. Although GGE might attract many prospective parents who have reproductive autonomy, it is likely to remain an experimental intervention to human reproduction for a long time due to the limitations of clinical study.

Moral Status of Human Embryos in GGE

It is worth exploring divergent views on prenatal life in prospective parents who consider GGE. In regard to the moral considerations regarding the human embryo and fetus, we can describe the three main outlooks as the ''all,'' ''none,'' and ''gradualist'' positions.[29]

The "All" position

Those adopting the ''all'' position hold that human embryos already possess full human status. For them, germline genetic modifications, including GGE, are in themselves acceptable because these unborn ''humans'' who are suffering from a genetic problem deserve medical attention.[30] At present, however, those holding the ''all'' position have come to regard such experimental interventions as unethical, since a relevant addendum to their position is that there be no risk of adverse events, no use of reproductive techniques, and few or no wasted embryos (humans).[31] Again, however, they would essentially accept the use of GGE for ensuring the welfare of unborn ''humans'' if GGE was perceived as safe, if it was not considered reproductive medicine, and if it would contribute to reduce the number of created embryos and, consequently, the number of spare embryos.

The "None" position

Conversely, the ''none'' position asserts that human embryos or fetuses have no moral status and therefore deserve no special moral concern before childbirth. As such, those holding the ''none'' position largely accept germline genetic modifications that can help them to have genetically related children.

The "Gradualist" position

The ''gradualist'' position regards human embryos as potential human beings, but not actual humans until birth. It also considers that human embryos possess a special status that deserves a certain degree of respect, which increases along with their development. Similarly to the ''none'' position, the ''gradualist'' position may accept germline genetic modifications for both clinical and research purposes. However, prospective parents are likely to encounter dilemmas when embryos and fetuses implanted in the mother are adversely affected. While they may accept that such embryos and fetuses are morally different from actual human beings, they may still feel that these entities should be treated with a higher degree of respect than genome-edited embryos that have yet to be implanted. In so doing, it seems unclear whether such affected implanted embryos and fetuses should be medically treated, as they were in the SCID gene therapy trial.

Scenarios After GGE and Psychological Aspects of Parents

To consider scenarios that might arise after GGE, the first human germline genetic modification is revisited below. From 1996 to around 2002, an American clinic performed a small study to test a germline mitochondrial modification technique for intractable female infertility, wherein the cytoplasm (containing mitochondria) from a donor egg is transferred into a patient's eggs (cytoplasmic transfer).[4] The study helped 13 couples have 17 genetically related children, but also resulted in a miscarriage, probably due to a chromosomal abnormality (Turner syndrome), and one selective fetal reduction from a twin pregnancy (Turner syndrome). While the pregnant woman might have been able to support both twins, she selectively aborted the affected fetus, where the mother's reproductive autonomy is generally permitted to take precedence over the life of the embryo or fetus.[32] Of note, it is virtually impossible to treat affected fetuses fundamentally in cases where many or all of the cells are genetically abnormal. Likewise, in GGE, while some parents may have genetically related children, others may choose to abort adversely affected fetuses.

Pregnant women frequently opt for abortion when prenatal testing reveals a genetic abnormality in their fetus.[33]

In terms of abortion in cases of germline genetic modification, further exploration of the psychological aspects of the parents is needed due to the potential risks to unborn children of parents who adopt the ''all position'' or to potential children for those who take the ''gradualist position.'' In contrast to the cytoplasmic transfer study, the Chinese GGE case and most GGE basic research intend to prevent the onset of a disease in offspring prenatally, despite etiological differences. If prospective parents who have consented to GGE for that medical purpose, they are not harboring a vague desire for healthy children but clearly wish to safeguard genetically related children against a specific disease. Nonetheless, if issues arise with the genetic intervention, they may be forced to choose whether to abort the affected fetuses—namely, whether to ''kill unborn or potential children.'' Such conflict would likely cause substantial grief to such couples.

Safety of GGE for Future Children

From a broader perspective, abortion can play a large role in integrating GGE into society. Indeed, the U.S. cytoplasmic transfer study, in which one affected fetus was aborted, helped 13 infertile couples to have genetically related children. However, 1 of the 17 children was subsequently diagnosed with a borderline pervasive developmental disorder, resulting in a regulatory intervention by the U.S. Food and Drug Administration.[34] For genome editing, a target sequence-binding molecule is designed primarily using a reference genome. However, the human genome differs slightly among individuals,[35] which may mislead scientists programming nucleases and result in unintended genetic modifications. In addition, there is no perfect prenatal testing. Should GGE become widespread, these limitations of genome editing and prenatal testing may lead to the birth of some adversely affected children or the later development of disease in some of these individuals after they grow. How then should the safety of experimental GGE be considered?

Once again, we might use as a paradigmatic example the discussions surrounding mitochondrial donation in the United Kingdom. In this case, a regulator's panel concluded in 2013 that ''evidence available at that time did not suggest that the techniques are unsafe,''[36] which in part led to the legalization. This suggests that possible points to consider regarding the safety of GGE include the presumed probability and seriousness of adverse events at birth and/or at some future points after preclinical research defined the degree of safety. Is it acceptable so long as the probability of adverse events after GGE is far less than one adverse event in 17 children in the case of cytoplasmic transfer? For some, such a lower probability may be acceptable, as most couples have genetically related healthy children. However, in addition to this probability, the seriousness of conditions of affected children should also be addressed.

On closer examination, unsafe mitochondrial donation has two different implications. First, the germline mitochondrial modification may fail to prevent the inheritance of serious mitochondrial disease to offspring. This failure would then result in the birth of children affected by a serious mitochondrial disease. Second, mitochondrial donation may affect the resultant children in an unexpected manner, imposing serious conditions other than mitochondrial disease upon them. Such conceivable risks to future children eventually did not halt the legalization of this procedure in the United Kingdom because those risks were perceived as acceptably low. Likewise, when preclinical research further advances, some countries could justify the clinical use of GGE for medical purposes, probably including the prenatal prevention of serious disease.

One might view the Chinese researcher's claim that twins were born safely via GGE as similar to the first successful birth using IVF in 1978, which led to wide use of technique worldwide. However, small off-target mutations might cause health complications in the children as they grow. It is therefore crucial to conduct long-term follow-up of children born via GGE.[37] Regarding mitochondrial donation, the regulator in the United Kingdom only requires that reproductive scientists prepare a long-term follow-up plan,[38] suggesting that the long-term follow-up after GGE will also be left up to scientists and parents. However, not all the children are likely to undergo the necessary follow-up. This is not simply because parents who consented to follow-up may later withdraw that consent. Of note, in the survey results of families who joined the cytoplasmic transfer study, only 1 out of 13 couples disclosed the use of the germline genetic modification to their children, which appears to be lower than the disclosure rate after the direct reproductive use of donor eggs.[4] This low disclosure rate is probably associated with the involvement of ''experimental'' infusion of mitochondria derived from donor eggs, which makes parents hesitate to disclose the fact of conception to their children. In such a situation, children who are not informed of the fact that they were born using experimental reproductive medicine may sometimes decline important hospital visits simply because they do not understand that their genome was edited, and that medical examinations are important to offset the potential risks associated with genome editing. Thus, their health may be threatened in the future.

The regulator in the United Kingdom requires reproductive scientists to report childbirth with mitochondrial

disease, birth defects, genetic abnormalities, or other adverse events after mitochondrial donation was performed.[38] Aside from issues with long-term safety, consider a situation in which the use of GGE for preventing serious disease in future children unexpectedly results in the birth of one child afflicted with another serious illness. It may be impossible to treat such a child fundamentally, as unsafe genome editing has unintentionally modified many or all of their cells. This would be a tragic event for all parents, regardless of their views on prenatal life, and some parents might even bring a wrongful birth lawsuit against reproductive scientists.[39] However, this lawsuit would likely become a protracted court case, as it would be unclear whether adverse events occurred due to the side effects of GGE or to the genomic instability frequently observed in the early embryos.[40] Furthermore, unsafe GGE may later result in family discord. In some countries, the affected child could bring a wrongful life lawsuit against their parents in addition to reproductive scientists, claiming that she or he should not have been born.[39] However, such actions are unlikely to be taken by the child because the parents refrain from disclosing the fact of their conception involving an experimental genetic intervention. The socially permissive politics that gave rise to the legality of wrongful birth lawsuits could backfire and diminish the rights of individuals whose parents attempt to guard preemptively against wrongful life lawsuits by withholding the facts of their conception.

Regarding conventional gene therapy, nearly 3,000 trials have been performed worldwide, with a dozen or more approved therapies.[26] Several SGE trials are ongoing at present. With prior review and patient consent, such clinical efforts are worthwhile because the development of SGE as well as gene therapy may satisfy the needs of current disease sufferers. This is in contrast to germline genetic modifications. Those who consent to such germline interventions are prospective parents. Current research reports suggest that most of the subjects would be preimplantation embryos with or without genetic defects that are unborn children for the ''all'' position but are morally different from existing humans to those adopting the ''gradualist'' or ''none'' position. For prospective parents with reproductive autonomy, GGE can be viewed as wish-fulfilling medicine.[41] Namely, GGE can help such parents create genetically related children with or without a specific trait. In contrast, some argue that germline genetic modification has few compelling needs and little social value.[21,22] In addition, unintended off-target mutations may adversely affect germline cells, potentially ruining the welfare of the resultant children, which conflicts with the ideals laid out in the NCB report. If countries judge that the potential harms to future children as well as the relative paucity of compelling needs or social value outweigh parental wishes for having genetically related children, they will maintain or prohibit GGE. On the other hand, in the United Kingdom, mitochondrial donation is legal for prospective parents pursuing genetically related healthy children. Therefore, the acceptability of GGE largely depends on its safety, as well as the degree of respect for parental reproductive autonomy. Is the risk of unintended victims of GGE an acceptable risk for such parents? At present, it is doubtful whether prospective parents sufficiently perceive the safety implications of GGE.

Conclusion

Given that GGE is more likely to spread than older germline genetic modifications, and that GGE carries a real risk of adversely affecting children with human rights in an irretrievable manner, it is vital for prospective parents with reproductive autonomy, as well as scientists and policy makers, to perceive the safety implications of experimental GGE for the genetically related future children. It also important to open a social debate on the necessity of disclosing the use of GGE to the Chinese twins and individuals whose birth involved the technique if we wish to guarantee their fundamental rights.

Author Disclosure Statement

No competing financial interests exist.

Funding Information

This work was supported by the Uehiro Foundation on Ethics and Education (T.I.), the JSPS KAKENHI Project Innovative Ethics (Graduate School of Humanities, Kobe University, Kobe, Japan; T.I.), and the Government of the Basque Country (Grant IT-1066-16; I.M.B.).

References

1. Frankel MS. CAR. Human inheritable genetic modifications. Assessing scientific, ethical, religious, and policy issues. Available online at www.aaas.org/sites/default/files/s3fs-public/germline.pdf (last accessed March 29, 2018).
2. Ishii T. Mitochondrial manipulation for infertility treatment and disease prevention In: Schatten H, ed. Human Reproduction: Updates and New Horizons. Hoboken, NJ: Wiley Blackwell, 2016:205–230.
3. Ishii T. Reproductive medicine involving mitochondrial DNA modification: evolution, legality, and ethics. *EMJ Repro Health* 2018;4:88–99. Available online at www.emjreviews.com/reproductive-health/article/reproductive-medicine-involving-mitochondrial-dna-modification-evolution-legality-and-ethics (last accessed XX-XX-XXXX)
4. Chen SH, Pascale C, Jackson M, et al. A limited survey-based uncontrolled follow-up study of children born after ooplasmic transplantation in a single centre. *Reprod Biomed Online* 2016;33:737–744. DOI: 10.1016/j.rbmo.2016.10.003.
5. Dimond R, Stephens N. Legalising Mitochondrial Donation: Enacting Ethical Futures in UK Biomedical Politics. London: Palgrave Pivot, 2018.
6. Barrangou R, Doudna JA. Applications of CRISPR technologies in research and beyond. *Nat Biotechnol* 2016;34:933–941. DOI: 10.1038/nbt.3659.

7. Liang P, Xu Y, Zhang X, et al. CRISPR/Cas9-mediated gene editing in human tripronuclear zygotes. *Protein Cell* 2015;6:363–372. DOI: 10.1007/s13238-015-0153-5.
8. Ma H, Marti-Gutierrez N, Park SW, et al. Correction of a pathogenic gene mutation in human embryos. *Nature* 2017;548:413–419. DOI: 10.1038/nature23305.
9. Kang X, He W, Huang Y, et al. Introducing precise genetic modifications into human 3PN embryos by CRISPR/Cas-mediated genome editing. *J Assist Reprod Genet* 2016;33:581–588. DOI: 10.1007/s10815-016-0710-8.
10. McCaughey T, Sanfilippo PG, Gooden GE, et al. A global social media survey of attitudes to human genome editing. *Cell Stem Cell* 2016;18:569–572. DOI: 10.1016/j.stem.2016.04.011.
11. Carolyn B. Do CRISPR germline ethics statements cut it? *CRISPR J* 2018;1:115–125. DOI: 10.1089/crispr.2017.0024.
12. National Academies of Science, Engineering, and Medicine. Human Genome Editing: Science, Ethics, and Governance. Washington, DC: National Academies Press, 2017.
13. Nuffield Council on Bioethics. Genome editing and human reproduction: social and ethical issues. Available online at http://nuffieldbioethics.org/project/genome-editing-human-reproduction (last accessed January 21, 2019).
14. He J. CCR5 gene editing in mouse, monkey and human embryos using CRISPR/Cas9, in Session 3 Human Embryo Editing at Second International Summit on Human Genome Editing. Available online at http://nationalacademies.org/gene-editing/2nd_summit/second_day/index.htm (last accessed January 21, 2019).
15. Xinhua. Guangdong releases preliminary investigation result of gene-edited babies. Available online at www.xinhuanet.com/english/2019-01/21/c_137762633.htm (last accessed January 21, 2019).
16. Hacein-Bey-Abina S, Garrigue A, Wang GP, et al. Insertional oncogenesis in 4 patients after retrovirus-mediated gene therapy of SCID-X1. *J Clin Invest* 2008;118:3132–3142. DOI: 10.1172/jci35700.
17. Howe SJ, Mansour MR, Schwarzwaelder K, et al. Insertional mutagenesis combined with acquired somatic mutations causes leukemogenesis following gene therapy of SCID-X1 patients. *J Clin Invest* 2008;118:3143–3150. DOI: 10.1172/jci35798.
18. Hacein-Bey-Abina S, Pai SY, Gaspar HB, et al. A modified gamma-retrovirus vector for X-linked severe combined immunodeficiency. *New Engl J Med* 2014;371:1407–1417. DOI: 10.1056/NEJMoa1404588.
19. Cho SW, Kim S, Kim Y, et al. Analysis of off-target effects of CRISPR/Cas-derived RNA-guided endonucleases and nickases. *Genome Res* 2014;24:132–141. DOI: 10.1101/gr.162339.113.
20. Araki M, Ishii T. Providing appropriate risk information on genome editing for patients. *Trends Biotechnol* 2016;34:86–90. DOI: 10.1016/j.tibtech.2015.12.002.
21. Baylis F. The ethics of creating children with three genetic parents. *Reprod Biomed Online* 2013;26:531–534. DOI: 10.1016/j.rbmo.2013.03.006.
22. Rulli T. What is the value of three-parent IVF? *Hastings Cent Rep* 2016;46:38–47. DOI: 10.1002/hast.594.
23. Brandt R, Wilkinson S, Williams N. The Donation and Sale of Human Eggs and Sperm. In: Zalta EN, ed. The Stanford Encyclopedia of Philosophy (Summer 2017 Edition). Available online at https://platostanfordedu/archives/sum2017/entries/gametes-donation-sale/ (last accessed March 11, 2019).
24. Frith L, Blyth E, Crawshaw M, et al. Secrets and disclosure in donor conception. *Sociol Health Illn* 2018;40:188–203. DOI: 10.1111/1467-9566.12633.
25. Becker G, Butler A, Nachtigall RD. Resemblance talk: a challenge for parents whose children were conceived with donor gametes in the US. *Soc Sci Med* 2005;61:1300–1309. DOI: 10.1016/j.socscimed.2005.01.018.
26. Ishii T. Reproductive medicine involving genome editing: clinical uncertainties and embryological needs. *Reprod Biomed Online* 2017;34:27–31. DOI: 10.1016/j.rbmo.2016.09.009.
27. Vassena R, Heindryckx B, Peco R, et al. Genome engineering through CRISPR/Cas9 technology in the human germline and pluripotent stem cells. *Hum Reprod Update* 2016;22:411–419. DOI: 10.1093/humupd/dmw005.
28. Evitt NH, Mascharak S, Altman RB. Human germline CRISPR-Cas modification: toward a regulatory framework. *Am J Bioeth* 2015;15:25–29. DOI: 10.1080/15265161.2015.1104160.
29. Tsai DF. Human embryonic stem cell research debates: a confucian argument. *J Med Ethics* 2005;31:635–640. DOI: 10.1136/jme.2005.011924.
30. Church G. Compelling reasons for repairing human germlines. *New Engl J Med* 2017;377:1909–1911. DOI: 10.1056/NEJMp1710370.
31. International Theological Commission. Communion and stewardship: human persons created in the image of God. Available online at www.vatican.va/roman_curia/congregations/cfaith/cti_documents/rc_con_cfaith_doc_20040723_communion-stewardship_en.html (last accessed January 21, 2019).
32. Berer M. Abortion law and policy around the world: in search of decriminalization. *Health Hum Rights* 2017;19:13–27.
33. Sullivan N, de Faoite E. Psychological impact of abortion due to fetal anomaly: a review of published research. *Issues Law Med* 2017;32:19–30.
34. Ishii T, Hibino Y. Mitochondrial manipulation in fertility clinics: regulation and responsibility. *Reprod Biomed Soc Online* 2018;5:93–109. DOI: https://doi.org/10.1016/j.rbms.2018.01.002.
35. National Human Genome Research Institute. Human genomic variation: understanding what makes each of us different and what makes us the same. Available online at www.genome.gov/27570931/april-06-human-genomic-variation/ (last accessed January 21, 2019).
36. Human Fertilisation and Embryology Authority. Mitochondria replacement consultation: advice to government. 2013. Available online at www.hfea.gov.uk/media/2618/mitochondria_replacement_consultation__advice_for_government.pdf (last accessed October 29, 2019).
37. Cwik B. Designing ethical trials of germline gene editing. *New Engl J Med* 2017;377:1911–1913. DOI: 10.1056/NEJMp1711000.
38. Human Fertilisation and Embryology Authority. Code of Practice Edition 9.0: 33. Mitochondrial donation. Follow-up arrangements. Available online at www.hfea.gov.uk/code-of-practice/33 (last accessed January 21, 2019).
39. Frati P, Fineschi V, Di Sanzo M, et al. Preimplantation and prenatal diagnosis, wrongful birth and wrongful life: a global view of bioethical and legal controversies. *Human Reprod Update* 2017;23:338–357. DOI: 10.1093/humupd/dmx002.
40. Vanneste E, Voet T, Le Caignec C, et al. Chromosome instability is common in human cleavage-stage embryos. *Nat Med* 2009;15:577–583. DOI: 10.1038/nm.1924.
41. Rubeis G. Human germline genome editing in the clinical context: the case of disease prevention. In: Braun M, Schickl H, Dabrock P, eds. Between Moral Hazard and Legal Uncertainty: Ethical, Legal and Societal Challenges of Human Genome Editing. Wiesbaden, Germany: Springer, 2018:149–160.

