Opinión / Opinion
Legal aspects of genetic databases for international biomedical research: the example of the International Cancer Genome Consortium (ICGC)*

VV.AA. **

_____________________

* This work (Carlos Romeo-Casabona and Pilar Nicolás’s work) was funded by Spanish Ministry of Economy and Competitiveness through the Instituto de Salud Carlos III (ISCIII).

** The authors thank Nikolajs Zeņš for his useful comments.

Carlos Romeo-Casabona. Member of the Ethics and Policy Committee (EPC) of the ICGC. Professor of Criminal Law. Director of the Interuniversity Chair in Law and the Human Genome, University of Deusto, University of the Basque Country UPV/EHU, Bilbao, Spain. Avenida de las Universidades 24, 48007 Bilbao, Spain. carlosmaria.romeo@ehu.es / cromeo@genomelaw.deusto.es

Pilar Nicolás. Member of the Ethics and Policy Committee (EPC) of the ICGC. Member of the International Data Access Committee (IDAC) of the ICGC. Senior Research fellow. Interuniversity Chair in Law and the Human Genome, University of Deusto, University of the Basque Country, Bilbao, Spain. Avenida de las Universidades 24, 48007 Bilbao, Spain.

Bartha Maria Knoppers. Member of the Ethics and Policy Committee (EPC) of the ICGC. Member of the International Data Access Committee (IDAC) of the ICGC. Professor. Faculty of Medicine, and Faculty of Law. Director of the Centre of Genomics and Policy. Faculty of Medicine, Dept. of Human Genetics. McGill University, 740 Dr. Penfield Avenue, Rm 5101 Montreal, QC., Canada H3A 1A4. bartha.knoppers@mcgill.ca

Yann Joly. Member of the Ethics and Policy Committee (EPC) of the ICGC. Member of the International Data Access Committee (IDAC) of the ICGC. Assistant Professor. Centre of Genomics and Policy. Faculty of Medicine, Dept. of Human Genetics, McGill University. 740 Dr. Penfield Avenue Rm 5101 Montreal, QC., Canada H3A 1A4. yann.joly@mail.mcgill.ca

Susan E. Wallace. Member of the Ethics and Policy Committee (EPC) of the ICGC. Lecturer of Population and Public Health Sciences. Department of Health Sciences. University of Leicester. Adrian Building, University Road. Leicester LE1 7RH. sew40@le.ac.uk

Don Chalmers. Member of the International Data Access Committee (IDAC) of the ICGC. Distinguished Professor. Director of the Centre for Law and Genetics. University of Tasmania. Sandy Bay Campus, Faculty of Law Building, 2.03. University of Tasmania. Australia. Don.Chalmers@utas.edu.au

Stephanie Dyke. Member of the Ethics and Policy Committee (EPC) of the ICGC. Policy Adviser. Wellcome Trust Sanger Institute. Wellcome Trust Genome Campus. Hinxton, Cambridgeshire, CB10 1SA. United Kingdom. sd4@sanger.ac.uk
Sumario / Summary: 1. International genetic databases and scientific research: legal framework. 2. Ethical and legal issues of international transfer of genetic data. 2.1. Consent and health data. 2.2. Personal data. 2.3. Confidentiality. 2.4. Privacy. 2.5. International transfer. 2.6. Participant’s ability to object. 2.7. Dissemination of research results. 2.8. Ability to withdraw. 2.9. Ethical and legal issues of international transfer of genetic data. Concern in the ICGC. 3. International Cancer Genome Consortium (ICGC): policy governing management and transfer of data. 3.1. National institutions. 3.2. The national research team and data controllers. 3.3. International persons or institutions. 3.4. International data transfer. 3.5. Oversight: the ICGC controlled-access database model. 4. Conclusions.

Resumen / Abstract: Existe una notable carencia de regulación internacional sobre el intercambio de datos personales y la gestión de la investigación. Este artículo arroja luz en este ámbito mediante la descripción de cómo el Consorcio Internacional del Genoma del Cáncer está desarrollando políticas y procedimientos para abordar las cuestiones éticas y jurídicas que plantea la transferencia internacional de datos y resultados. El objetivo de estas políticas y procedimientos, es, en primer y más importante lugar, salvaguardar los intereses de los participantes en la investigación y de otros actores involucrados y, en segundo lugar, facilitar el intercambio de datos y resultados a fin de obtener mayores beneficios de este tipo de investigación genética de colaboración internacional.

There is a noticeable lack of international regulation on personal data exchange and management in research. This article sheds light in this area by describing how the International Cancer Genome Consortium is developing policies and procedures to address the ethical and legal issues raised by the international transfer of data and results. These policies and procedures aim, first and most importantly, to safeguard the interests of the research participants and other involved stakeholders and, secondly, to facilitate the sharing of data and results to realize greater benefits from this kind of internationally collaborative genetic research.
1. International genetic databases and scientific research: legal framework

Human genetic databases are essential tools used to translate biomedical and pharmaceutical research into improvements in health care and in the identification, diagnosis and treatment of diseases. Increasingly, researchers are collaborating and exchanging results between different genetic research databases in different countries. Apart from critical issues of privacy, genetic databases raise a broad range of ethical and legal issues relating to consent, governance, the handling of human tissue, confidentiality, benefit-sharing and international collaboration. These databases rely on the participation of patients and healthy volunteers. This research must be conducted with integrity and according to the highest ethical standards, as public trust is an essential pre-condition for the successful operation and maximization of future benefits derived from genetic research databases.

The International Cancer Genome Consortium (ICGC) is an international project involving collaboration, data exchange and linkage between databases in fourteen member countries. The ICGC coordinates large-scale cancer genome studies that are studying tumors from 50 different cancer types and/or subtypes. A focus is placed on cancers that have clinical and societal importance. An international consortium approach was chosen as founders agreed that no one country could

---

1 For a helpful list of ethical tensions and issues in biobanking, including consent, ownership and IP, governance, public engagement, data-sharing, research access, security, privacy, benefit-sharing, commercialisation, discrimination, public good, cultural sensitivity and international harmonisation, see CAMBON-THOMSEN, Anne et ál., “Biobanks for genomics and genomics for biobanks”, Comparative and Functional Genomics, Vol. 4, 2003, pp. 628-634.


conduct the project on its own. Also, some cancers are more common in certain parts of the world and more widely studied. By bringing these different research projects together, the ICGC gives researchers opportunities to combine their resources and work together to accelerate research and the dissemination of data and analytical methods to the scientific community. The ICGC member projects are expected to act in conformity with their own national laws and the international guidelines applicable to biobank research, while adhering the requirements of the ICGC. To assist them, the ICGC has and continues to develop and publish policies and procedures\(^4\) both to safeguard the interests of the research participants and other involved parties, and to facilitate data and results sharing of its member projects to optimize the benefits of this kind of international collaborative genetic research.

Despite the growth of international research tools like genetic databases and collaborations such as the ICGC, there is still no single source of law governing their international operation. There is a range of international declarations, recommendations, guidelines and literature on the ethical and legal implications of human genetic databases\(^5\), but these are generally not legally-binding (except for broad legal instruments that are not genetic specific and address only ancillary topics). The legally binding rules that do exist are mostly present at the national

---

\(^{4}\) Available at: http://www.icgc.org/daco


level (with some transnational agreements, such as in the EU\(^6\)). Consequently, any international project using databases must be conducted according to international best practice standards harmonized with the national laws of the countries involved.

At the national level, privacy laws operate in many countries and, in developed countries, enshrine the OECD *Guidelines on the Protection of Privacy and Transborder Flow of Personal Data* of 1980\(^7\). These apply to the collection of genetic data, its processing and use, and particularly to the consent required for its legitimate use. One key point is that data collection and processing should be carried out only for specified, explicit and legitimate purposes. Some countries also have legislation, regulation, and policies that govern the operation of genetic databases. However, since national laws and regulations cannot be applied at the international level, there should be identified other instruments operating at that level. The 2009 *Madrid Resolution on International Standards on the Protection of Personal Data and Privacy*\(^8\) takes the main principles of the 1980 OECD Guidelines, the Convention 108 of the Council of Europe\(^9\) and the Directive 95/46/EC\(^10\) and situates them in a broader geographical context. This can be seen as an important step in the establishment of a common legal framework under which international research projects can operate providing clearer and more secure criteria for greater interoperability.

\(^6\) There are norms on personal data that are binding on European countries, such as the Convention 108 of the Council of Europe for the Protection of Individuals with regard to Automatic Processing of Personal Data and the Directive 95/46/EC on data protection.

\(^7\) Available at: [http://www.oecd.org/document/18/0,3343,en_2649_34255_1815186_1_1_1_1,00.html](http://www.oecd.org/document/18/0,3343,en_2649_34255_1815186_1_1_1_1,00.html) [Accessed 21 Dec 2011].


At the international level, the 2003 UNESCO *International Declaration on Human Genetic Data* addresses issues of data protection and exchange as well as the scientific value of processing such data to warrant the establishment and continuation of a genetic database\textsuperscript{11}. Similarly, the 2009 OECD *Guidelines on Human Biobanks and Genetic Research Databases* sets out comprehensive series of guidelines on genetic databases. Though both of these documents are influential internationally, neither is legally binding. International approaches are needed to facilitate the development of ethical collaborative research. Hence, there must be a system to ensure compatibility and comparability of results from different research studies and the appropriate delivery of benefits. Formal, legally binding agreements relating to data management and sharing are being drafted. Such agreements will help create a network between professional and academic bodies assigned to make decisions affecting data integration, respond to requests for access, and determine the basis upon which access is to be granted. However, the enforcement of such agreements across borders is not easy\textsuperscript{12}.