The Geneva Statement on Heritable Human Genome Editing: a criticism

1.- Introduction

In January 2019, an international group comprising public interest advocates, policy experts, bioethicists, and scientists, met at the Brocher Foundation near Geneva, Switzerland, to assess and discuss public engagement and the governance of heritable human genome editing. The outcome of this meeting was a statement (the "Geneva Statement on Heritable Human Genome Editing: The Need for Course Correction") that has been published recently by Trends in Biotechnology [1]. According to its signatories, this new document is aimed at reorienting the conversation around heritable human genome editing *"by identifying misrepresentations and misunderstandings that muddy the discourse and by encouraging a robust consideration of the social, historical, and commercial contexts that would influence the development of heritable human genome editing and shape its societal effects."*

Reading these laudable intentions, I could not help but feel pleased. The debate about germline gene editing is indeed riddled with confusion. That a group of colleagues as highly qualified as those who signed it set out to resolve these issues seemed hopeful to me. Unfortunately, my expectations were soon frustrated. In my opinion, this new declaration not only fails to achieve its objective but, rather, contributes significantly to maintaining and even increasing the problems it seeks to alleviate. In this short comment I justify my criticism by exposing the weaknesses that can be found in the fundamental basis of the Statement, this is, the main assumptions made in its "Clarifying Misconceptions" section.

2.- Does heritable human genome editing "treat, cure or prevent disease in any existing matter"?

The first bullet point of the "Clarifying Misconceptions" section reads: "*Heritable human genome editing would not treat, cure, or prevent disease in any existing person*

(...)[It] should be understood not as a medical intervention, but as a way to satisfy parental desires for genetically related children or for children with specific genetic traits." However, this paragraph does not accommodate well with reality.

To begin with, the first phrase sounds quite enigmatic to me. What does it really mean? At first glance, it seems to take for granted that the embryos subjected to germline gene editing (GEE onwards) are not "existing" persons. However, there are millions of people who would disagree with a statement that is supposed to be "consensual". I am, of course, thinking about those people who believe that human life starts at fecundation. Moreover, this belief constitutes, for instance, a crucial point for institutions such as the Catholic Church (Indeed, do all the signatories really share the idea that embryos are not people?). Therefore, it seems to me that this statement is somewhat extreme to constitute a kind of common basis to build upon.

Moreover, even if this were not the case, that is, even if we were to accept that the GGE would not treat, cure, or prevent disease in any existing person because embryos cannot be considered as such, this does not mean that it would not produce this result in people who will exist in the future. In fact, if it ever works in the way it is expected, GGE might reduce someone's genetic predisposition to some types of cancer or prevent a person from transmitting hemophilia or suffering from Huntington Disease. Are these reasons not strong enough to support an intervention? Are they not, in fact, the reasons that justify the use of preimplantation genetic diagnosis (PGD), or that make in utero interventions to improve the health of a fetus (which is also not an existing person by the way) a moral obligation? [2,3] But if this is so, what is the sense in concluding - as this part of the Statement does -that *"Heritable human genome editing should be understood not as a medical intervention, but as a way to satisfy parental desires for genetically related children or for children with specific genetic traits"*? Is this really the case in all circumstances? I, frankly, do not share this idea, and I would safely assume that I am not the only one. Therefore, trying to build a consensus on these statements does not seem to be in any way constructive.

3.- Risk and Prospective parents' choices

The following bullet points in the "Clarifying Misconceptions" section are also very unconvincing. The second states that *"Modifying genes in early embryos, gametes,*

or gamete precursor cells could produce unanticipated biological effects in resulting children and in their offspring, creating harm rather than preventing it." This is true, of course. That is why it is deeply inmoral to implement these techniques right now. However, what the statement does not say is that one day that particular risk may be much lower than it is today. So much so, in fact, that what may constitute harm, in this case produced by omission, will be the failure to modify the genome of a human being who will, as a consequence, be much more susceptible to suffer from cancer [4].

Similarly, the third bullet point states that *"Prospective parents at risk of transmitting a genetic condition already have several options to avoid doing so, should they find them acceptable. For example, prospective parents may seek to have unaffected children via third-party gametes or adoption."* This quite naïve statement raises two important issues. First, such alternatives would never serve to provide genetically related descendance to all those people who suffer from the impossibility to reach this aim without using biotechnology. Indeed, in some -certainly scarce- cases their disability could only be solved through GGE [5]. Second, if we hold the signatories' suggestion, we should also claim for a general avoidance of PGD, since it is pretty clear that third-party gametes or adoption would also serve for this same purpose. I cannot find any substantial difference in both cases. However, is this really what some of the signatories of this Statement are willing to claim? Are they therefore claiming for a general ban of PGD? This might be the case, but I do not think we would find a general consensus on this either.

3.- GE and PGD

However, it is the final bullet point, in the section that I am commenting on, that concerns me the most, since it involves a substantial number of misunderstandings and half-truths. It is true, of course, that *"prospective parents at risk of transmitting a genetic condition who wish to avoid doing so and to have genetically related children can accomplish this with the existing embryo screening technique preimplantation genetic diagnosis (PGD)"*. What is not so true is that PGD and GGE are (or might be) equally useful [5,6]. GGE might allow us to obtain results that PGD could never provide, such as reducing predisposition to certain diseases or improving our immune system, for example. Omitting this information means omitting a fundamental part of the basis to debate the issues at stake.

Furthermore, even though the signatories state the opposite, it is not so clear that *"Genome editing cannot be considered an alternative to PGD, because PGD would remain a necessary step in any embryo editing procedure."* If one day GGE proves to be truly efficient and safe, we could edit all our embryos preventively, without having to examine them afterwards through PGD. Thus, for example, a family carrying a gene expression that predisposes to Huntington Disease could modify all their embryos and then proceed to direct implantation, without any PGD involved. Moreover, even if this were not the case, in a scenario in which GGE becomes real, PGD would lose all the eugenic component it currently has [7] because it would be executed to decide which embryos should be cured, not which ones should be destroyed. This, in my view, is another fundamental issue that the Statement completely ignores .

4.- Conclusion

From all that I have pointed out in the previous sections, I must conclude that the Geneva Statement on Heritable Human Genome Editing will hardly be able to reach a reasonable consensus on the issues involved in the GGE of human beings. Unfortunately, the description of the facts included in the document is very far from reflecting the consensus that we could all share. It neither shows the therapeutic possibilities of GGE nor describes well how it could improve the performance of PGD or alter its eugenic character, just to summarize some of its fundamental flaws. Therefore, even though I totally share the thought that *"we need to address and clarify several misrepresentations that have distorted public understanding of heritable human genome modification",* I do not share the idea that this document could serve for this purpose. We will, unfortunately, have to wait and see the fruit of future, and hopefully better balanced efforts.

5.- Acknowledgement

This project has received funding from the European Union's Horizon 2020 research and innovation programme under grant agreement No 788039 PANELFIT (this article reflects only the author's view and the Agency is not responsible for any use that may be made of the information it contains).

List of references

1. Andorno, R. et al, Geneva Statement on Heritable Human Genome Editing: The Need for Course Correction, *Trends in Biotechnology,* Published online 31 January 2020, available at: https://www.cell.com/trends/biotechnology/fulltext/S0167-7799(19)30317-8#articleInformation
2. Gyngell Christopher, Thomas, Douglas, and Julian, Savulescu. (2016). The ethics of germline gene editing. *Journal of Applied Philosophy* 34 (3): 1–16.
3. Savulescu, J., Pugh, J., Douglas, T. & Gyngell, C. (2015). The moral imperative to continue gene editing research on human embryos, *Protein & Cell*, 6 (7), p. 476. Available: https://link.springer.com/content/pdf/10.1007%2Fs13238-015-0184-y.pdf, accessed: January 14th 2020. DOI: 10.1007/s13238-015-0184-y
4. de Miguel Beriain, I. Human dignity and gene editing: Using human dignity as an argument against modifying the human genome and germline is a logical fallacy, *EMBO Reports*, 2018, 10 (19), p. 2. Available: https://www.embopress.org/doi/pdf/10.15252/embr.201846789, accessed: January 14th 2020. DOI:10.15252/embr.201846789
5. Ranisch, R. Germline genome editing versus preimplantation genetic diagnosis: Is there a case in favour of germline interventions? *Bioethics*. 2020; 34: 60–69. https://doi.org/10.1111/bioe.12635
6. Cavaliere, G. (2018). Genome editing and assisted reproduction: Curing embryos, society or prospective parents? Medicine, Health Care and Philosophy. Volume 21, Issue 2, pp 215–225. doi:10.1007/s11019-017-9793-y
7. MacKellar C, Bechtel C. The ethics of the new eugenics. New York: Berghahn Books; 2014
8. de, Miguel Beriain, I. and Ishii, T. Comment on "Should gene editing replace embryo selection following PGD? Some comments on the debate held by the International Society for Prenatal Diagnosis". *Prenatal Diagnosis*, 2019, 39: 1170-1172. doi:10.1002/pd.5542

Query Details Back to Main Page

1. The citation "De Miguel Beriain 2019" has been changed to "De Miguel Beriain, 2019a" to match the author name/date in the reference list. Please check if the change is fine in this occurrence and modify the subsequent occurrences, if necessary.

This is fine with me, thank you

2. References [De Miguel Beriain, 2017, Raposo, 2012] were provided in the reference list; however, this was not mentioned or cited in the manuscript. As a rule, all references given in the list of references should be cited in the main body. Please provide its citation in the body text.

I would prefer that you delete both references in the list of references, if possible be.

Human Dignity and Gene Editing: Additional Support for Raposo's Arguments

Iñigo de Miguel Beriain, [1,2]✉

Email inigo.demiguelb@ehu.eus

Begoña Sanz, [3,4]

[1] Chair in Law and the Human Genome Research Group, Department of Public Law, University of the Basque Country, UPV/EHU, Barrio Sarriena s/n, Leioa, Bizkaia, Spain

[2] IKERBASQUE, Basque Foundation for Science, Bilbao, Spain

[3] Department of Physiology, Faculty of Medicine and Nursery, University of the Basque Country (UPV/EHU), Barrio Sarriena s/n, E-48940 Leioa, Bizkaia, Spain

[4] Biocruces Bizkaia Health Research Institute, Plaza de Cruces 12, E-48903 Barakaldo, Bizkaia, Spain

Received: 3 April 2019 / Accepted: 26 February 2020

Abstract

The aim of the present paper is to reinforce some of the affirmations made by Vera Lucia Raposo in a recent paper published by the *Journal of Bioethical Inquiry*. According to her, germline gene editing does not violate human dignity at all. This article offers some complementary ideas supporting her statement. In particular, four main arguments are stressed. Firstly, not only is the idea of human dignity unclear, but the idea of the human genome suffers from a general lack of concreteness, which has dramatic consequences for the debate. Secondly, it is highlighted that if we believe that the immutability of the human genome underpins human dignity, then it should be our duty to use the tools of genetic modification to reverse any accidental changes that occur in nature. Thirdly, it is showed that if the alteration of germline constitutes an attack on human dignity, then we should also refrain from performing medical practices such as chemotherapy, which cause precisely this effect. Finally, we argue that modification of germline is not contrary to human dignity but an excellent expression of our autonomy.

Keywords

Gene editing
Human dignity
CRISPR
Enhancement

Introduction

Recently, an interesting article was published in the *Journal of Bioethical Inquiry* by Vera Lucía Raposo (Raposo 2019), in which it was argued that genetic modification is not an offense against human dignity but a manifestation of that same dignity. In our opinion, this conclusion is entirely correct and this article will try to show that it should be generally accepted. The aim of the present paper is to reinforce some of the affirmations through discussion of complementary ideas. In particular, four main arguments are stressed. Firstly, not only is the idea of human dignity unclear, but the idea of the human genome suffers from a general lack of concreteness, which has dramatic consequences for the debate. Secondly, it is highlighted that if we believe that the immutability of the human genome underpins human dignity, then it should be our duty to use the tools of genetic modification to reverse any accidental changes that occur in nature. Thirdly, it is showed that if the alteration of germline[1] constitutes an

attack on human dignity, then we should also refrain from performing medical practices such as chemotherapy, which cause precisely this effect. Finally, it is argued that modification of germline is not contrary to human dignity but an excellent expression of our autonomy.

The Lack of Concreteness of the Idea of the Human Genome

It is typically highlighted that modification of the human germline is immoral, since it implies a change in the human genome, which is considered the source of dignity. This belief is often linked to article 24 of the Universal Declaration on the Human Genome and Human Rights, adopted by UNESCO on November 11, 1997, which states “practices that could be contrary to human dignity, such as germline interventions” (2). However, this clause does not necessarily define germline interventions as being against human dignity but simply claims that they could be, which is entirely different, since it allows distinction between modifications that violate human dignity and those that do not. Nevertheless, authors continue to have doubts regarding the association between the human genome and human dignity. Indeed, Raposo is correct when highlighting that these types of assertions are extremely difficult to contradict, since the concept of human dignity suffers from a general lack of definition (Raposo 2016; Macklin 2003). Therefore, the statement regarding a loss of human dignity due to changes in our genome cannot be held consistently.

However, it would be even more accurate to say that the concept of modification of the human genome is not clear either (De Miguel Beriain and del Cano 2018). Thus, it would be necessary to define whether any change in the genes of an individual's somatic line constitutes a modification of the human genome or whether this concept should be limited to modifications affecting the human germline. Moreover, it is also possible to think that there will only be a real change in the genome when the modification involves the introduction of a novelty in our gene pool[2] (National Academies of Sciences, Engineering, and Medicine 2017). For instance, if we limit ourselves to replacing a gene for Huntington’s disease in a human embryo, we would certainly be altering the germline of that embryo but not the human gene pool, which would remain the same (De Miguel Beriain 2019a). These differences are extremely important, since it would be possible to determine whether an action will affect human dignity (assuming that a change in the genome alters human dignity, which is not so easy to assume) depending on the notion of the human genome. Unfortunately, those who sustain this genome−dignity connection do not typically clarify this point.

AQ1

The Immutability of the Human Genome

Opponents of genetic modification often defend the idea that the human genome, as the foundation of dignity, is sacred and must be preserved in its natural state (Rifkin 1983; Kass 2004). As authors such as Raposo correctly highlight, this is a somewhat extravagant concept, since it seems to assume that the genome is a fixed and immutable entity. On the contrary, it is well known that our genes are constantly changing. Only in this way has it been possible to evolve—through natural selection—and adapt to the environment around us. Therefore, the appeal to respect the genome in its present form is, at the very least, difficult to understand. It is even less understandable that the proposed normative consequence of this belief is to refrain from modifying the human genome. Instead, those who are willing to preserve the human genome should postulate an active intervention in nature that would eliminate any novel gene combination in humans, since conservation of the human genome implies precisely that: preventing the changes (all changes, including the changes introduced by nature) from transforming the human gene pool. However, as already mentioned, they propose exactly the opposite, that is, to refrain from using gene editing at all. However, this is absurd, since if we do revoke the random alterations that nature produces, the human genome will undoubtedly vary.

Some may say that this refutation falsifies the main argument, since what is usually argued is not that the genome should be maintained in its present form but that we should not usurp the role of nature as the sole author of this change (Habermas 2003). However, these kinds of statements implicitly assume a belief that is difficult to share: that either nature is an entity endowed with intelligence and/or will that also watches over the destiny of humans; or that we must distrust human ingenuity to the point of preferring the logic of randomness to the intervention of humans in the configuration of the gene pool. The first of these two hypotheses contains an undoubted metaphysical element that is only compatible with some form of pantheism (De Miguel Beriain 2019b; Ereshefsky 2010). The second is decidedly pessimistic, since humans have been carrying out activities that interfere with nature in some way for centuries (such as vaccination, for example), and life has not worsened, in fact quite the opposite. In short, it does not seem that the refutation of the thesis put forward adds too much weight to our objection.

The Need to Refrain From Any Form of Attack on Human Dignity

Thirdly, those who support need to rid the human genome of any modification produced by humans do not fully assume the logical consequences of their claim. If they really believed in the soundness of their argument, they would have to

advocate the absolute and eternal prohibition of any human activity that produced such an effect (De Miguel Beriain 2018). Obviously, this would imply the need to veto treatments such as chemotherapy, which could affect the patient's germline, even if they do not currently have any efficient substitute (Isasi et al. 2016). However, the authors of this article are not aware of anyone adhering to this proposal, even though it is the only one consistent with the main postulate of the argument that is criticized in this article.

Faced with this accusation, some will argue that there is no comparison between an action that directly promotes the modification of the human genome and another action that only accepts this result as an inevitable consequence of an intervention that pursues a morally elevated goal, such as saving a human life (Costam 2019). However, this defence is not solvent. Its main defect is that it is only sustainable if we think that human dignity is a transactional good, such as life itself. One can give one's life to defend one's country, for example, or one can risk one human life in exchange for saving thousands of lives. This does not happen with human dignity, which is an absolute good, that is, a good that cannot be compared with any other good. We cannot renounce our freedom and enslave ourselves, no matter how much we do it to save human lives, for example. Consequently, we cannot accept conduct that alters dignity through modification of the human genome, even though this is only an indirect effect of the action. Once we are aware that the consequence of our intervention will be to undermine human dignity, there is no choice but to avoid it at all costs. Any other consideration is utterly absurd from a moral point of view.

Human Dignity and Autonomy

Finally, one last argument must be highlighted that reinforces Raposo's thesis, even though it was not explicitly expressed in the paper. The author rightly highlights that modification of germline might even be considered an excellent expression of our autonomy. Moreover, since such autonomy is the foundation of dignity, there is no immorality, but rather, intervention in our gene pool is a moral requirement for the benefit of individuals who will be affected by the change.

This statement would be considerably strengthened if we bear in mind that a non-intervention would also be an expression of our autonomy. For instance, if we choose not to save the life of a sick human because such intervention would alter the human genome, we are also expressing our autonomy, suggesting that it is not true that the action will increase it or the omission will diminish it. What has endowed us with greater autonomy has been the mere fact of having the power to intervene in our genome. Whether we exercise it or not will only affect

our moral elevation. Whether to save a life will generate an undoubted moral responsibility, regardless of which option is chosen. It is the power given to us by the techniques of genetic modification—not their exercise—that makes us more autonomous, and therefore more responsible. This is why it makes no sense to think that acting as if nothing has happened is the option that respects the idea of human dignity the most. We must react as responsible adult humans, recognize our new capabilities, and begin to use them (or not) according to our moral compass. Our dignity will not be altered by the decisions taken, but our moral status will surely be affected by the reasons involved in them.

Acknowledgements

Iñigo de Miguel Beriain's work was supported by the Government of the Basque Country, Grant IT-1066-16 and the EU Commission, H2020 SWAFS Programme, PANELFIT Project, research grant number 788039.

Funding

Iñigo de Miguel Beriain's work was supported by the Government of the Basque Country, Grant IT-1066-16 and the EU Commission, H2020 SWAFS Programme, PANELFIT Project, research grant number 788039.

References

AQ2

Costam, D. B. 2019. Comment on "Human dignity and gene editing". *EMBO reports* 20: e47220.

De Miguel Beriain, I. 2017. Legal issues regarding gene editing at the beginning of life: An EU perspective. *Regenerative Medicine* 12(6): 669–679.

———. 2018. Human dignity and gene editing. *EMBO Reports* 19(10): e46789.

———. 2019a. Should human germ line editing be allowed? Some suggestions on the basis of the existing regulatory framework. *Bioethics* 33(1): 105-111.

———. 2019b. Response by the Author. *EMBO reports* 20: e47346.

De Miguel Beriain, I. and A. Marcos del Cano. 2018. Gene editing in human embryos. A comment on the ethical issues involved. In *The ethics of reproductive genetics. Between utility, principles, and virtues*, edited by M. Soniewicka, 173-187. Springer.

Ereshefsky, M. 2010. What's wrong with the new biological essentialism. *Philosophy of Science* 77(5): 674-685.

Habermas, J. 2003. *The future of human nature*. Malden: Polity.

Isasi, R., E. Kleiderman, and B. M. Knoppers. 2016. Editing policy to fit the genome? *Science* 22: 337-339.

Kass, L. R. 2004. *Life, liberty and the defense of dignity*. San Francisco: Encounter Books.

Macklin, R. 2003. Dignity is a useless concept. *British Medical Journal* 327: 1419-1420.

National Academies of Sciences, Engineering, and Medicine 2017. *Human genome editing: Science, ethics, and governance*. Washington, DC: The National Academies Press.

Raposo, V. L. 2019. Gene editing, the mystic threat to human dignity. *Journal of Bioethical Inquiry* 16(2): 249-257.

———. 2016. The convention of human rights and biomedicine revisited: Critical assessment. *The International Journal of Human Rights* 20(8): 1277–1294.

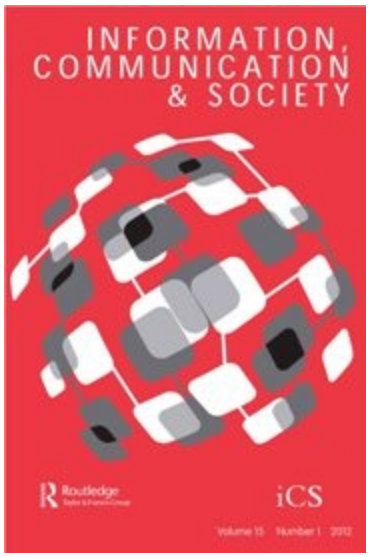
———. 2012. Biodireitos: The new kids on the juridical playground. In *Estudos em homenagem ao Pr of. Doutor José Joaquim Gomes Canotilho*, edited by F. A. Correia, J. Machado, et al., 1–833. Coimbra: Coimbra Editora.

Rifkin, J. 1983. *Algeny*. New York: Viking.

---

[1] In the context of this text, this term refers to the edition of human germline genome.

[2] In the text, this term refers to all possible genetic variations of the human species.



Machine Learning in the EU health care context: exploring the ethical, legal and social issues

Journal:	*Information, Communication and Society*
Manuscript ID	Draft
Manuscript Type:	Special Issue Paper
Keywords:	Big Data, ICTs, Law, Regulation, Systems Design, e-Health

Machine Learning in the EU health care context: exploring the ethical, legal and social issues

Abstract

The diagnosis and clinical decision making based on Machine Learning technologies are showing significant advances that may change the functioning of our health care systems. These advances promise more effective and efficient healthcare, at a lower cost. This may allow healthcare professionals to recover 'high-touch' time with their patients. The evidence suggests that all these promises have yet to be demonstrated in clinical practice, but what is undeniable is that these technologies are resignifying the relationships in the health landscape, particularly the physician-patient relationship, which we could already redefine as "physician-computer-patient relationship". Although it is true that today fully automated decision systems are scarce in comparison with integrative decision support systems, we cannot fail to observe the horizon they define. Our most recent regulatory framework, defined by the General Regulation on Data Protection, has tried to avoid this scenario by including the right not to be subject to a decision based solely on automated processing. In this paper, however, we argue that this legal tool is adequate but not sufficient to address the legal, ethical and social challenges that Machine Learning technologies pose to patients' rights and health care givers' capacities.

1. Introduction. AI As a New Promised Land for Health Care Systems: The Pearls and the Perils

In 1955, John McCarthy coined the term 'Artificial Intelligence' (AI) as the name for the science and engineering of making intelligent machines (Hamet & Tremblay, 2017). Today, AI focuses on how computers learn from data and mimic human thought processes (Noorbakhsh-Sabet et al., 2019). The field of medicine is one of the most promising application areas for AI (Yu et al., 2018), due to AI's ability to handle and optimise very complex data sets from very complex systems (Bini, 2018).

Software programs and Machine Learning (ML) are able to convert big data into algorithms, providing advantages such as flexibility and scalability; the ability to analyse diverse data types for disease risk, diagnosis, prognosis, and appropriate treatments; as well tackling unique challenges for model training and refinement; and managing the need for data pre-processing and making crucial ethical considerations (Ngiam & Khor, 2019).

ML algorithms can be classified into two groups. The first involves deep learning platforms, such as IBM's Watson Oncology -Dr. Watson-, which is fed everything written, in every language, at any time, that is related to cancer diagnostics and treatment (Londhe & Bhasin, 2019). The more information Watson has about a patient, the more accurate it's assistance will be; but actually, it is not yet perfect (Bini, 2018). In the second group, the algorithms fall under the denomination of pattern medicine, based on data collected through imaging techniques such as x-rays (Kallianos et al., 2019), mammogram images (Le et al., 2019), immunohistochemical stains (Niazi et al., 2019), and retinal images (Schmidt-Erfurth et al. 2018), among others. Some of the pattern medicine algorithms have been approved by U.S. Food and Drug Administration, and most of them have been validated by comparison to the precision exhibited by human

beings (Ngiam & Khor, 2019). Nevertheless, the future of AI in diagnosis and treatment should be based on hybrid strategies, since specific medical diagnostic and prognostic success for each concrete matter depends on the nature of the task, type of data, and available information about the related disease (Shahid & Singh, 2019).

All that being said, ML has limited exploratory power: algorithms might be able to identify correlations, but not necessarily prove causation. So, despite their differences, ML and evidence-based medicine can and should complement one another (Scott, 2018). In this scenario, the clinician's role is to be a bridge between machine and decision (Coeira, 2019), and professionals across different fields, speaking different languages, should be trained and integrated with the real benefits and applicability of developed algorithms in health care (Nuñez-Reiz, 2019).

The success of AI, therefore, can bring about a dramatic change in the way medicine is understood, and in the functioning and sustainability of public health systems. However, it also poses considerable challenges. To begin with, it inevitably affects the core of medical practice: the relationship between the care giver and the patient. The emergence of AI means that doctors must consider their own roles. They will not only be responsible for their patients' health; they will be managers of their patients' personal data, with a commensurate obligation to inform them about the use of automated decision-making systems that physicians do not fully understand and the recommendations of which they do not always share. In effect, doctors will be forced to rethink the way they manage the information at their disposal and the very idea of data confidentiality.

This new scenario may cause patients to feel helpless against the use of opaque tools and automated decision-making processes that affect essential aspects of their lives. Faced with this dilemma, the European Union has developed a regulatory framework focused on defending the rights of the data subject, in this case patients. Its General Data

Protection Regulation has proclaimed a patient's right to information and a right not to be subjected to profiling and automated decision-making processes which, it is hoped, will serve as an efficient mechanism to protect patients from the misuse of their data. However, this general framework shows some gaps and deficiencies that need to be clarified.

This paper is intended for this purpose. Its aim is to explain how the implementation of AI can pose problems for patients' and doctors' interests. It analyses the mechanisms created to address these issues, highlighting their weaknesses and incorporating suggestions on how to resolve them. We begin by showing the main technical obstacles that make the guarantee of adequate decision making by patients, doctors, and others responsible for health systems extremely complex. Then, we propose some measure that might contribute to face these challenges successfully.

2. Understanding AI: Intrinsic Issues That Render Transparency Highly Complicated

As discussed, the implementation of AI in health care systems will only respect patients' rights if patients are allowed to make the final decision on whether or not to use automated decision-making systems. However, this is very difficult if patients lack sufficient information, which should be provided by their physicians or health care providers. Achieving the goal of sufficient information transfer is complicated because there are multiple factors that seriously hinder an efficient transmission of information and subsequent decision making. These include: the difficulty of assimilating the paradigm shift introduced by AI; the (current) deficiencies of AI systems applied to the health care sector; the difficulty of reconciling business interests and transparency; and the shortcomings inherent in the construction of algorithms. In this section, we analyse each of these issues.

2.1 The difficulty of assimilating the paradigm shift introduced by Artificial Intelligence

The first issue involving the use of algorithms produced through Deep Learning tools is that the whole philosophy underlying their production differs substantially from the way science has been conducted at least since the main scientific methods were developed. Science generally advances through the formulation of hypotheses (i.e. possible causal associations between two events), which can be subjected to a consistency test through contrast with reality. It is true that in recent years the complexity of some disciplines has forced us to accept alternative models of scientific evidence (Sterky & Lundeberg, 2000), but these subtle exceptions to the general rule have not yet been assimilated by health professionals or their patients. In this particular part of science, the rigid rules of hypothesis-reality contrast continue to apply.

Luckily or regrettably, the algorithms do not fit this form of epistemological functioning. An algorithm does not formulate a hypothesis to contrast with data extracted from the real world, but rather the hypotheses are precisely the result of the analysis of these data. An algorithm only discovers correlations that can predict, not causalities that can explain. In this sense, AI, in its current development, is a complement rather than a substitute for science (Ellis & Silk 2014). If science is assumed to have both the ability to explain and the ability to predict, this part of AI is limited exclusively to the latter (Anderson, 2018). However, as we have anticipated, the value of AI lies in the fact that it is able to achieve acceptably accurate diagnoses with a more efficient use of resources, and much more accurate prognoses.

The question is whether this enormous limitation – the practical impossibility of giving a causal explanation for specific recommendations – will make the use of AI in health care acceptable to health professionals and patients. In the case of professionals,

the generalisation of AI will, to a large extent, be an amendment to their entire training, as they will often need to adjust their performance to a mechanism that does not provide them with reasons, but with probabilities. Therefore, the ability to interpret these probabilities clearly and sensitively represents an additional—and essential—educational demand for patients and their families (Wartman & Combs, 2019). In the case of patients, it seems at first glance that the situation may be less complex, but, in a world where conspiracy theories are becoming increasingly predictable, knowing what the reaction will be to the use of an eminently opaque technology is a mystery.

Faced with this situation, it is obvious that the key to transmitting adequate information lies in efficient training of health professionals, which does not exist at the moment. As Char et al. stated, 'Physicians who use machine learning systems can become more educated about their construction, the data sets they are built on, and their lim. HEtations. Remaining ignorant about the construction of machine-learning systems or allowing them to be constructed as black boxes could lead to ethically problematic outcomes' (Char, Shah, & Magnus, 2018). Thus, training is a key concept in terms of efficient information.

2.2. The (current) shortcomings of AI systems applied to the healthcare sector: The 'black box' medicine

The problems described in the previous section would probably be less important were it not for the fact that many of the algorithms developed by machine learning systems are inherently opaque tools. As Ferretti, Schneider and Blasimme (2018) have rightly described:

> "While most people recognize the promise of applying AI systems to medical diagnosis and decision-making, many are worried about the use of partly autonomous computer programs for medical purposes. This fear has to do with a

> characteristic of many ML methods. AI systems that incorporate ML learn with a varying degree of supervision which rules they need to follow in order to perform their task. The programmer sets up the system so that it can learn to do something. However, he or she does not decide, nor is necessarily aware of the rules the AI system has learnt and is following in order to do what it is supposed to do. This characteristic is often referred to as the opacity of ML. For the same reason, AI systems based on ML are often called black boxes, to stress that it is hard or even impossible for human users to open them up, so to say, and see for themselves what the machine is doing (or, which is the same, what rule the machine has learnt and is employing). The possibility that these systems could remain opaque to their own creators as well as to their end-users is a cause of concern."

The issue, in short, is simple: it is very difficult to talk about algorithmic transparency in the case of ML technologies because the operation of these techniques makes it almost impossible to understand how their inferences operate (de Miguel Beriain, 2018); not even their programmer could do it. Indeed, Consequently, the fact must be faced that there is a part of the information that is not available to patients, physicians, or health care providers. In this context, a field of research called explainable AI (xAI) is raising. It is aimed at producing methods that make algorithmic decision-making systems more trustworthy and accountable (Mittelstadt et al., 2019). Nevertheless, further work in this field is mandatory since explanatory systems are focused on programmers or IA experts, not on end users or policy-makers (Gilpin et al., 2018).

2.3. The difficulty of reconciling business interest and transparency

The description of the facts is not complete without an account of a third factor that contributes substantially to the difficulty of understanding the new reality faced by patients and health care givers. The companies that develop the algorithms invest considerable resources in their development. This includes both the need to procure large

and well-ordered databases -Smart data- and the development of the AI mechanism itself.

Non-public companies seek a return on that investment. Therefore, opacity is an intentional form of self-protection that attempts to keep trade secrets and the competitive advantages involved (Burrell, 2016). In other spheres of human activity, this is often achieved through mechanisms such as patents or copyright. In the case of data, the system of intellectual property protection offers notable shortcomings. Hence, within the EU, what is known as sui generis database rights, a property right settled by Directive 96/9/EC of 11 March 1996 on the legal protection of databases, according to which

> Member States shall provide for a right for the maker of a database which shows that there has been qualitatively and/or quantitatively a substantial investment in either the obtaining, verification or presentation of the contents to prevent extraction and/or re-utilization of the whole or of a substantial part, evaluated qualitatively and/or quantitatively, of the contents of that database (article 7.1).

Unfortunately, until now we have not been able to develop in parallel a 'sui generis algorithm right'. Therefore, algorithms continue to be considered as ideas; creations of the mind that do not find accommodation in the intellectual property protection regime, unless we accept the theses proposed by authors such as Minssen and Pierce (2018), who consider that patenting algorithms could be possible in the EU arena. Otherwise, companies have no choice but to hide their algorithms under the trade secret layer in order to maximize their returns. Consequently, the inherent opacity is often exacerbated by this deliberately sought-after form of opacity.

2.4. Inherent flaws in algorithm construction

Finally, distrust of algorithms is by no means unjustified. Evidence shows that machine learning algorithms are often biased and may lead to discrimination based on classes like race and gender (Buolamwini & Gebru, 2018). The content of the dataset determines how

the algorithm will make decisions on real-world cases (Wellner & Rothman, 2019). Therefore, significant problems arise from errors and biases latent in data training sets that tend to be reproduced in the outputs of these tools (Zerilli et al., 2018). For example, a database comprised mostly of information about white males will surely produce an algorithm much less accurate for Hispanic women. In other cases, failures stem from the deficiencies generated by a machine learning system that induces unlucky correlations through the incorporation of a human collective thinking system that cannot avoid being biased. As Char et al. have rightly pointed out, 'Subtle discrimination inherent in health care delivery may be harder to anticipate; as a result, it may be more difficult to prevent an algorithm from learning and incorporating this type of bias' (Char et al., 2018).

Patients may therefore legitimately ask whether the algorithm being used to make a diagnosis or assess their response to treatment is adequately adapted to their personal circumstances, or whether it is not. Unfortunately, these questions can only be clarified if care givers ensure that the AI mechanisms have been subjected to validation methods and monitoring systems capable of verifying that there are no biases or errors incompatible with their use in the health care system. And a care giver, of course, will hardly be able to provide the patient with any information other than whether or not these quality control systems have been implemented.

3. Protection of data subjects under EU legislation with respect to AI applications. Right to information and prohibition on fully automated decisions

On the basis of the above limitations (that the operating logic of AI differs substantially from that of science, and that the algorithms are inherently opaque), the legislator has attempted to protect data subjects (patients, in this case) without banishing AI. At the EU level, this attempt has resulted in the development of two normative initiatives: the

proclamation of a fundamental right to information and the prohibition of solely automated decision-making. In the following sections, we will attempt to set out the fundamental bases of these policies, and their limitations.

3.1 The right to information

One of the first issues involved in giving patients adequate information about the use of AI in the health care process is that patients need to be made aware that these mechanisms are being used to make decisions that affect them personally. In some cases, it will be easy to guess this, as in order to use AI efficiently it will be necessary to request a huge amount of data from the patient, making it very complex to hide its use. In other cases, however, the controller could use only data that were already available, such as the data already included in the patient's clinical history. It is, therefore, necessary to avoid this possibility by making a rule that obliges the data controller to inform the patient of the intervention of AI mechanisms in decision-making. This is what the Regulation provides for in Articles 13 and 14. Article 13 (Information to be provided where personal data are collected from the data subject) states:

> "1. Where personal data relating to a data subject are collected from the data subject, the controller shall, at the time when personal data are obtained, provide the data subject with all of the following information: (…) c) the purposes of the processing for which the personal data are intended as well as the legal basis for the processing."

Furthermore, its number 2 states that

> "2. In addition to the information referred to in paragraph 1, the controller shall, at the time when personal data are obtained, provide the data subject with the following further information necessary to ensure fair and transparent processing: (...) g) the existence of automated decision-making, including profiling, referred to in Article 22(1) and (4) and, at least in those cases, meaningful information about the logic

> involved, as well as the significance and the envisaged consequences of such processing for the data subject"

Thus, the Regulation grants the patient the right to be fully aware of the use of personal data collected by the controller if these data are to be used for automated decision-making purposes. It is necessary to point out that the Regulation does not use the term 'solely automated decision-making', but only 'automated decision-making'. This seems reasonable, since otherwise the obligation to communicate the fact that an AI tool would be involved would be reduced if the process included some form of human supervision. In this way, the GDPR confronts the secret use of automated decision systems, which has been claimed to be harmful (O'Neil, 2016): every patient has the right to know if her personal data has been subjected to this kind of automated processing. It is important, on the other hand, to underline the fact that this obligation applies not only to automated decision-making but also to profiling, that is:

> "any form of automated processing of personal data consisting of the use of personal data to evaluate certain personal aspects relating to a natural person, in particular to analyse or predict aspects concerning that natural person's performance at work, economic situation, health, personal preferences, interests, reliability, behaviour, location or movements"[art. 4].

The provisions of Article 13, which relate to personal data provided by the subject to the controller, are complemented by those of Article 14, which applies to information to be provided where personal data have not been obtained from the data subject. This clause states that, for such data, controllers must also inform the data subject about the purposes of the intended processing of the personal data as well as the legal basis for that processing [Art. 14.1.c], and about the existence of automated decision-making, including profiling, and *'meaningful information about the logic involved, as well as the significance and the envisaged consequences of such processing for the data subject'*.

To summarize, the Regulation has shaped a scenario in which the right to an explanation about the use of AI tools for profiling or automated decision-making plays a dominant role. This is of crucial importance in terms of the reliability of the health system, as it avoids reasonable suspicions about the ultimate purpose of introducing AI into the process. Moreover, the proclamation of this right contributes to a reinforcement of the trust between the health personnel (who can exercise the role of the data controller, as has been said) and the patient. With regard to this caregiver–patient relationship, it is mandatory to evaluate the impact that the introduction of AI in clinical decision-making will have. The right to information avoids giving patients the impression that *"they are being marginalized in decisional processes regarding their health, thus affecting their decisional autonomy and their sense of self-determination. In light of these considerations, restricting disclosure to solely-automated activities may turn out to be insufficient"* (Ferreti et al., 2018).

However, this apparently strong legal structure hides some important holes, mainly related to the content of this general right, proclaimed in the legislation, to receive an explanation. What does this right mean in practical terms? Does it mean that patients are given a right to know about the technicalities of the decision-making tool? Does it only mean that they should be informed that an AI tool will be used? In the legal arena, this issue has raised a profound discussion, which is still far from being resolved (Brkan, 2019; Goodman & Flaxman, 2017; Selbst & Powles, 2017; Wachter, Mittelstadt & Floridi, 2018). To enter into the subtleties of this discussion would clearly go beyond the boundaries of this text. However, we believe it is possible to set out some of the issues that seem most pertinent now.

3.2. What the right to explanation is not: a right to disclosure

First, we must highlight that the right to an explanation by no means implies that the data

subject is empowered to have access to the algorithm as such. This would clearly render industrial secrecy impossible and would deprive the developer of the algorithm of any way to exploit the result of his investment commercially. This result is unacceptable for both legal and practical reasons. From the legal point of view, it would contradict the spirit of the Regulation, whose Recital 63 states *"that right should not adversely affect the rights or freedoms of others, including trade secrets or intellectual property and in particular the copyright protecting the software"*.

As Ferretti et al. (2018) wrote,

> "It follows that, while data controllers must disclose that they are conducting profiling or automated data processing, they are not obliged to reveal all details about their AI systems. In practical terms, this entails that data controllers may still be required to provide information regarding the general characteristics of their system, but they may not be compelled to explain what rules the AI system follows, how it has reached a conclusion, or how it has taken a given decision about a particular data subject."

Moreover, from a practical point of view, disclosing the algorithm would just provide patients with information that they could not really understand, a situation that is far removed from their needs and from the spirit of the Regulation. Indeed, one must consider that information about the logic must be meaningful to the data subject, who is, notably, a human being who can be presumed to have no particular technical expertise (Selbst & Powles, 2017).

Therefore, we must conclude that the right to an explanation does not include disclosure and, furthermore, that the right could not be satisfied by disclosing the algorithm (that is, the controller would not comply just by providing the patient with the algorithm used in the automated decision-making). Obviously, this does not mean that

controllers can rely on the protection of their trade secrets as an excuse to deny access, and nor can they refuse to provide information to the data subject (A29WP, 2018).

3.3. What the right to explanation must include: a right to know the type of information that is being used and the general principles involved in the design of the algorithm

In our opinion, patients may, in short, assume that they will probably never know exactly how the algorithmic mechanism that will intervene in a crucial decision in their life works. This is not necessarily new; the sorts of explanations we cannot obtain from AI are the same as those we cannot obtain from humans either (Zerilli et al., 2018). However, this does not mean that patients cannot be provided with any form of relevant information. Indeed, there are some fruitful ways to guarantee that the explanation is sufficient to facilitate the exercise by patients of the rights granted to them by the GDPR and human rights law. For instance, Article 12 emphasises intelligibility and contains the requirement that '[t]he controller shall facilitate the exercise of data subject rights' (Selbst & Powles, 2017).

To begin with, it is perfectly possible to provide a layperson with general information about how an algorithm has been constructed or what type of data categories it uses. This has been understood, for example, by the Article 29 Data Protection Working Party, an advisory body made up of a representative from the data protection authority of each EU Member State, which played a prominent role in terms of the interpretation of the Regulation until it was replaced by the European Data Protection Board (EDPB) under the GDPR. The Working Party has stated:

> "Article 15 gives the data subject the right to obtain details of any personal data used for profiling, including the categories of data used to construct a profile. In addition to general information about the processing, pursuant to Article 15(3), the controller has a duty to make available the data used as input to create the profile as well as

> access to information on the profile and details of which segments the data subject has been placed into." (A29WP, 2018)

Similarly, patients must be made aware of the importance of the contribution made by the AI system in the final decision, including receiving all available information on the main factors in the decision, whether changing a certain factor would or would not have changed the decision, and why different decisions are reached in similar-looking cases, or the same decision in different-looking cases (Doshi-Velez et al., 2017). On the one hand, this also means that from the very first moment patients should know about the use that might be made of their data and the foreseeable consequences of the data processing for this purpose, as, indeed, is required by the Regulation; however, this requirement could be very limited in an actual scenario of big data analytics, where new data are created from inferred and derived data. Looking at how the automated processing of data and profiling works, it is undeniably true that the GDPR focuses primarily on mechanisms to manage the input side of the processing, and that the legal mechanisms that address the outputs of the processing, including inferred and derived data, profiles, and decisions, are far weaker (Wachter & Mittelstadt, 2019). On the other hand, it also means that physicians and/or health care providers must explain to patients the weight that automated decision-making and profiling represented in their final decision, and provide understandable explanations for why the automated decision-maker's suggestions were or were not followed. It might happen, indeed, that physicians have to confess that the only reason they followed the machine's advice is that they could simply find no justification to contradict its opaque conclusion. But, if this is the case, this information and no other should be shared with the patients. For this purpose, a flexible, functional approach will be most appropriate for understanding the term 'meaningful information' that is included in the right to an explanation (Selbst & Powles, 2017).

3.4. But… the issues that remain

The construction of an apparently sound legislative framework, such as the one we have described, will not, however, serve to address all the problems that the introduction of AI will bring to the management of health information. To begin with, it is difficult to know how it will be possible to reconcile a patient's right to restrict the use of his or her health data with increasingly automated health systems. If in the future most decisions are made on the basis of AI recommendations, patients who refuse to provide their data for that purpose will have to rely on physicians who will probably have lost some of the skills of traditional medicine. Thus, the configuration of the medicine of the future may end up dividing patients into two groups, those who are reconciled to the use of their data in AI systems and those who refuse to take this step. It is not clear what the consequences of this division will be, or whether we should start warning of these dangers right now.

From the doctor's point of view, the introduction of AI creates a growing challenge in terms of the concept of confidentiality and the fiduciary relationship between a patient and a physician. As Char et al. have written,

> In the era of electronic medical records, the traditional understanding of confidentiality requires that a physician withhold information from the medical record in order to truly keep it confidential. Once machinelearning–based decision support is integrated into clinical care, withholding information from electronic records will become increasingly difficult, since patients whose data aren't recorded can't benefit from machine-learning analyses. The implementation of machine-learning systems will therefore require a reimagining of confidentiality and other core tenets of professional ethics. What's more, a learning health care system will have agency, which will also need to be factored into ethical considerations surrounding patient care." (Char et al., 2018)

Third, even if it is convenient that physicians' skill sets include collaborating with and managing AI devices that aggregate big data (Wartman & Combs, 2019), one cannot

ignore the fact that it will be hard for physicians to acquire all the technical capacities needed to provide accurate information about those devices directly. Therefore, taking care of these issues may take us into a highly undesirable scenario in which patients do not receive accurate information and physicians are stressed by the need to perform tasks they are not trained to perform. In our view, this could be prevented if we let physicians primarily communicate to the patient how the use of AI has influenced their diagnosis or choice of treatment, including the reasons that would have supported that conclusion. It would be better if health care providers could designate other professionals who are more familiar with AI to convey technical information about how AI works in each particular case. For this reason, the creation of new roles, such as that of health information counsellors (HICs) (Fiske, Buyx, & Prainsack, 2018), is of particular interest. These counsellors would be professionals with a broad knowledge of various kinds of health data and data quality evaluation techniques, as well as analytical skills in statistics and data interpretation, who could offer patients information about AI much more efficiently than health care givers. As Fiske et al. (2018) propose, *'trained also in interpersonal communication, health management, insurance systems, and medico-legal aspects of data privacy, HICs would know enough about clinical medicine to advise on the relevance of any kind of data for prevention, diagnosis, and treatment'*. Therefore, both patients and physicians would profit from the intervention of this new role.

Last, but not least, we must keep in mind that the right to information concerns not only what we have historically referred to as health data, but also what the GDPR calls data concerning health: *'personal data related to the physical or mental health of a natural person, including the provision of health care services, which reveal information about his or her health status'* (Article 4.15). According to this definition, the concept extends to an increasing variety of data generated and collected outside the clinical

setting, such as lifestyle data, data about dietary habits, socio-economic data, and data included in patients' health records, but also data collected through smartphones, direct-to-consumer testing, online platforms, apps, and wearables (Frisse, 2016). This means that the obligation to provide explanations may extend to data controllers who are not health care providers as such. If this is the case, we should design new policies regarding informed consent that apply to the use of these devices and deal with the obligations to which these providers are subject. Quite a number of tasks to perform there.

4. The general prohibition on fully automated individual decision-making

The General Data Protection Regulation has directly addressed its concern for decisions based solely on automated data processing, especially when it affects special categories of data, a concept which includes 'data concerning health' (article 9.1). In this sense, its Recital 71 states that

> The data subject should have the right not to be subject to a decision, which may include a measure, evaluating personal aspects relating to him or her which is based solely on automated processing and which produces legal effects concerning him or her or similarly significantly affects him or her (…) Such processing includes 'profiling' that consists of any form of automated processing of personal data evaluating the personal aspects relating to a natural person, in particular to analyse or predict aspects concerning the data subject's performance at work, economic situation, health, personal preferences or interests, reliability or behaviour, location or movements, where it produces legal effects concerning him or her or similarly significantly affects him or her (…) In any case, such processing should be subject to suitable safeguards, which should include specific information to the data subject and the right to obtain human intervention, to express his or her point of view, to obtain an explanation of the decision reached after such assessment and to challenge the decision (…) Automated decision-making and profiling based on special categories of personal data should be allowed only under specific conditions.

The binding part of the Regulation reflects the intentions made in the Recital in its article 22, which is quite complex, but might be summarized by stating that *"the data subject shall have the right not to be subject to a decision based solely on automated processing, including profiling, which produces legal effects concerning him or her or similarly significantly affects him or her"* (article 22.1). This clause does not rule whether the decision is: (a) necessary for entering into, or performance of, a contract (b) authorised by Union or Member State law or (c) based on the data subject's explicit consent. Additionally, where the automated processing is based on special categories of personal data, such as data concerning health, data subjects have to explicitly consent to the use of such data or processing needs to be justified by a substantial public interest, and the data controller must adopt suitable measures to safeguard the data subject's rights and freedoms and legitimate interests (article 22.4). As we have seen above, according to Recital no. 71 those measures include providing specific information to the data subject and the right to obtain human intervention, to express his or her point of view, to obtain an explanation of the decision reached after such assessment and to challenge the decision.

A first crucial issue that seemed unclear is the nature of this right (Brkan, 2019). It could be understood both as a right to object, where automated decision-making is restricted only to cases in which the data subject actively objects, or as a prohibition, where data controllers will not be allowed to make automated decisions about a data subject until one of the legal requirements is met (Wachter et al., 2017). This understanding is crucial since it is in no way realistic to believe that there is effective control of personal information through consent – or objection in this specific case – and the rights that complement it (Cotino, 2017). Fortunately, this point has been addressed

the Article 29 Data Protection Working Party, the guidelines on automated individual decision-making and profiling declared *that a general prohibition on this type of processing exists to reflect the potential risks to individuals' rights and freedoms*.

Therefore, unless we met a legal exception to the general prohibition, there is no room for solely automated decision-making in the EU zone if it *produces legal effects concerning him or her or similarly significantly affects him or her*. This seems to be the case of most of the clinical decisions, even if it remains arguable what kind of decision-making significantly affects such an individual (Brkan, 2019). The risk categorization framework proposed by the FDA for the use of AI systems (U.S. Food & Drug Administration, 2019), based on the state of healthcare situation or condition of the patient and the significance of the information provided by the system to the healthcare decision, might be useful in this scenario.

However, the Regulation also fails to make clear what counts as a decision based solely on automated decision-making. Indeed, one cannot deduce from the literacy of the clauses what kind of human intervention is needed to make the difference between automated and solely automated decision-making. The only concretion made by the Regulation is that *"the controller must put in place suitable measures to safeguard the data subject's rights and freedoms and legitimate interests"*. This, definitively, does not provide too much concretion (Zarsky, 2017), what is certainly worrying, particularly if the condition 'solely' in 'decisions made solely by automation' is interpreted narrowly, because the safeguards and associated requirements of meaningful information will have limited applicability (Andrew et al., 2017).

Once again, the Article 29 Data Protection Working Party has provided some clarifications on what should be understood by solely automated decision-making,

> "The controller cannot avoid the Article 22 provisions by fabricating human involvement. For example, if someone routinely applies automatically generated profiles to individuals without any actual influence on the result, this would still be a decision based solely on automated processing. To qualify as human intervention, the controller must ensure that any oversight of the decision is meaningful, rather than just a token gesture. It should be carried out by someone who has the authority and competence to change the decision. As part of the analysis, they should consider all the available input and output data." (A29WP, 2018)

The current state of the art (Yu et al., 2018) suggests that, for the moment, fully automated clinical decision systems, which could be understood as solely automated decision systems, are scarce in comparison with integrative decision support systems, where clinicians still need to make the final decision and, therefore, they are invested with authority and competence to change the algorithmic decision. Of course, neither the Regulation, nor the Statement by the Working Party focus specifically on the health care arena (indeed, this last document only mentions health care marginally). At first sight, it is perfectly clear that patients are entitled to claim for the intervention of a human being in the process of decision-making, but although the intervention of a human with authority and capability to change the decision may be legally appropriate and societally desirable, it might present enormous difficulties in practice (Brkan, 2019). This approach does not serve us well to specify what kind of obligations physicians and caregivers who play the role of data controllers will have to assume. Limitations are both relevant for fully automated clinical decision-making systems, where human intervention is claimed, and for decisions based on integrative decision support systems, where human intervention remains in control. Moreover, we have to focus on the kind of practical consequences this system will bring and which dynamics might arise within the health caregiving community.

In our opinion, this is quite difficult to determine, since circumstances can change substantially depending on the time of care – diagnosis or therapy – or even depending on the circumstances of each specific case. To begin with, it is obvious that a physician must, in any case, supervise that the diagnosis or treatment recommendation provided by the AI does not blatantly contradict what medical science has been able to determine in a well-known situation. Thus, for example, if the AI recommends a treatment or a dose of medicine that would surely cause unnecessary harm or even death to a patient, physicians must be able to detect it and impose their judgement on the machine (and report the failure to improve the system, obviously).

Much more complex is determining what to do in cases where a machine recommendation challenges what we might call its intuitions. In these circumstances, we face a complex dilemma. On the one hand, if we concede that it has to be the medical criterion, we would be largely denying one of the bases that justify the use of AI: its ability to make a diagnosis or recommend treatment more efficiently than a human in unclear circumstances. On the other hand, it seems complex to force physicians to act against their own inclinations. Another relevant point in the evaluations of those dynamics is how automation may reinforce the (mal)practice of 'defensive medicine' (Perin, 2019). Surely, the solutions to these dilemmas can only be traced by leaving the final decision in the hands of a patient who has been adequately informed of the circumstances at hand. This will include, of course, the possible consequences of an error in the suggestion of the machine or in human intuition. Once again, it seems necessary to resort to some kind of advice that goes beyond that which can be provided by the doctor who is directly involved in the dilemma. And once again it seems recommendable to introduce the figure of the Health Information Counsellor in the equation.

References

Afzal Hussain Shahid, A. H. & Singh, M. P. (2019). Computational intelligence techniques for medical diagnosis and prognosis: Problems and current developments. Biocybernetics and Biomedical Engineering (in press). https://doi.org/10.1016/j.bbe.2019.05.010

A29WP WP251. Guidelines on Automated individual decision-making and Profiling for the purposes of Regulation 2016/679. European Commission, 22 August, 2018.

Beer, D. (2017). The social power of algorithms. Information Communication and Society, 20(1), 1-13.

Bini, S.A. (2018). Artificial Intelligence, Machine Learning, Deep Learning, and Cognitive Computing: What Do These Terms Mean and How Will They Impact Health Care? The Journal of Arthroplasty, 33, 2358-2361.

Buolamwini, J., & Gebru, T. (2018). Gender Shades: Intersectional Accuracy Disparities in Commercial Gender Classification∗ Joy. Proceedings of Machine Learning Research, 81, 1–15.

Burrell, J. (2016). How the machine 'thinks': Understanding opacity in machine learning algorithms. Big Data & Society, 3(1), 1–12.

Char, D. S., Shah, N. H., & Magnus, D. (2018). Implementing Machine Learning in Health Care: Addressing Ethical Challenges. New England Journal of Medicine, 378, 981–983.

Coeira, E., (2019). On algorithms, machines, and medicine. The Lancet Oncology, 20(2), 166–167.

Cotino, L. (2017). Big Data and Artificial Intelligence. An Approach from a Legal Point of View about Fundamental Rights. Dilemata, 24, 131–150.

de Miguel Beriain, I. (2018). Does the use of risk assessments in sentences respect the right to due process? A critical analysis of the Wisconsin v. Loomis ruling. Law, Probability and Risk, 17, 45–53.

Doshi-Velez, F., Kortz, M., Budish, R., Bavitz, C., Gershman, S., O'Brien, … Wood, A. (2017). Accountability of AI under the law: The role of explanation. arXiv preprint arXiv:1711.01134, 2017.

Ellis, G. & J. Silk (2014). Defend the integrity of physics, Nature, 516, 321-323.

Gilpin, L.H., Bau, D., Yuan, B.Z., Bajwa, A., Specter, M.A., & Kagal, L. (2018). Explaining Explanations: An Overview of Interpretability of Machine Learning. 2018 IEEE 5th International Conference on Data Science and Advanced Analytics (DSAA), 80-89.

Ferretti, A., Schneider, M., & Blasimme, A. (2018). Machine Learning in Medicine: European Data Protection. Law Review, 4, 320 – 332.

Fiske, A., Buyx, A., & Prainsack, B. (2018). Health Information Counselors: A New Profession for the Age of Big Data. Academic Medicine : Journal of the Association of American Medical Colleges, 94, 37–41.

Frisse, M.E. (2016). The business of trust. Academic Medicine, 91: 462–464.

Goodman, B. & Flaxman, S. (2017). European Union Regulations on Algorithmic DecisionMaking and a "Right to Explanation". AI Magazine, 38, 50–57.

Hamet, P. & Tremblay, J.. (2017) Artificial intelligence in medicine. Metabolism, 69, S36–S40.

Hoffmann, A. L. (2019). Where fairness fails: data, algorithms, and the limits of antidiscrimination discourse. Information, Communication & Society, 22(7), 900–915.

Kallianos, K., Mongan J., Antani, S., Henry, T., Taylor, A., J.Abuya, J.& Kohli, M. (2019). How far have we come? Artificial intelligence for chest radiograph interpretation. Clinical Radiology, 74, 338-345.

Kemper, J., & Kolkman, D. (2018). Transparent to whom? No algorithmic accountability without a critical audience. Information, Communication & Society, Published online: 18 Jun 2018, 1–16. https://doi.org/10.1080/1369118X.2018.1477967

Le, E. P. V., Wang, Y., Huang, Y., Hickman S. & Gilbert F. J. (2019). Artificial intelligence in breast imaging. Clinical Radiology, 74, 357-366.

Lipton, Z. C. (2018). The Mythos of Model Interpretability. Queue, 16(3), 31–57. https://doi.org/10.1145/3236386.3241340

Londhe, V. Y. & Bhasin, B. (2019) Artificial intelligence and its potential in oncology. Drug Discovery Today, 24, 1228-232.

Mazzocchi, F. (2015). Could Big Data be the end of theory in science? EMBO Reports, 16(10), 1250–1255.

Minssen, T., & Pierce, J. (2018). Big Data and Intellectual Property Rights in the Health and Life Sciences. In I. Cohen, H. Lynch, E. Vayena, & U. Gasser (Eds.), Big

Data, Health Law, and Bioethics (pp. 311-323). Cambridge: Cambridge University Press.

Ngiam, K. Y. & Khor, I. W. (2019). Big data and machine learning algorithms for health-care delivery. Lancet Oncology, 20: e262–73.

Niazi, M. K. K., Parwani, A. V. & Gurcan, M. N. (2019) Digital pathology and artificial intelligence. Lancet Oncol, 20: e253–61.

Noorbakhsh-Sabet, N., Zand, R., Zhang, Y. & Abedi, V. (2019). Artificial Intelligence Transforms the Future of Health Care. The American Journal of Medicine (in press). https://doi.org/10.1016/j.amjmed.2019.01.017

Nuñez-Reiz, A. (2019). Big data and machine learning in critical care: Opportunities for collaborative research. Medicina Intensiva, 43, 52-57.

O'neil, C. (2016). Weapons of Math Destruction: How Big Data Increases Inequality and Threatens Democracy. (Random House LCC US, Ed.).

Perin, A. (2019). Standardizzazione, automazione e responsabilità medica. Dalle recenti riforme alla definizione di un modello d'imputazione solidaristico e liberale. Biolaw Journal - Rivista Di BioDiritto, (1), 207–235.

Schmidt-Erfurth, U., Sadeghipour, A., Bianca S. Gerendas, B. S., Waldstein, S. M. & Bogunović H. (2018) Artificial intelligence in retina. Progress in Retinal and Eye Research, 67, 1-29.

Scott, I. A. (2018) Machine Learning and Evidence-Based Medicine. Annals of Internal Medicine, 169, 44-46.

Selbst, A. D. & Powles, J. (2017). Meaningful Information and the Right to Explanation. International Data Privacy Law, 7, 233–242.

Sterky, F. & Lundeberg, J. (2000). Sequence analysis of genes and genomes. Journal of Biotechnology 76, 1-31.

U.S Food & Drug Administration. Proposed Regulatory Framework for Modifications to Artificial Intelligence/Machine Learning (AI/ML)-Based Software as a Medical Device (SaMD) - Discussion Paper and Request for Feedback. Published online: April 2, 2019. Retrieved from: https://www.fda.gov/media/122535/download

Wachter, S., Mittelstadt, B., & Floridi, L. (2017). Why a Right to Explanation of Automated Decision-Making Does Not Exist in the General Data Protection Regulation. International Data Privacy Law, 7, 76–99.

Wartman, S. A., & Combs, C. D. (2019). Reimagining Medical Education in the Age of AI. AMA Journal of Ethics, 21(2), 146–152.

Wellner, G., & Rothman, T. (2019). Feminist AI: Can We Expect Our AI Systems to Become Feminist? Philosophy & Technology. https://doi.org/10.1007/s13347-019-00352-z

Yu, K.-H., Beam, A. L. & Kohane, I. S. (2018). Artificial intelligence in healthcare. Nature Biomedical Engineering, 2, 719–731.

Zarsky, T. (2017). Incompatible: The GDPR in the Age of Big Data (2017) Seton Hall Law Review, 47, 995-1019.

Zerilli, J., Knott, A., Maclaurin, J., & Gavaghan, C. (2018). Transparency in Algorithmic and Human Decision-Making: Is There a Double Standard? Philosophy & Technology. https://doi.org/10.1007/s13347-018-0330-6

This is a post-peer-review, pre-copyedit version of an article published in Medicine, Health Care and Philosophy (2020) 23:247-252.
The final authenticated version is available online at: http://dx.doi.org/10.1007/s11019-020-09939-2

Should we have a right to refuse diagnostics and treatment planning by artificial intelligence?

Abstract

Should we be allowed to refuse any involvement of artificial intelligence (AI) technology in diagnosis and treatment planning? This is the relevant question posed by Ploug and Holm in a recent article in *Medicine, Health Care and Philosophy*. In this article, I adhere to their conclusions, but not necessarily to the rationale that supports them. First, I argue that the idea that we should recognize this right on the basis of a rational interest defence is not plausible, unless we are willing to judge each patient's ideology or religion. Instead, I consider that the right must be recognized by virtue of values such as social pluralism or individual autonomy. Second, I point out that the scope of such a right should be limited at least under three circumstances: 1) if it is against a physician's obligation to not cause unnecessary harm to a patient or to not provide futile treatment, 2) in cases where the costs of implementing this right are too high, or 3) if recognizing the right would deprive other patients of their own rights to adequate health care.

Keywords: Artificial Intelligence, right to refuse treatment, health care, patients autonomy

1. Introduction

In July 2019, *Medicine, Health Care and Philosophy* published an extraordinarily interesting article. Thomas Ploug and Søren Holm (2019) argued the need to protect the right to refuse diagnostics and treatment planning by artificial intelligence (AI). Nevertheless, the authors showed the possibility of distinguishing between a strong version of this right, which would allow the holder to refuse any involvement of AI technology in diagnosis and treatment planning, and a weak version, which would only allow recognition of the claim to physician involvement in the diagnostic and treatment planning process. The authors seemed to favour the strong version of the right, albeit with

limitations, when patients' objections are 'based on rational concerns about the systemic effects of AI use'.

In this article, I adhere to their conclusions, but not necessarily to the rationale that supports them. Instead, I criticize some of the weaknesses I found in the authors' arguments and provide some alternative arguments that might serve better to support their proposals. To this purpose, I will start by stating that it is not necessary to introduce a discussion on the weak version of the right, but on its extension. As Ploug and Holm correctly state, the right as such has been clearly recognized by the current European Union (EU) legal framework, even though we are yet to define its boundaries.

Instead, I will focus on the idea that we must adopt the strong version of the right to refuse diagnostics and treatment planning by AI, but subject to severe restrictions. To that end, I will separate myself substantially from Ploug and Holm's argumentation. First, I will argue that the idea that we cannot root this right based on a rational interest defence. I will show that this is not plausible, unless we are willing to judge each patient's ideology or religion and this is against fundamental principles included both in the Charter of Fundamental Rights of the European Union and in the General Data Protection Regulation. Instead, I will argue that the proposed right must be connected with values such as social diversity or individual autonomy and responsibility. Afterward, I will point out that that the scope of such a right should be limited at least under three circumstances: 1) if it is against a physician's obligation to not cause unnecessary harm to a patient or to not provide futile treatment, 2) in cases where the costs of implementing this right are too high, or 3) if recognizing the right would deprive other patients of their rights to adequate health care.

2. The current EU legal framework: A weak version of the right to refuse diagnostics and treatment planning by AI in the GDPR

The EU legal framework on the application of AI to human health is resolved by Article 22(1) of Regulation (EU) 2016/679 of the European Parliament and of the Council of 27 April 2016 on the protection of natural persons with regard to the processing of personal data and on the free movement of such data, and repealing Directive 95/46/EC (General Data Protection Regulation, thereinafter GDPR), which reads *'The data subject shall have the right not to be subject to a decision based solely on automated processing,*

including profiling, which produces legal effects concerning him or her or similarly significantly affects him or her.'

Obviously, the wording of this clause clearly indicates that it is not possible to use AI if there is no human element involved in the decision-making process. Therefore, we can definitely hold that the weak version of the right invoked by Ploug and Holm has already been recognised by the EU law. However, they are perfectly right when they point out that its scope is yet to be defined. Indeed, the article does not make explicit is the degree of intervention that must be considered necessary to conclude that the requirement is covered (Mitchell and Ploem 2018). Therefore, it is particularly important to highlight the statement made by the Article 29 Data Protection Working Party, an advisory body comprising a representative from the data protection authority of each EU member state, which played a prominent role in terms of interpretation of the Regulation until it was replaced by the European Data Protection Board (EDPB) under the GDPR. In 2017, the Party clarified the scope of the prohibition by stating that:

> *'The controller cannot avoid the Article 22 provisions by fabricating human involvement. For example, if someone routinely applies automatically generated profiles to individuals without any actual influence on the result, this would still be a decision based solely on automated processing. To qualify as human intervention, the controller must ensure that any oversight of the decision is meaningful, rather than just a token gesture. It should be carried out by someone who has the authority and competence to change the decision. As part of the analysis, they should consider all the available input and output data.'*

Therefore, the legal framework is clear in one part: there is no room for solely automated decision-making, at least in the EU zone (Dreyer and Shulz 2019). However, the degree of concrete involvement of physicians in the final decision and the tools implemented to guarantee that this not become a mere tramite remains unclear. It is hard to see at the present moment how could we avoid that physicians routinely adopt the recommendations made by artificial intelligence due to defensive medicine considerations, for instance. The implementation of the weak version of the right and its concretisation includes a wide range of options and we have not faced this issue yet. Thus, there is an urgent need for discussion on this essential point and Ploug and Hold are perfectly right when they claim for it, since social concerns must play a fundamental role

in the decisions made. Hopefully, this will bring us a better agreement on the concrete degree of human involvement that the weak version of the right involves and the best ways to guarantee it.

Anyway, we could at least conclude that at the present moment, the weak version of the right to refuse diagnostics and treatment planning by AI, that is, the 'claim to physician involvement in the diagnostic and treatment planning process', has been endorsed by the EU regulation (Wachter et al. 2017). However, this does not at all mean that a strong version of the right, that is, the right 'to refuse any involvement of AI technology in diagnosis and treatment planning', is against the EU regulation. Indeed, in the next section I will argue that such a strong version of the right works well with some of the values that are widely accepted in the EU context, and thus there are some good reasons to support it.

3. The argument for the recognition of the strong version of the right

One of the parts I found most disturbing in the article by Ploug and Holm is that in the section entitled 'Rational concerns and dystopies' they defend the idea that the strong version of the right must be based on the patient's rational fears and concerns. In fact, the authors make a great effort to demonstrate that if a patient raises an objection to the use of AI for those purposes on the grounds of a possible undesirable societal effect, then we should respect the patient's claim and recognize their right to refuse diagnostics and treatment planning by AI. In my opinion, this is an unfortunate argumentation, as it concedes, in the negative sense that, if there is no rational explanation of the reasons for refusing AI, then the strong version of the right does not apply. This implies assuming the need to situate our focus on the reasonableness of a request, which means questioning the rationality of an ideology or a faith, an attitude that violates Article 21 of the Charter of Fundamental Rights of the European Union (CFR). Indeed, this is not a typical course of action, of course. Take Jehovah's Witnesses case, for example. Do we really protect their right to refuse certain treatments on the basis of the rationality of their beliefs? In the answer to this question is the reason for my rejection of the theses of the authors of the article I now criticize (Petrini 2014).

Of course, I am not holding here that the right to refuse diagnostics and treatment planning by artificial intelligence must be considered as a part of the general right to

refuse treatment. I think that Ploug and Hold argue in a very convincing manner that both rights are different. Instead, I hold that the principles that refrain us for judging the ideology or religious beliefs that support refusing a treatment should also apply to the right that the authors of the paper that I am commenting are describing. Furthermore, I consider that there are no good reasons to oblige patients to declare the reasons why they are opposing to the use of AI in the decision process, provided that the conditions I mentioned in the introduction and will explore in the following sections apply. If this is the case, then it would only be the patient who would suffer the consequences of his or her negative. Therefore, I cannot see any strong reason to oblige him or her to reveal any kind of information about his or her ideology. Moreover, that would be contrary to the principle of data minimization, an essential ethical principle that has been incorporated to the European Union's General Data Protection Regulation (GDPR). This principle means that data processing should only use as much data as is required to successfully accomplish a given task. Provided that we can base the right refuse diagnostics and treatment planning by artificial intelligence on reasons other than the rationality of a belief (as I will hold immediately), I do not think that we have any reason to oblige patients to reveal these very sensitive personal data. Instead, if the conditions mentioned apply, then the reasons that guide the patients' decision would be totally irrelevant, since the right would not be applicable.

Rather, I believe that we must opt for the strong version of the right based on value pluralism and the patient's autonomy and responsibility. Value pluralism means that 'people's views diverge about a range of fundamental questions, political ethical and religious. This diversity appears to be inevitable and irresolvable. It is not possible to determine a single correct view or set of values (Turner 2004). As a consequence, negotiation, tolerance and compromise are necessary' (Wilkinson and Savulescu, 2018). Indeed, this value has been embedded in the EU Chart of Fundamental Rights in its Preamble ('The Union contributes to the preservation and to the development of these common values while respecting the diversity of the cultures and traditions of the peoples of Europe') and plays a key role in EU societies at present.

My second argument relies on the concept of autonomy, a concept that is certainly mentioned in the paper by Ploug and Hold, but in a quite different sense. They place autonomy in the basis of our acting as rational beings. In my view, autonomy refers here to the capacity of the patients to make their own decisions according to their principles

and values Indeed, I think that respect of patient autonomy is guaranteed under Article 8 of the European Convention on Human Rights, which protects the right to private and family life. Thus, it serves as an excellent root to the right to refuse AI intervention in health care. This is probably due to the fact that I think that autonomy must be understood not only as a right to refuse a treatment, but to make decisions on the whole treatment process, as autonomy is rooted in the importance of self-government and freedom to live according to one's goals (Varelius 2006; Hartzband and Groopman 2009).

Therefore, I believe that the strong version of the right we are considering is directly connected with basic values such as patient autonomy and value pluralism and therefore it must be fully accepted in the EU context. Indeed, the focus should be on the reasons we could oppose or at least request the restrictive use of a right that is directly linked to these fundamental principles and values. What could be the reasons for defining boundaries to the right to refuse diagnostics and treatment planning by AI in its strong version? In my opinion, there are two: the need to reconcile this right with respect to physicians' ethical concerns and the costs it might involve for health care systems. I will analyse both in the next sections.

4. The argument of physicians' right to make an informed decision

First, one might oppose the right we are considering by stating that physicians are meant to have a say in the diagnostic and treatment procedures used in the development of their work. However, if we do not adopt a paternalistic approach to medicine, I doubt that this statement involves a general right for physicians to make decisions without considering the patient's values and interests. For example, in the case of Jehovah's Witnesses, we concede to such patients the right to decide on how surgery should be performed, not only the right to decide whether or not they want to undergo it. Thus, it does not seem reasonable to consider that patients cannot decide on the diagnostic and treatment tools and the possibility of avoiding AI for these purposes. However, I think that this general right only applies if this does not yield as a consequence a violation of the physician's right to not act against globally recognized medical ethical principles, such as non-maleficence or beneficence, for example (Macklin 2003). And this might certainly happen under some circumstances if we recognize a strong version of the right.

Indeed, there are cases when a physician might be unsure on whether a concrete treatment may be futile or even harmful to a patient. Imagine, for example, that a patient with cancer requests chemotherapy, but the oncologist does not know if this could be really effective in this concrete case. Under these circumstances, AI might be the only means of making a decision about it. If the patient exercises the strong version of the right, physicians would have to face a situation in which they might infringe their ethical duties: they might finally act without knowing if the intervention will not cause harm or death, not to mention the futile use of public resources, even though it would be possible to solve this dilemma using AI tools. In my opinion, cases such this show that the right to veto could undermine physicians' right to use the most accurate resources available to ensure that they are not disregarding the essential ethical principles in health care I have mentioned earlier. Indeed, physician refusal to provide futile or harmful care is supported by the ethical principle of non-maleficence, which seems particularly relevant in the situation described (Luce 1995).

Ploug and Holm might point out that this same happens in the case of Jehovah's Witnesses and yet physicians have a duty to proceed with the alternative treatment, but I think that both scenarios are not at all the same. In the case of Jehovah's Witnesses, physicians are forced to choose an alternative treatment that, in any case, would always work better than no treatment at all or death. In the case described, it might perfectly happen that the treatment would help the patient face their cancer, but it might also cause unnecessary pain to the patient. Thus, physicians forced to provide treatment without using AI would be aware that they could be causing an avoidable harm, but they would also know that if they were to not provide the treatment, they could flout the beneficence principle. The question of certitude is key in such situations, but this is precisely what the patient opting for the right would be stealing from physicians, and I think this is unfair to physicians. Moreover, I think this is a misunderstanding of the informed consent framework (Paris 2010).

Therefore, we must conclude that the right of patients to not use AI in decisions about their treatment cannot be extended to the point of forcing doctors to act against commonly accepted medical ethical principles. This could be expressed either by establishing this circumstance as a limit on the exercising of the right, or by accepting the right to conscientious objection from the health professional who is to provide the treatment. In my opinion, it seems more reasonable to adopt the first option, because if

the principle of non-maleficence is a basic principle in medical ethics, we should not think that its respect implies the need to invoke the right to conscientious objection. Furthermore, as the principles at stake are universally accepted, would it make any sense to finally put into practice a treatment that could be futile or harmful only because the patient manages to find a doctor who does not mind carrying it out? In my opinion, such a physician would be flouting the principle of non-maleficence on the basis of the alibi provided by the principle of patient autonomy.

To sum up, I consider that respect of health care workers' principles and values is a strong enough reason to conclude that the right to refuse diagnostics and treatment planning by AI cannot be an unlimited right. Indeed, it seems reasonable to think that if AI can determine whether a treatment will serve, or instead cause harm or be futile, it is the obligation of a physician to make good use of it, due to the prevalence of the non-maleficence principle, which overrides the autonomy of the patient. I concede that Ploug and Holm might be right on the idea that in the future some physicians "may come to be biased toward the decisions made by the IA technology and less sensitive to the particular preference and interests of the individual". However, if this were the case they would not be practising medicine according to the goals and standards of their profession and, thus, this type of attitudes could never become a legitimate boundary to the right we are discussing now.

5. Health care system sustainability and the rights of other patients

The second factor when considering the limits of the right to refuse diagnostics and treatment planning by AI are those that derive from the costs that the recognition of this right could cause to the health care system. Ploug and Holm address this issue in a senseful way in their paper and I cannot but adhere to what they state, even though I would like to make some remarks to it. I do not share their belief that in some circumstances, "allowing some patients to refuse AI involvement (…) might lead to cost savings because patients who are strongly opposed to AI would avoid seeking health care until their conditions have progressed to a serious state", I dare say that, in those cases, the use of AI in a previous state could lead to a better diagnosis and a more effective treatment, rendering the costly treatment that we usually have to administrate to a serious state unnecessary. To the contrary, I suspect that recognizing the strong version of the

right will probably reduce the savings that the implementation of AI in health care might bring. However, I do not think that this fact, even if confirmed by evidence, should play a definitive role in order to oppose to the recognition of the right. It is very common, in fact, for the exercise of a right derived from patient autonomy to harm public health. This happens, for example, if we accept that a patient rejects an optimal treatment, opting instead for another that will ultimately would lead to higher public health costs.

However, we usually accept this result based on the defence of principles such as the need to respect the plurality of values in non-uniform societies or the importance of respecting each person's life plans. Thus, for example, no patient is forced to undergo a kidney transplant even though the alternative (that is, long-term dialysis) is much more expensive for the system. Nor has compulsory vaccination been introduced against influenza, even though this substantially increases healthcare costs. Nor, of course, are patients penalized in general for not strictly following the recommended treatment, even though this may lead to relapses and higher costs. Moreover, there are strong reasons that support such policies (Howard 2008; Schmidt 2007). I therefore understand that the concept of increased public health costs should not serve to veto, in general, the strongest version of the law we are analysing.

However, I believe that there are exceptions to this general rule. If a treatment is particularly costly, for example, it should not be administered without first having recourse to the advice of the AI if the efficiency of the corresponding predictive algorithm had been demonstrated. This usually happens in health systems, which set specific indexes for decision-making on financing treatments (which as in the UK happens with the incremental cost-effectiveness ratio [ICER]) (Nikolentzos et al 2008). I believe that this type of threshold could perfectly well be applied even in the case of recognising this right, setting objective limits to its exercise.

However, I believe that the main objection against the strong version of the right we are analysing comes from other types of situations. More specifically, it is necessary to consider cases in which the exercising of this right would cause obvious harm to third parties, who would be deprived of adequate care as a consequence. This might be better understood through an example. Imagine that eight people want to gain access to very expensive treatment. Furthermore, the statistics show that only half of the patients with that concrete condition respond to the treatment in a minimally reasonable way, following the prevailing criteria for allocating resources in that health system. Interestingly, there is

an algorithm capable of precisely guessing which of these eight people will benefit from the treatment at a reasonable cost and which will not. However, four of them refuse to have the AI used in the analysis of their specific case.

Imagine now that the AI is used for the other four and that the algorithm determines that two of them are not treatable under the underlying cost conditions. This means that there are six people left who are likely to enter into the final selection of the four candidates. If we believe that the appeal to the right should not lead to any discrimination against those who exercise it, it would be logical to draw lots among the remaining six. Thus, fortune would decide impartially who will and will not be treated. However, statistically, this would imply that at least one person capable of healing would be excluded and one for whom treatment is futile would be treated.

In my view, however, this final distribution of resources would be absurd. The logical approach would be to administer treatment to the two people for whom the AI has made an encouraging prognosis and to circumvent the other two candidates among the four who want to exercise their right to not have these mechanisms used to decide on their treatment. The opposite would be to arrive at an inefficient and unfair result based on personal ideology. However, if this is the case, then it is clear that the right we are talking about has to be limited based on the costs for the health system, the need to optimise resouce allocation, but above all, on the right of third parties to access efficient treatment. It could, of course, be pointed out that the case I have put forward is exceptional and should not serve as a rule. I do not think that is true. It is a case that arises every time there is a drug shortage, and we must design a system for allocating scarce resources among patients who are likely to take advantage of it (or not). In my opinion, if AI were able to suggest an efficient form of allocation, we should not allow the right to refuse treatment planning by AI to deny scarce health resources to patients who are able to benefit from it.

References

1. Ploug T, Holm S. The right to refuse diagnostics and treatment planning by artificial intelligence. Medicine, Health Care and Philosophy, 2019. Jul 29. doi: 10.1007/s11019-019-09912-8. [Epub ahead of print]

2. Mitchell C, Ploem C. Legal challenges for the implementation of advanced clinical digital decision support systems in Europe. J Clin Transl Res. 2018;3(Suppl 3):424–430. Published 2018 Aug 18.

3. Dreyer S, Schulz W, The General Data Protection Regulation and Automated Decision-making: Will it deliver? Potentials and limitations in ensuring the rights and freedoms of individuals, groups and society as a whole. Working Paper. January 2019 Bertelsmann Stiftung. Available from: https://www.bertelsmann-stiftung.de/fileadmin/files/BSt/Publikationen/GrauePublikationen/GDPR.pdf. Accessed 11 August 2019

4. Turner L, Bioethics in pluralistic societies. Medicine, Health Care and Philosophy, 2004: 7(2), 201–208.

5 Art. 22 GDPR. 2018. Automated Individual Decision-Making, Including Profiling|General Data Protection Regulation (GDPR). General Data Protection Regulation (GDPR). https://gdpr-info.eu/art-22-gdpr/. Accessed 11 August 2019

6 Article 29 Data Protection Working Party. 2018. ARTICLE29 Guidelines on Automated Individual Decision-Making and Profiling for the Purposes of Regulation 2016/679 (wp251rev.01). https://ec.europa.eu/newsroom/article29/item-detail.cfm?item_id=612053. Accessed 20 July 2019

7. Wachter S, Mittelstadt B, and Floridi L, Why a Right to Explanation of Automated Decision-Making Does Not Exist in the General Data Protection Regulation. International Data Privacy Law, 2017, 7(2):76–99.

8. Charter of Fundamental Rights of the European Union (CFR), 2012/C 326/02. https://eur-lex.europa.eu/legal-content/EN/TXT/?uri=CELEX:12012P/TXT. Accessed 20 July 2019

9. Petrini C. Ethical and legal aspects of refusal of blood transfusions by Jehovah's Witnesses, with particular reference to Italy. Blood Transfus. 2014;12 Suppl 1(Suppl 1):s395–s401. doi:10.2450/2013.0017-13

10. Wilkinson D, Savulescu J, Cost-equivalence and Pluralism in Publicly-funded Health-care Systems. Health Care Analysis, 2018, Volume 26, Number 4, Page 287

11. Howard BM, First, Do Not Punish: Individual Incentives in Health Policy, Virtual Mentor. 2008;10(11):719-723. doi: 10.1001/virtualmentor.2008.10.11.conl1-0811.

12. Schmidt H. Patients' charters and health responsibilities. BMJ. 2007;335(7631):1188

13 European Convention on Human Rights. https://www.echr.coe.int/Documents/Convention_ENG.pdf. Accessed 20 August 2019

14. Varelius J. The value of autonomy in medical ethics. Med Health Care Philos. 2006;9(3):377-788.

15. Macklin R. Applying the four principles. Journal of Medical Ethics 2003;29:275-280

16 Hartzband P, Groopman J. Keeping the patient in the equation—Humanism and health care reform. New England Journal of Medicine 2009, 361: 554–555

17. Luce JM, Physicians do not have a responsibility to provide futile or unreasonable care if a patient or family insists. Crit Care Med. 1995;23:760–766

18 Paris JJ, Autonomy does not confer sovereignty on the patient: A commentary on the Golubchuk case. American Journal of Bioethics, 2010: 10(3), 54–56

19. Nikolentzos A, Nolte E, Mays N, Paying for (expensive) drugs in the statutory system: an overview of experiences in 13 countries. London: London School of Hygiene & Tropical Medicine. 2008. Available from: http://www.dh.gov.uk/en/Publicationsandstatistics/Publications/PublicationsPolicyAndGuidance/DH_ 089990. Accessed 20 August 2019.

Acknowledgements:

This project has received funding from the European Union's Horizon 2020 research and innovation programme under grant agreement No 788039

Journal of Law and the Biosciences, 1–19
doi:10.1093/jlb/lsaa040
Original Article

‘Digital pills’ for mental diseases: an ethical and social analysis of the issues behind the concept

Iñigo de Miguel Beriain[1,*] and Marina Morla González[2]

[1]Derecho Público. University of the Basque Country, Leioa, Spain, and IKERBASQUE, Basque Foundation for Science. Bilbao. Spain
[2]Departamento de Derecho Público, Universidad de Leon, Area de Derecho Eclesiástico, Spain
*Corresponding author. E-mail: inigo.demiguelb@ehu.eus

ABSTRACT

Recently, the US Food and Drug Administration has given a landmark approval to the very first digital pill with a sensor embedded in the inside. These are complex systems that include a drug and an electronic tracker that is activated when the patient takes the pill. Accordingly, they might be an excellent tool for monitoring and potentially improving patients’ adherence to prescriptions. This would serve well to avoid unnecessary healthcare costs and reduce the anxiety of patients and their relatives. However, digital pills might also diminish patient autonomy, reduce privacy, or promote inadequate use of pharmaceutical resources. This article is aimed at contributing to adequate use of this new tool by showing the main ethical and social issues they involve and proposing measures meant to address them. Finally, we conclude by defending the idea that these new systems should be seen as means of complementing traditional strategies to promote adherence to treatment, and not as substitutes.

KEYWORDS: Digital health, digital pills, adherence, ingestible sensor, data protection and privacy, patient autonomy

INTRODUCTION: A GAME-CHANGER TECHNOLOGY IS BORN

On November 13, 2017, a pharmaceutical company Otsuka Pharmaceutical Co., Ltd (Otsuka), based in Maryland, USA, and a Silicon Valley company, Proteus Digital Health (Proteus), announced that the US Food and Drug Administration (FDA) had approved a digital medicine system called Abilify MyCite® (AMC, aripiprazole tablets

with a sensor). This is a drug–device combination product comprised of Otsuka's oral aripiprazole tablets embedded with an ingestible event marker (IEM) sensor.[1] Aripiprazole is an antipsychotic used to treat adults with schizophrenia, bipolar I disorder, and major depressive disorder. The drug is part of a more complex product, i.e. the Abilify MyCite® System, which comprises the Abilify MyCite® and the following components: a wearable sensor developed by Proteus, i.e. the MyCite® Patch; a smartphone application (app) called MyCite® app, which can display information about the patient on a compatible smartphone, and web-based portals for healthcare providers and caregivers that display a summary of aripiprazole ingestion over time.[2]

The Abilify MyCite® System offers healthcare providers an astonishing outcome: it records real-time medication ingestion by patients and collects data on activity level as well as self-reported rest and mood. The processing is easily described: after the daily antipsychotic pill is swallowed, a digital sensor the size of a grain of sand (made of copper, magnesium, and silicon, which Proteus states are all found in food) functions like a battery by releasing an electric signal to the patch when it has reached the stomach acid. Thus, the adhesive patch on the patient's torso collects information on the date and time the pill was taken, blood pressure, temperature, and level of activity. Then, the patch sends a signal to an app on the patient's smartphone. At this stage, patients can add self-reported mental health data about how they are feeling. The app uploads the data to a secure website on a cloud-based system for viewing by doctors.[3] As a final result, all information gathered by the system can be communicated to patients and healthcare providers through the electronic devices incorporated with the product. In this manner, it is possible to obtain an objective summary of drug ingestion over time.[4] It is good to highlight that patients can decide who has access to their data at any moment among other authorized parties, such as Otsuka and its vendors, their selected healthcare providers, their family and friends, their pharmacy, or their health plan.[5]

At present, there are good reasons to believe that Abilify MyCite® will soon be followed by other digital pills. Based on the information gathered, the industry is producing apps for substance abuse treatment, diabetes management, and heart and blood pressure monitoring at a rapid clip. At the same time, studies are underway for digital pills for addressing other mental health pathologies, cancer, cardiovascular conditions, and infectious diseases, such as preexposure prophylaxis medications for preventing human immunodeficiency virus.[6] Therefore, a new generation of intelligent

1 PR N. Otsuka and Proteus® Announce the First U.S. FDA Approval of a Digital Medicine System: ABILIFY MYCITE® (aripiprazole tablets with sensor). PR Newswire US [serial online]. Nov. 14, 2017: Regional Business News (accessed Jan. 21, 2020).

2 *Id*.

3 Christopher Rowland, *This $1,650 Pill Will Tell Your Doctors Whether You Have Haken It. Is It the Future of Medicine?* The Washington Post, Apr. 9, 2019, https://www.msn.com/en-ie/news/indepth/this-dollar1650-pill-will-tell-your-doctors-whether-youve-taken-it-is-it-the-future-of-medicine/ar-BBWorND?li=BBPCQrg (accessed Jan. 29, 2020).

4 Anthony Ryan Hatch, *Digital Mental Health Drug Raises Troubling Questions*, PhillyVoice.com (June 15, 2018), https://www.phillyvoice.com/digital-mental-health-drug-cyborg-ethics-abilify-mycite/ (accessed Jan. 21, 2020).

5 Otsuka, Ablify MyCite system Terms of Use, Privacy Notice, and Authorization & Consent. Patient authorization & consent, https://www.otsuka-us.com/media/static/Abilify-Mycite-Patient-Consent.pdf (accessed Apr. 30, 2020).

6 *Supra*, note 3.

drugs is arriving and we must properly address the benefits and challenges posed by them, while preserving our most valuable ethical principles.

Indeed, digital pills might be an excellent tool for monitoring and potentially improving patients' adherence to prescriptions, which could result in an impressive mechanism for avoiding unnecessary healthcare costs and an efficient and excellent tool for reducing anxiety in patients and their relatives. However, they could also diminish patient autonomy or reduce their privacy. Keeping this in mind, this article is aimed at contributing to the adequate use of digital pills by showing the main ethical issues digital pills involve and proposing measures meant to address them. To this end, we start by showing the main benefits digital pills might provide to us all, mainly their potential toward better adherence to treatments. Subsequently, we focus on the main ethical dilemmas this innovation poses, such as marketing pressures that have contributed to the emergence of this cutting-edge product, as well as other challenges facing patient autonomy.

DIGITAL PILLS: THE PEARLS

As mentioned, digital pills might be an extremely useful tool for reliable identification and minimization of medication non-adherence, a crucial issue in terms of healthcare systems governance. The lack of adherence to treatment causes huge dysfunctions in the healthcare sector. In accordance with internationally recognized standards by the medical profession, a patient is observing a treatment if the average ratio between medication intake and prescription is $\geq$80 per cent.[7] The World Health Organization considers that, in the case of chronic diseases, at least 50 per cent of patients show poor adherence to treatment in global terms,[8] a percentage that is even lower in certain cases.[9] In France, for example, a study has shown very low levels of treatment observation: 36 per cent of heart failure cases, 37 per cent of Type 2 diabetes cases, 40 per cent of hypertension cases, 44 per cent of hypercholesterolemia cases, or 53 per cent of osteoporosis cases.[10] These figures are particularly worrying in the case of psychiatric illnesses.[11] Non-adherence causes terrible consequences. Indeed, it causes death or higher complications to a huge number of patients.[12] For example, non-adherence is the largest driver of relapse and hospitalization among patients with disorders such as schizophrenia, diabetes, and asthma.[13]

Furthermore, non-adherence leads to considerable yearly cost overruns. In terms of health from an economic perspective, in the USA, non-optimized medication therapy

7 Grégoire Moutel et al. *Le Medicament Connecté, Entre Bienveillance et Surveillance,* 34 Analyse des enjeux éthiques, médecine/sciences 717–22 (2018).

8 Sabate E. *Adherence to Long-Term Therapies: Evidence for Action.* Report 2003. Geneva: World Health Organisation, n° 92 4 154599 2 (2003).

9 Aurel O. Iuga & Maura J. McGuire, *Adherence and Health Care Costs*, 7 Risk Manag Healthc Policy 35, at 4 (2014).

10 Moutel, *supra* note 7.

11 Palazzo P. *Observance Médicamenteuse Et Rechutes Dans La Schizophrénie : Des Neuroleptiques Classiques Aux APAP,* 167 Annal. Médico-Psychol. 308–17 (2009).

12 P. M. Ho, J. S. Rumsfeld, F. A. Masoudi, et al. *Effect of Medication Non-adherence on Hospitalization and Mortality among Patients with Diabetes Mellitus,* 166 Arch. Int. Med. 1836–1841 (2006). doi:10.1001/archinte.166.17.1836

13 Iuga, A. O. & M. J. McGuire. 2014. *Adherence and Healthcare Costs*, 7 Risk Manag. Healthc. Policy 35–44. doi:10.2147/RMHP.S19801

costs up to $528.4 billion, which is equivalent to 16 per cent of the total US healthcare expenditure in 2016.[14] For European states, non-adherence is estimated to turn into an economic loss of €125,000 million each year.[15] In terms of medical practice, non-adherence constitutes a fundamental obstacle to adequate practice of good care, as 'When patients do not respond to a medication, it can be difficult to determine whether the lack of response is due to non-adherence or whether the medication itself is not effective'.[16]

To date, physicians cannot really do much to solve the adherence issue. They are in general entirely dependent on patients' self-reporting. However, this source is not reliable. Some patients do not report adequately because they are unable to keep good records or they are not willing to do so due to reasons such as failure to understand the instructions, lack of resources, and adverse effects[17].

Therefore, the need to improve adherence to treatment is undoubtedly an essential task for healthcare systems. In recent years, multiple studies have been carried out with the view to achieving this objective. For example, between 2009 and 12, the European Commission financed the ABC research project (Ascertaining Barriers for Compliance: Policies for safe, effective and cost-effective use of medicines in Europe) within its Seventh Framework Program. This and other studies placed the focus of better adherence on the need to strengthen the relationship of trust between patient and doctor, because it is in the direct relationship between the two where it is easiest to assess the actual observance of treatment. Several additional approaches have been developed to support adherence, such as the establishment of therapeutic education groups in healthcare services and patient discussion groups.[18]

Traditionally, healthcare providers could use directly observed therapy (DOT) when needed to ensure that patients adhered to the treatment on schedule. Now, new technologies are turning into a useful tool for physicians to measure adherence with the same objectivity rates, while overcoming some disadvantages shown by DOTs. These alternatives include the issuance of follow-up notebooks to be completed by the patient, which allows the patient to check their catches and omissions, and the doctor to advise the patient during consultations. Alternative tools include electronic medication

14 Jonathan H. Watanabe, Terry McInnis, *Cost of Prescription Drug-Related Morbidity and Mortality,* 52 9 Ann. Pharmacother., 829, at 832 (2018).

15 Unidad de Bioindustrias y Farmacia. *Antares Consulting*, https://www.antares-consulting.com/es_ES/main/detallepublicacion/Publicacion/79/apartado/B/idUnidad/1 (accessed Jan. 16, 2020). See: National Council on Patient Information and Education, Enhancing Prescription Medicine Adherence: A National Action Plan, 7 (2007): Almost half of those polled (49%) said they had forgotten to take a prescribed medicine; nearly one-third (31%) had not filled a prescription they were given; nearly three out of 10 (29%) had stopped taking a medicine before the supply ran out; and almost one-quarter (24%) had taken less than the recommended dosage.

16 C. M. Klugman et al. *The Ethics of Smart Pills and Self-acting Devices: Autonomy, Truth-telling, and Trust at the Dawn of Digital Medicine,* 18 Am. J. Bioethics, 38–47 (2018).

17 M. J. Stirratt, J. Dunbar-Jacob, H. M. Crane et al. *Self-report Measures of Medication Adherence Behavior: Recommendations on Optimal Use,* 5 Trans. Behav. Med. 470–82 (2015). doi:10.1007/s13142-015-0315-2. In the case of schizophrenia, see: Peter M. Haddad, Cecilia Brain & Jan Scott, *Nonadherence with Antipsychotic Medication in Schizophrenia: Challenges and Management Strategies.* Patient Relat. Outcome Meas. 43, at 48 (2014).

18 Moutel, *supra* note 7.

container lids[19] or boxes (called pill boxes, sometimes electronic) containing as many boxes as there are doses to be taken in a day, which the patient can program to trigger alerts on their mobile phone.[20] Many of these tools raise awareness of new and precise information provided by the device on a daily basis, and during consultations, they provide 'feedback' with the professional and allow for dialogue.[21] Besides, these medical–device applications not only help patients play an active role in the decision-making process, but also constitute a means of supplying the lack of time not dedicated by the physician,[22] as they provide continuous monitoring that allows patients and physicians immediate access to the patient's relevant health data.[23] Yet, they all rely on the patient's will to monitor their adherence to the drug prescribed. Thus, some scholars have pointed out the need for better alternatives for measuring adherence.[24]

Are digital pills the response to this query? At first sight, it looks like it. Unlike in the previous scenario, in a world with digital pills, a cooperative attitude on the part of the patient is no longer necessary to obtain accurate knowledge of adherence to treatment. It is enough for patients to agree to use the pills (or for the system to force them to adopt them) so that their physicians know perfectly what the real adherence to treatment is. An additional advantage is that this technology could serve to help patients overcome some of the difficulties they face when trying to follow a treatment, a situation that is particularly stressful in the case of the elderly or people with mental conditions.

Thus, this wirelessly observed therapy offers better features than the supporting technologies already described, which still rely on the patient's capacities and will (what if a patient misuses the notebooks or simply does not take the pill even if they remove it from the box?). Indeed, unlike traditional tools, digital pills register observance automatically, providing patients with the means to ensure optimal monitoring of their drug administration, avoiding missed or duplicated doses.[25]

Nevertheless, it is important to underscore that reasons behind bad adherence rates can be diverse and multiple (not always they consist on a mere distraction to be solved through a tracking system). When we talk about bad adherence to treatments we are addressing a complex biosocial phenomenon, as health sciences and social sciences literature show us. To this regard, if we assume that the operating mode of digital pills could offer a good solution to solve the adherence issue, we should be aware of the professional perspective we are adopting—in which no report from the patient is needed—since from the non-adherent patient perspective, the system could be far from approaching the true reasons behind bad adherence rates.

19 Klugman, *supra* note 16.

20 B.B. Granger, S.C. Locke et al. *The Digital Drag and Drop Pillbox: Design and Feasibility of a Skill-based Education Model to Improve Medication Management,* 32 J. Cardiovasc. Nurs. E14–20 (2017).

21 Moutel, *supra* note 7.

22 Dimitra Petrakakia, Eva Hilbergb, Justin Waringc, *Between Empowerment and Self-discipline: Governing Patients' Conduct through Technological Self-care,* 213 Soc. Sci. Med. 146, at 150 (2018).

23 Ho, *infra* note 27.

24 S. Garfield, S. et al., *Suitability of Measures of Self-reported Medication Adherence for Routine Clinical Use: A Systematic Review*, 11 BMC Med. Res. Methodol. 149 (2011). doi:10.1186/1471-2288-11-149

25 P.R. Chai, Rosen R., Boyer E.W. *Ingestible Biosensors for Real-time Medical Adherence Monitoring: myTMed,* 16 Proc. Annu. Hawaii Int. Conf. Syst. Sci. 416–23 (2016).

To sum up, digital pills provide health systems with precise data on patients' medication taking[26] while informing physicians on whether the failure of a prescribed treatment is due to the ineffectiveness of the treatment or a significant failure in its administration. However, this does not necessarily mean that digital pills must be considered a kind of panacea for adherence issues. Indeed, their use involves relevant issues that should be balanced against their benefits.

THE PERILS: A COPERNICAN TURN IN THE PATIENT–PHYSICIAN RELATIONSHIP

First, one needs to understand that the use of digital pills for monitoring patient adherence constitutes a radical turn in the way we focus this issue. Our current healthcare system is built on a mentality in which trust between clinicians, caregivers, or social workers and the patient is a fundamental piece. The introduction of digital pills replaces this framework with a new policy in which monitoring and control play a key role. It is no longer the patients who reveal data to the physician on a voluntary basis. Instead, the physician becomes a kind of 'Big Brother' who knows everything about the patient even though they are unwilling to share such information.

Of course, one might reply that this does not necessarily have to happen. Indeed, this is hardly the case if the patient is willing to use the digital pill. On the other hand, it is also possible to think that, as the patient will be aware of the knowledge acquired by the physician, it would be much easier for them to discuss the reasons they are not observing their treatment, instead of lying to the healthcare provider. This might indeed happen and it is quite difficult to know in advance whether digital pills might cause a real loss of trust in the physician–patient relationship.

However, the dysfunctions caused by digital pills to the way we approach the functioning of the healthcare system go beyond the loss (or not) of the notion of trust. They extend to the possible erosion of the personal relationship between patients and their physicians. By now, patients usually discuss with their doctor the problems arising from the follow-up of the prescribed treatment. Nevertheless, in the new scenario, patients somehow become the object of inspection of the health system, which watches closely for any deviation from the correct administration of treatment.

It is very important that patients have sufficient confidence in their doctors to discuss with them the reasons they are reluctant to take the prescribed medication. It is also essential that the system provides both with the possibility of building that relationship through adequate means. Thus, with the use of such smart devices, trust would be compromised from both the professional and patient perspectives. First, data generated by the device may cast doubt on the truthfulness of the patient's self-report. Conversely, patients may distrust physicians and their therapeutic recommendations if they receive a different diagnosis from that suggested by the device on which they rely.[27] On the other hand, digital pills open a major gateway to distant and mediated interaction between doctors and patients, thereby decreasing the need for face-to-

26 J. Frias et al. *Effectiveness of Digital Medicines to Improve Clinical Outcomes in Patients with Uncontrolled Hypertension and Type 2 Diabetes: Prospective, Open-label, Cluster-randomized Pilot Clinical Trial,* 19 J. Med. Internet Res. e246 (2017).

27 Ho, A. & Quick, O. *Leaving Patients to Their Own Devices? Smart Technology, Safety and Therapeutic Relationships,* 19 BMC Med. Ethics 18 (2018). https://doi.org/10.1186/s12910-018-0255-8

face communication. Finally, the use of these new technologies may over-technify the monitoring of treatment or decision-making about a patient. This new scenario, which constitutes a serious challenge in the health care arena, is by no means inevitable, but requires the adoption of an appropriate mentality and measures capable of preventing it. It is essential to keep in mind that technologies should serve to enhance the physician–patient relationship, rather than to replace it.[28]. For the sake of maximizing the usefulness of these cutting-edge medical technologies in the way we conceive medicine of even the integral care of patients, we must make a proper use of them in terms of both safety and confidence,[29] otherwise, a key aspect of the patient–physician relationship would be broken: trust. The question, in short, is whether the possible increase in adherence to treatment would compensate for the decrease in this fundamental value, confidence, if patients were forced to use to this new technology.[30] We sincerely believe that this is not the case. That is why we advocate a system that is respectful of patient autonomy and that only allows the use of digital pills in cases in which the patient encourages it, unless the defense of a public good, such as health or safety, makes it essential. We will return to this issue later.

PHARMA BENEFITS VERSUS PATIENT INTERESTS: ETHICAL ISSUES FROM A MARKET PERSPECTIVE

One of the most important ethical dilemmas posed by digital pills comes from the business model on which they are based. Traditionally, the quality of a pharmaceutical product depends on its capacity to improve a patient's health. On this basis, it is possible to draw up cost–benefit analyses, indexes of limitations of coverage in public healthcare, or limits on the provision of funds by insurers. In the case of digital pills, the scenario is much more complicated, as what is offered is not only a medicine, but also a complex pharmaceutical product that combines both that medicine and a monitoring system based on cutting-edge technology. Hence, many challenging dilemmas arise. First, it becomes complicated to compare a system that includes a drug that may not be the most appropriate for a patient with a drug that may be more efficient in treating the patient's specific pathology, but that cannot provide information about adherence. This could obviously be solved by adapting the monitoring system such that it can be incorporated into any medicine, but for the moment this scenario is far from reality.

Furthermore, we must not forget that the pharma industry is guided by a strong interest in enhancing human health, while making a profitable business of it. It may happen that, for this purpose, it focuses its attention on the monitoring system rather than on the medicine it incorporates, or worse still, the system is used as a means of revaluing a medicine that would otherwise be almost obsolete.[31]

In this respect, the first digital pill approval paves the way for future marketing of similar drug-device combination products, encouraging other applicants to innovate similarly over older drugs. It is important to notice that the way that ingestible sensor can accompany the drug is particularly relevant from the regulatory process perspective,

28 J. Torous & L. W. Roberts. *The Ethical Use of Mobile Health Technology in Clinical Psychiatry*, 205 J. Nervous Mental Dis. 4–8 (2017). doi:10.1097/NMD.0000000000000596

29 Ho, *supra* note 27.

30 Klugman, *supra* note 16.

31 Hatch, *supra* note 4.

hence for the entrance into the market. Ingestible sensor physically integrated inside the drug, as is the case for Abilify MyCite® capsules, requires a New Drug Application approval—since it falls under the Section 3.2 (e) (1) of Title 21 of the Code of Federal Regulations, and the aripiprazole is the combination product primary mode of action. Nevertheless, in case the sensor is not physically integrated in the pill, but embedded separately inside the same capsule, applicant can take advantage of no requirements to undergo a new round of regulatory approval. In this way, no FDA approval was necessary in a recent use of Proteus sensor in a digital oncology pill, within a program developed with cancer patients in cooperation with University of Minnesota and Fairview Health Services, since the sensor was 'loosely packaged' with the drug in the capsule.[32] Avoiding a time-consuming and costly regulatory process could therefore constitute a great incentive for applicants to place sooner on the market innovative products as digital pills.

Innovation in the pharmaceutical sector and business strategies are closely linked, all the more so since from the beginning of the last decade, the pharmaceutical industry has been experiencing a phenomenon known as 'patent cliff': a massive expiration of pharmaceutical patents.[33] Even Abilify MyCite® developers have mentioned it as a significant factor for their progress. The pharmaceutical market is based on free market and innovation under the umbrella of solid intellectual property regulation[34] . This translates into a situation where once the patent holder of a blockbuster drug loses the patent, they automatically lose the market gap occupied by that drug. From then on, it will be occupied by generic formulations at a lower cost. Against this background, patent holders deploy various business strategies with the intention of patching the hole in their incomes, or to delay entry of most upcoming generic versions into the market.[35]

How has this phenomenon affected Otsuka lately? The market for the previous Abilify formulation—without the ingestible sensor—of the digital pill version, entailed a total of $7.5 billion in the USA for the company,[36] and operations with this drug in North America constituted about 40 per cent of Otsuka global sales.[37] This put Abilify ahead in the top-selling drugs in the USA between 2013 and 14,[38] the year before the patent expired in 2015.[39] The entry of generic versions into the market after patent expiry would result in a calamity for the patent holder. And that was the starting point. In 2015, after several attempts to delay entry of the generics (materialized in various

32 Sara Gerke et al., *Ethical and Legal Issues of Ingestible Electronic Sensors,* 2 Nat. Electron. 329, at 331 (2019).

33 Jack DeRuiter & Pamela L. Holston, *Drug Patent Expirations and the "Patent Cliff"*, 37 6, U.S. Pharm. 12 (2012).

34 European Commission, Competition DG. *Pharmaceutical Sector Inquiry. Final Report.* https://ec.europa.eu/competition/sectors/pharmaceuticals/inquiry/staff_working_paper_part1.pdf (accessed May 5, 2020) and World Trade Organization, *Trade-Related Aspects of Intellectual Property Rights,* at 11 https://www.wto.org/english/docs_e/legal_e/27-trips_01_e.htm (accessed May 5, 2020).

35 Chie Hoon Song & Jeung-Whan Han, *Patent Cliff and Strategic Switch: Exploring Strategic Design Possibilities in the Pharmaceutical Industry,* 5,1 692 SpringerPlus 1, at 3 (2016).

36 Otsuka, Annual Report, at 17, (2014).

37 *Id.*, at 40 (2014), and Leah Ida Harris, *The Rise of the Digital Asylum,* Mad in America, Sept. 15, 2020, at 15.

38 *Id.*, at 17 (2014).

39 The protection period for the substance patent of ABILIFY will expire in Jan. 2016 in Japan (including the 2-year pediatric exclusivity), in Apr. 2015 in the USA (including the 6-month pediatric exclusivity) and in Oct. 2014 in Europe. *Id.*, at 53 (2014).

litigations),[40] the first generic-version aripiprazole entered the market. In the face of this situation, Otsuka, with cooperation from Proteus, introduced an innovation to the obsolete product—the ingestible sensor—which made it new again, and thus allowed for 'evergreen'[41] patenting. This strategy enabled the maintenance of their leadership in the market, at least for the market share represented by patients that did not meet the proper medication-taking adherence.

The answer to the question of why digital pills have entered the market appears to be clear: non-adherence to medication constitutes a major problem—i.e. especially the case for antipsychotics.[42] Hence a large market share would demand a product that monitors treatment adherence. But this does not necessarily mean that digital pills constitute the solution for non-adherence to antipsychotics. Some more reasons are needed for that.

In this sense, we find some social factors that would support digital pills' market entry can be added. The first is a favorable public opinion of treatment compliance by such patients, as non-compliance could involve a hazard to public safety in case they behave dangerously toward themselves, their family, or third parties.[43] They also have the potential advantage of reducing possible tensions within the family, or reducing family anxiety, about treatment non-compliance. Compliance would warrant public—and private—safety, and digital pills constitute a major step for this purpose, as they are not subject to the limitations shown by previous electronic reminders in ingestion tracking.[44]

The second factor is a favorable attitude from healthcare professionals toward a treatment that would substitute the monitoring ingestion alternative: the long-acting injectable antipsychotics (LAIs)—apart from other advantages they might find for such treatment. LAIs are a means of managing treatment periodically, so taking the medication does not depend on the patient, hence neither does compliance nor non-compliance to the patterns given by the physician. Although LAIs make non-adherence impossible, they have been observed to have some limitations as well—such as difficulties in finding the proper dose—and are not suitable for all patients.[45] In addition, digital pills overcome the challenges presented by other alternatives posed by professionals for increasing adherence, such as psychosocial interventions (i.e. psychoeducation), electronic reminders (i.e. smart pill bottle, SMS), other service interventions (i.e. access to emergency services, interventions for reducing medicine prices), or financial

40 *Id.*, at 54 (2014); Wolters Kluwer, *Drug Makers Given Green Light to Market Generic Versions of Otsuka's Abilify Drug*, Intellect. Prop. Law Daily, Apr. 17, 2015 and WCG, FDA, *FDA News*, https://www.fdanews.com/articles/183464-otsuka-loses-again-in-challenge-to-fda-green-light-for-competing-drug (accessed Jan. 16, 2020).

41 Cosgrove et al. define evergreening as 'a strategy used by industry to effectively extend patent protection by making small changes to existing products, changes that have almost no added benefit to the patient', see *infra* note 49, at 236.

42 Jonathan P. Lacro et al., *Prevalence of and Risk Factors for Medication Nonadherence in Patients with Schizophrenia: A Comprehensive Review of Recent Literature*, 63 10 J. Clin. Psychiatry 892, at 892 (2002) and Leah Ida Harris, *supra* note 37, at 12.

43 Peter M. Haddad, Cecilia Brain & Jan Scott, *supra* note 17, at 46–47 (2014).

44 Leah Ida Harris, *supra* note 37, at 14.

45 *Id.*

incentives (i.e. payment in return for taking the medicines, although this last option raises ethical issues).[46]

Finally, the third key factor relates to a questionable favorable attitude—already regarded by some as 'spin'[47]—in the scientific literature and news reports that somehow impacts on both attitudes just mentioned: that of the public and the professionals. There is an underlying concern regarding the scientific support and favorable opinion presented in several reports, which revolves around the real comparative effectiveness of this new-generation drug.[48] Some authors have already highlighted that the approval of this first version of digital pills was based on weak clinical trial evidence. Abilify MyCite is not indicated for adherence, and its impact on it has not been demonstrated. Cosgrove et al.[49] underscore with their systematic review of clinical trials submitted to the FDA three relevant facts: first, in the reviewed clinical trials, no higher or lower efficacy is proved in comparison with the previous nondigital drug, or with other active drug comparators (approved in the USA for the same indication), or with placebo, while at the same time no clear information about drug safety is provided. Second, the clinical trials could only prove that the treatment fulfilled the purpose for which it is indicated: tracking the ingestion; they failed to prove that fact would increase adherence, and therefore there is no way of knowing for certain if this sort of treatment would improve patient quality of life, symptoms, or relapses. It only succeeds in demonstrating that the sensor works properly.[50] Third, Cosgrove et al. also point out an emergent scientific and news tide that distorts interpretation of the evidence shown by clinical trials, which is manifestly biased by conflicts of interests, presenting a greater impression of the benefits than that provided by the data.[51]

Furthermore, it is important to highlight that medicine prices constitute a barrier to adherence.[52] Regarding the data between 1999 and 2015 in the USA about cost-related prescription non-adherence, a study reported that millions of people do not

46 Peter M. Haddad, Cecilia Brain & Jan Scott, *supra* note 17, at 55.

47 Cosgrove et al. define spin as 'a specific way of reporting, intentional or not, to highlight that the beneficial effect of the experimental treatment, in terms of efficacy or safety, is greater than that shown by the results", see *infra* note 49, at 232.

48 Leah Ida Harris, *supra* note 37, at 16.

49 Lisa Cosgrove et al., *Digital Aripiprazole or Digital Evergreening? A Systematic Review of the Evidence and Its Dissemination in the Scientific Literature and in the Media*, 24 6 BMJ EBM 231 (2019).

50 Daniel J. Lee. *Center for Drug Evaluation and Research, Application Number: 207202orig1s000 Clinical review(s)*, at 11 https://www.accessdata.fda.gov/drugsatfda_docs/nda/2017/207202Orig1s000MedR.pdf (accessed Jan. 16, 2020).

51 Cosgrove el al. showed that 10 out of 14 papers that reported on the two studies taken into account did not address the lack of efficacy of the trials. Thirteen out of 14 did not mention the scarcity of data on safety or the fact that no comparator studies were conducted. In 10 out of 14 papers, authors gave an unsupported impression of benefit, and in eight out of 14 there was at least one author who had economic links with Otsuka or Proteus; moreover, in six out of 14 papers, the authors were employees in those companies. When analyzing news reports, lack of efficacy was not acknowledged in 40 out of 70 cases studied, and 65 out of 70 reports omitted information about the lack of safety data and did not include any nondigital comparator. In 52 out of 70 cases, benefits not supported by evidence were reported. In 54 out of 70 cases, experts were cited, but in 21 of those 54 cases, those experts had economic ties with the companies mentioned. See Lisa Cosgrove et al., *supra* note 49.

52 Maria Kelly, Suzanne McCarthy & Laura J. Sahm, *Knowledge, Attitudes and Beliefs of Patients and Carers Regarding Medication Adherence: A Review of Qualitative Literature*, 70 Eur. J. Clin. Pharmacol. 1423, at 1427 (2014).

fill a prescription, postpone a prescription fill, take less medication than prescribed, or skip doses to save money. These figures increase among working-aged adults, women, African Americans, the uninsured, people with disabilities, among others.[53] This new pharmaceutical product costs nearly $1700 per month, whereas the generic-version aripiprazole without the sensor costs $20 per month,[54] which seems to be relevant inasmuch as access to treatments is important for patients with long-term health conditions. From a funding prescription perspective, it is foreseeable that health insurers will pay for this innovative treatment according to the provided cost-effectiveness. To this regard, they would find digital aripiprazole preferable over the nondigital version if its use translates into reduced costs for the coverage of the patient, hence making it worthwhile to opt for.[55] Thus far, there is not enough comparative evidence that shows a major ability for Abilify MyCite® to improve patient's health over the nondigital version. As highlighted above, we can ensure ingestion will be tracked with a high precision, but we cannot anticipate if this circumstance would translate, in all events, in a patient's health improvement.

In addition, the very characteristics of the final users of these digital pills (surveillance paranoia and similar) appear to be discouraging for the approval of a pharmaceutical product that takes surveillance to a higher level. In this sense, potential hazards over the patients derived from the intake of these pills must be approached in a specific and more in-depth study that has not been performed yet.[56] But even in the case that these ethical barriers are overcome by the benefit an eventual high adherence rate would generate, then, as already pointed out by some, the lack of effective outcomes or/and harmful adverse effects that high rates of adherence would generate in patients on long-term therapies, should be considered.[57]

All this being said, latest news reveal the short way gone for the once promising millionaire deal between Proteus and Otsuka. Recently, Proteus has announced that it will now focus its interests on some other fields such as oncology and infectious diseases, bringing the agreement with Otsuka to an end. The reason behind its pivoting direction remains, as pointed by some, in the thorny way chosen by Proteus trying to first expand its system between patients and healthcare providers in the area of mental illnesses: not a lot of them seemed comfortable with this new kind of combined product, a circumstance that turned into too low sales for Otsuka, and to an unprofitable and discouraging economic situation for Proteus. Meanwhile, Otsuka would continue

53 Jae Kennedy & Elizabeth Geneva, *Medication Costs and Adherence of Treatment Before and After the Affordable Care Act: 1999-2015*, 106 10, AJPH, 2016, 1804-1806, at 1806.

54 Lisa Cosgrove et al., *supra* note 49, at 236.

55 Klugman, *supra* note 16, at 42.

56 Lisa Rosenbaum, *Swallowing a Spy—The Potential Uses of Digital Adherence Monitoring*, 318 2 N. Engl. J. Med. 101, 102 (2018) and Lisa Cosgrove et al., *supra* note 49, at 236.

57 Thomas Insel, *Post by Former NIMH Director Thomas Insel: Antipsychotics: Taking the Long View* https://www.nimh.nih.gov/about/directors/thomas-insel/blog/2013/antipsychotics-taking-the-long-view.shtml#1 (accessed Jan. 18, 2020); Lex Wunderink et al., *Recovery in Remitted First-Episode Psychosis at 7 Years of Follow-up of an Early Dose Reduction/Discontinuation or Maintenance Treatment Strategy Long-term Follow-up of a 2-Year Randomized Clinical Trial*, 70 JAMA Psychiatry 913, at 919 (2013); Leah Ida Harris, *supra* note 37, at 19.

developing medicines with the use of Proteus system, in a kind of a fully paid-up license conceded by Proteus for a transitional period.[58]

We can conclude that this scenario could anticipate the eventual consequences generated by the entrance in the market of some cutting-edge digital health products.

RESPECT FOR PATIENTS' AUTONOMY

Patients' autonomy is a fundamental value. It might be the most important value in the way we understand medicine these days. It means that we have finally accepted that patients have the last word in making decisions that will have consequences in their own health or life. Therefore, patients' self-determination constitutes a sort of last boundary that should never be violated by physicians, healthcare providers, social workers, or any other person who may be involved in a caring relationship with a patient. This iron rule cannot be overridden by considerations such as the best interest of the patient (beneficence). Otherwise, we would be indulging in paternalism, a practice that has lasted for too long in healthcare. Do digital pills involve a restriction on patient autonomy?

The response to this crucial question is not easy to provide, as many different variables play a role in the answer. Synthetically, we dare say in advance that they neither violate patient autonomy if this autonomy does not exist, nor violate it without justification if there are good reasons to annul it (such as public interest), nor restrict it at all if patients are competent to consent and show willingness to use the tool, as long as they provide a real informed consent—which means they have been informed properly—and freely consent to it. Instead, they would definitively violate such autonomy if the patient would not provide real informed consent, a scenario that might be present under several common circumstances. To this respect we must address very cautiously the information issue in the user agreements these tracking systems imply, since a lack of agreement could translate into a lack of access to the treatment, and then into a pressure over patients to accept some clauses they might not really agree with, thus, making not an optimally autonomous choice. We will approach this question in Section 6.

Starting from the easiest scenario, we concede that digital pills involve no risk for autonomy if the patient consents to their use under such circumstances just mentioned: a really informed consent. This might happen for multiple reasons. For example, patients with memory loss might be willing to use a system that would serve them well to avoid overdoses while reducing the anxiety stemming from the doubt of whether they have taken the pill. Similarly, patients could be looking forward to benefitting from a tool that allows them to demonstrate to their doctors that they are following the provided treatment strictly. Alternatively, they could be proud to use a modern technology that allows them to incorporate their own impressions about the treatment in an agile way.

The reasons for acceptance are indeed uncountable, and we do not think that our mission should be to focus on them. Instead, we should concentrate in cases in which patients are not willing to adhere to the use of digital pills. In our opinion, this would

58 Dave Muoio, *Proteus Parts Ways with Otsuka As It Pivots Toward Oncology, Infectious Disease Treatment Adherence*, Mobile Health News, Jan. 14, 2020, and Rebecca Robbins, *A Forerunner in 'Smart Pills' Adopts a New Tack as Key Pharma Partnership Unravels*, STAT, Jan. 14, 2020.

not necessarily act as the definitive reason for avoiding their use. First, one must think whether acting against the patient's will violates their autonomy, and then consider some situations that would yet justify this violation.

The first thing to take into account is the patient's legal capacity. In case of legally incapacitated patients, their legal representatives have to decide on whether to adopt digital pills—respecting the ultimate patient's interests and counting with their participation in the decision-making process as far as possible. This would be the case of minors or legally incapacitated people.

More complex is the case of other people who do not have a permanent or lasting restriction on their autonomy, but who find themselves in circumstances that advise the use of digital pills. Imagine, for example, the case of a person affected by a particularly serious contagious disease that requires the administration of a specific medication for treatment; or the case of a mental patient whose pathology is associated with violent outbursts that may endanger other people. In all these cases, the patient poses a threat to public health. It is therefore necessary to adopt measures capable of neutralizing it. At present, this is done through mechanisms such as quarantining or confining the patient in a health facility, where the medication is administered in a forced manner.

The appearance of digital pills promotes an alternative to this situation, as it allows monitoring the administration of treatment without confining the patient (unless the danger of contagion is unavoidable, or patient's values and preferences are in accordance with the confinement, in which cases confinement is legally supported). In these circumstances, recourse to this new technology would undoubtedly be contrary to patient autonomy, but much less than the alternative possibility of confinement. This fact would justify its use even against the patient's will.

Finally, we must consider the case of the largest group of patients, i.e. those who possess full faculties for consenting to a treatment and whose pathologies do not pose a public health or public safety problem. In all these cases, it is not possible, in our opinion, to justify the use of digital pills if it is not through the consent of the affected person. Moreover, consent must be obtained through a process that provides the patient with adequate information and guarantees freedom of choice. This is particularly relevant when we are talking about vulnerable populations, such as the mentally ill, the elderly, or people with low levels of education, as well as in people with very little social support. In these cases, apparent acceptance often hides a desire to not lose the approval of their scarce social links. As Dotolo et al.[59] wrote, 'When the technology embedded in AMC is introduced to clients and families by prescribers, its use is normalized, if not tacitly endorsed. Although formal policy may require informed consent for AMC prescription use, social workers understand that freely given consent in practice is often complicated by difficulty understanding consent forms and processes (Schenker, Fernandez, Sudore & Schillinger, 2011), power asymmetries (Barusch, 1987), and borderline coercive practices in the context of caregiving (Berridge, 2017). Once the technology is broadly adopted and normalized, it may be featured in mandated treatment or coercively encouraged by family members and service providers in the name of beneficence and safety'.

59 D. Dotolo, Petros, R., & Berridge, C. *A Hard Pill to Swallow: Ethical Problems of Digital Medication*, 63 Soc. Work 370–372 (2018). doi:10.1093/sw/swy038

In such situations, where using smart devices for healthcare becomes normalized, not participating in such a self-care technological paradigm could even be contemplated as a basis for exclusion from access to health services.[60] In addition, we may consider that being aware of the risks a person is exposed to appears to condition the person into adopting the necessary measures for safeguarding themselves by controlling those already known risks,[61] a scenario in which smart devices could be extremely useful.

Therefore, we need to be particularly vigilant to ensure that patients have good understanding of the implications of the use of digital tagging. On top of that, we will have to strive to provide a framework that allows them to express their opinions freely and support their decisions, attempting to reduce the hostility they may arouse in their social support networks, family, or friends.

In any case, the dilemma must not be seen as an all-or-nothing decision. It is not true that the alternative to the adoption of digital tagging is the loss of absolute control over the patient's behavior. Today, medication management tools that serve these purposes well even though they limit the patient's autonomy much less are already in use. It is true that they probably do not provide with such exact information. However, it will be necessary to assess in which cases the difference in precision would endorse the imposition of a measure—the use of digital pills—which represents the considerable loss of a person's autonomy.

Finally, it is good to remember that, in general, people are allowed to refuse a medical treatment due to a number of reasons that are not necessarily rational. It would be unusual to make an exception in the case of digital pills. It is widely accepted that choosing between welfare and peace of conscience is a decision to be taken from one's own deepest autonomy, as an expression of the ownership of rights, without the State or third parties playing a role in the decision-making—except legal incapacity cases.[62] We could consider the paradigmatic example of a Jehovah's Witness's decision in rejecting a blood transfusion, and hence deciding to preserve their freedom of consciousness at the expense of their health or even their life. To support this, we can follow Stuart Mill: 'The only part of the conduct of any one, for which he is amenable to society, is that which concerns others [. . .] Over himself, over his own body and mind, the individual is sovereign'.[63] This is the idea that has constituted the guideline in designing a healthcare system based on autonomy, stripped of paternalism patterns.

AUTONOMY AND USER AGREEMENTS

The use of digital pills implies the need to address a particularly complex issue in terms of informed consent, which is not present in all physician–patient relationships. As we have explained, digital pill systems include both a drug and a digital tracking system, i.e. three electronic devices: the IEM, the patch, and the mobile app. The issue is that the use of these devices requires the acceptance of some conditions of service, i.e. of a consent that is unrelated with the consent related to the administration of the drug.

60 Dimitra Petrakakia, Eva Hilbergb & Justin Waringc, *supra* note 22, at 149.

61 Sonja Erikainen et al., *Patienthood and Participation in the Digital era*, 5 Digital Health 1, at 6 (2019).

62 Francisco Bueno Arús, *El consentimiento del paciente en el tratamiento médico-quirúrgico y la Ley General de Sanidad*, *in* Estudios de Derecho penal y criminología, 163 (UNED, 1989).

63 J. Stuart Mill, On Liberty, at 13 (1859).

Here, we may find the first concern about this issue: this privacy policy is sort of a take-it-or-leave-it contract (usually termed an adherence contract). These agreements often consist of hundreds of pages written in technical language. Sometimes they hide clauses that enable manufacturing companies to manage the data collected for purposes other than monitoring the treatment of the patient involved.[64] In these—mostly common—situations, the patient may be authorizing uses they would not be able to understand due to the complex terminology. This question keeps the door open to a wider discussion about data privacy: future data use, eventual collections of identifiable patient information, access by stakeholders to patient and physician data collected the mobile app, and the web portal used by them, etc. The second concern we find is that when the privacy policy is provided only by the mobile app, there is a risk that document only refers to the app, overlooking sensor and patch privacy issues about which the user should be informed.

In both situations, the patient would not be making a properly informed decision: in the first situation, it is because of the lack of understanding about what information is collected and how it is used; in the second situation, because the patient has no way of knowing the risks implied. Hence, such privacy policies fail to protect the consumer that, in this case, meets a patient status whose autonomy is infringed.[65]

Moreover, such consent is far removed from the norm in the practice of medicine, as it is not based on face-to-face information and a negotiation of the terms of treatment, but on a user agreement that cannot be discussed with the provider. Pretending that a person can, by their own means, provide informed and free consent to the use of these devices is, in these conditions, not very credible. Although some of our current regulations (such as the General Data Protection Regulation in the European Union context) accept that a mere 'box ticking' serves to capture the existence of consent, the truth is that this rarely happens. What really happens is that very often we sign a consent form to access a service without having any idea of the terms of the contract. This, which is worrying in any sphere of human life, is even more so in the field of health. This might become even worse if acceptance of the use agreement becomes a condition of access to the drug. In such cases, we could think about an absolute perversion of the system of consent to treatment.

It is therefore necessary to create new mechanisms capable of tackling this problem effectively, ensuring an effective defense of the patient's interests. Some authors have postulated an adaptation of the traditional informed consent.[66] This way, healthcare providers would be the player committed to informing the patient about such privacy issues. There are some advantages to this proposal: the patient will be informed before buying the treatment, and will likely better understand when that information is communicated face-to-face by a trusted person (doctor) instead of from a legal document. Some would say that studying privacy policies would take a long time for physicians, apart from exceeding their competences, but, that is, when doctors may

64 Klugman, *supra* note 16.

65 Amelia Montgomery, *Just What the Doctor Ordered: Protecting Privacy Without Impeding Development of Digital Pills*. 19 1 Vand. J. Ent. & Tech. L. 147, at 168 (2016).

66 Article 29 Data Protection Working Party, *Opinion 8/2014 on the on Recent Developments on the Internet of Things,* at 7, https://www.dataprotection.ro/servlet/ViewDocument?id=1088 (accessed 5 May, 2020).

ponder the sacrifices they must make and the benefits obtained by prescribing a digital pill treatment.[67]

We could certainly think of many other alternatives, which should be carefully explored in the future. This may well result in a form of paternalism, but in our view, it would not be an immoral type. Paternalism is only reprehensible when someone tries to supplant the will of the patient on the basis of the alleged pursuit of his welfare. However, if the patient is incapable of giving consent because the process makes it impossible in practice to be adequately informed, then we are faced with a situation of vulnerability, in which the intervention of a third party to protect their interests is unavoidable from an ethical point of view. Thus, an administrative intervention capable of putting conditions on use agreements and their updates, or a system that allows patients to access reliable information on the real content of these agreements, seems to be a more than reasonable option.

THE ISSUE OF PRIVACY

The use of digital pills has strong implications on patient privacy. To begin with, it is necessary to stress that the introduction of this technology puts an end to the monopoly of power over information on the observance of the treatment possessed by patients. Up until now, and despite the existence of mechanisms that allow adherence to treatment to be monitored in some way, the truth is that patients are still the only ones who know for certain whether they are following the indicated doses. This is due to the simple fact that, ultimately, only the patients know whether they are taking the prescribed tablet and when. With the inception of this technology, however, that monopoly was broken. There is an alternative source for the doctor capable of providing extremely accurate information. Consequently, the patient's privacy is unavoidably threatened. The security of secrecy is no longer in their hands alone. To begin with, their doctor will have direct access to the information without having to consult with the patient. This in itself is not the worst threat to the patient. Healthcare professionals have been used to professional secrecy for generations and the law protects patients against indiscretion.

Nevertheless, this scenario introduces, in addition, a third party to play a role in between physician and patient, and who will have access to all the data collected by the device: the device developer. Developers need that access to procure the minimum safety and effectiveness levels for the service they are offering. When using that data properly anonymized and for legally contemplated purposes (such as investigations), no explicit consent is needed as long as developers comply with the applicable laws. But when we talk about highly protected information—health data—this circumstance raises serious challenges related with data breaches and deanonymization.[68] Consequently, while they are accessing the same data as the physician, developers should not abdicate the same secrecy and confidentiality responsibilities demanded of healthcare professionals.

Furthermore, we must keep in mind that friends or relatives around patients might access the data. Different from healthcare providers, they are not legally obliged to keep

67 *Id.*, at 173 and 174.

68 Glenn Cohen et al., *The Legal And Ethical Concerns That Arise From Using Complex Predictive Analytics In Health Care*, 33 7 Health Aff. 1139, at 1141 (2014).

the data confidential, nor have they been trained to do so. On the other hand, it is quite obvious that the mere knowledge that patients are using this type of device can lead to great pressure being exerted from their environment to share the data. Take, for example, the case of a bipolar person living in the house of a brother, son, or their parents. Do we not think that there will be many cases in which the relationship of economic dependence is used to gain access to data? Even more dangerous are the pressures that insurance companies could exert on their policyholders to gain access to the data. If it were legally possible, it is likely that some would try to condition the funding of these devices based on the possibility of appropriating the resulting information, or at least, offering discounts to policyholders who allow it. Data provided by tracking devices could be used in health decision-making, either for ensuring compliance with therapeutic recommendations, or as a consequent fairer distribution of health resources ('if you don't lose weight/don't take the pills on schedule, then you lose the right to undergo a surgical procedure/to be covered up to this insurance policy').[69]

Finally, we must not underestimate the possibility of the stored data being used in police/judicial instances, perhaps as a condition for a convict to be released on bail ('you take the pills on schedule, or you go back to jail'), or perhaps as evidence against the patient himself. Some bioethicists consider this method more reliable than just trusting a detainee/convict's word.[70] This judicial use has already happened in some cases, for example, 'in one reported case, police sought a search warrant to access pacemaker data of a patient they suspected of arson'.[71]

THE SOCIAL PRESSURE FACTOR

Finally, it is important to bear in mind that the existence of an objective measure of adherence to treatment can lead to moral, social, or even legal censorship of patients who sustain in general a lack of adherence to treatment. Lack of adherence to treatment constitutes, thus, a neglect of responsibilities assumed by the person once they are aware of their behavior and the risks derived from it. In addition, datafication of patients provides the possibility of creating new categories of patients according to the information they generate;[72] consequently, we run the risk of constructing a scenario in which it is possible to distinguish between 'good' and 'bad' patients depending on their adherence to treatment. We might even be tempted to impose sanctions on the second group, a temptation that has already given rise to action in this regard.[73] We could assume that if the individual behavior generates harm over the interests of others, without the existence of a higher duty that obliges a person to behave that way, they deserve moral disapproval. Not so if the consequences of that behavior only affect their own interests.[74] The question is whether (or when) this behavior—not taking pills when prescribed by the physician—constitutes a damage of the interests of others, and hence could be punished by the community.

69 Sarah Chan, *Bioethics in the Big Data Era: Health Care and Beyond*, 41 Rev. Bio. y. Der. 3, at 11 (2017).

70 R. Brandom, *The Frightening Promise of Self-Tracking Pills*, The Verge, Oct. 7, 2015, at 1.

71 *Telltale Heart: Pacemaker Data Leads to Arson, Fraud Charges*, Fox News U.S. (2017), http://www.foxnews.com/us/2017/02/08/police-use-data-on-mans-pacemaker-to-chargehim-with-ohio-arson.html

72 Sonja Erikainen et al., *supra* note 61, at 6.

73 Moutel, *supra* note 7.

74 Stuart Mill, *supra* note 63.

In our view, differentiating between 'good' and 'bad' patients and imposing sanctions on the latter would be a fatal mistake for many reasons. In general, constructing the figure of the 'guilty patient' is a mistake, a moral injustice that can eventually lead to State interference in private life. Second, this vision of the world sacralizes science, thinking that it is possible to objectively set optimal treatment guidelines that everyone should follow faithfully, although this is not how things work. As we have mentioned, reasons for non-adherence could be multiple and complex, and we should not fall into the error of thinking that this phenomenon is only understandable and approachable in one way. If there is one thing the evidence shows us, it is that each patient responds individually to a treatment, so unless we are able to optimize the doses for each patient, we will have to assume a margin of error. Moreover, we have to assume that this margin empowers the patient to deviate from the intended dosage without there being any evidence that this will lead to worse treatment performance. Moreover, we must keep in mind that there are times when strict adherence to treatment can be very difficult or even harmful for the patient, either because of the physical adverse effects it causes or because of the lifestyle changes it inevitably imposes. While the healthcare professional may prioritize healthcare understood to mean perfect adherence to the treatment, the patient may prioritize well-being in a wider sense, more related with a quality of life concept.[75] Therefore, it seems reasonable to assume that patients would know better how to deal with a treatment so that their lives are improved effectively instead of the opposite. This is usually known as 'self-efficacy', a concept developed by Bandura,[76] which shows the need to pay attention to the circumstances at stake and the wisdom of the patient's decision.

FINAL REMARKS

It is quite difficult to deny that smart pills could be useful for increasing adherence to treatment. If we are able to force patients to consume a medical device that inform healthcare providers if they have taken the pill, and combine this with the threat of punishing any lack of adherence with forced confinement of the patient, surely the intended objective will be achieved. However, this is not, in our view, the ideal way of ensuring better adherence. If studies show anything, it is that adherence improves with better understanding of the need for medication and a fluid and permanent dialogue between the patient and healthcare workers, which makes it possible to reduce harmful adverse effects. Obviously, there will be patients for whom all this is impossible. There will also be others where resistance to treatment is irrational. But, in general, we believe that these new systems should ideally be seen as a means of complementing traditional strategies for promoting adherence to treatment, and not as a substitute. Only in this manner can we obtain a final result that is not reduced to an increase in adherence rates subject to inadequate limitation of patient autonomy. As stated, 'Automatic and computerized data collection, related to the follow-up of a treatment, will require us to consider the following the question of benefit/risk assessment. The evaluation of

75 Victoria Camps, Una vida de calidad. Reflexiones sobre bioética, at 72 and 73 (2001) and Rosana Triviño, El peso de la conciencia. La objeción en el ejercicio de las profesiones sanitarias (2014).

76 Bandura A. *Health Functioning*, in SELF-EFFICACY: THE EXERCISE OF CONTROL, 259–318 (SF Brennan ed., 1997). W.H. Freeman and Company.

a connected device will require to study how its use improves or not the quality of the follow-up and, ultimately, the patient's quality of life. And to analyze if risks would not offset these potential benefits (whether or not fundamental freedoms are infringed, psychological impact of fear of surveillance, increased anxiety, etc.)'.[77]

FUNDING

Iñigo de Miguel Beriain's work received funding from the EU Commission, H2020 SWAFS Programme, PANELFIT Project, research grant number 788039. Marina Morla's work received funding from the Programa Propio de Investigación de la Universidad de León.

77 Moutel, *supra* note 7.

Journal of Law and the Biosciences, 1–19
doi:10.1093/jlb/lsaa040
Original Article

Acknowledgements:

This project has received funding from the European Union's Horizon 2020 research and innovation programme under grant agreement No 788039

Iñigo de Miguel Beriain, Marina Morla González, 'Digital pills' for mental diseases: an ethical and social analysis of the issues behind the concept, *Journal of Law and the Biosciences*, lsaa040, https://doi.org/10.1093/jlb/lsaa040