This paper will analyse how the ICGC policies and procedures on data sharing and data protection could promote the development of international best practices and serve as an example for collaborative cross-border human genetic research, notwithstanding the current lack of an established international legal framework. The ICGC policies and procedures will supplement existing Research Ethics Board (REB) requirements within institutions. As such, the ICGC may serve as a model for similar projects, currently underway or in the future.

2. Ethical and legal issues of international transfer of genetic data

As noted above, international collaborations between genetic databases raise a broad range of ethical and legal issues such as consent, confidentiality, privacy, data transfer and participant withdrawal from research projects.


2.1. Consent and health data

Personal health data is generally deemed to be sensitive data. Since genetic data is considered to be health data it is also classified as sensitive data. The express consent of the participant is therefore usually required for processing such data including transferring it to someone other than the data controller or processor. Thus, it is important to establish what information must be provided to participants to ensure that they understand the terms upon which they are consenting to the transfer of their data, including possible transfer to an international database. Even in cases where the data to be transferred to the international database is not considered personal, in certain circumstances, information about the transfer must be provided to the research subject, because the duty to provide such information is not only derived from the right to privacy, but also from the principle of autonomy and from the individual’s right to know the terms of his or her contribution and participation in research. Therefore, valid consent by a participant requires information about:

a) the identity of the data controller or representative, if any;

b) the purpose of the processing;

c) further relevant information, such as:

- possible future commercialization and intellectual property
- the recipients or categories of recipients of the data;
- the opportunity to ask further questions concerning any doubt;
- their ability to access and to rectify the data (for example, if applicable, when further information is necessary having regard

---


14 For the terms “controller” and “processor” we use the definitions provided by the Data Protection Directive 95/46/EC.

15 It should be taken into account also that anonymization process requires the affected person’s prior consent or whether it is necessary at least to inform him / her of this process. ROMEO CASABONA, Carlos María, “Anonymization and pseudonymization: the legal framework at a European level”, The Data protection Directive and Medical Research Across Europe, Ashgate, United Kingdom, 2004, pp. 33-50: 42, 43.

16 The information given before consent to use genetic data should include the circumstances about their “subsequent processing, use and storage”. See the UNESCO International Declaration on Human Genetic Data, 2003, article 8 a.
to the specific circumstances in which the data are (were) collected), measures to ensure fair processing in respect of the data subject and, the nature and obligations of the international consortium (organization) which will receive the data; and

- where responsibility for compliance rests within the national database.

With regard to personal data processing, a Research Ethics Board (REB) will usually consider three distinct aspects: First, the operation of a database under national law; second, participant consent and knowledge about the transfer to and storage and use of their data by an international database; and third, participant informed consent to the possible transfer of their data from the international database to other researchers. Obviously, all research participants should be informed of any proposed intellectual property claims and commercialization potential.

2.2. Personal data

There has been much discussion concerning personal data and difficulties in determining when particular genetic/medical data constitutes personal data. Despite the consensus that personal data is data that can be linked to a real person, several situations are problematic. Sometimes the possibility of identifying the person exists but is very remote; or it is only the processor, the controller or other “qualified” user of the database who could do it (but this is not possible for other users of the database). Moreover, often in the context of scientific research, the same data are in two files, and only one person has responsibility for the key codes. Must the coded data still be seen as personal in every situation and for everyone?  

According to Article 29 European Commission Data Protection Working Party: “Where identification of the data subject is not included in the purpose of the processing, the technical measures to prevent identification have a very important role to play. Putting in place the appropriate state-of-the-art technical and organizational measures to protect the data against identification may make the difference to consider that the persons are not identifiable, taking account of all the means likely reasonably to be used by the controller or by any other person to identify the individuals. In this case, the implementation of those measures are not the consequence of a legal obligation arising from Article 17 of the Directive” (which only applies if the information is personal data in the first place), but rather a condition for the information precisely not to be considered to be personal data and its processing not to be subject to the Directive” Article 29 Data Protection Working Party, European Commission, Opinion 4/2007 on the concept of personal data, page 17.
It should be noted that current legal requirements in Europe are bi-modal and that similar obligations also apply in the United States: “if the information is identifiable, then all the legal protections are applicable; if the information is not identifiable, then there are no protections whatsoever”\textsuperscript{18}. The right to personal data protection refers to any information related to an identified or identifiable person. This right requires consent prior to every data transfer. If this principle were to be applied rigidly, international genetic databases would be unable to operate unless data were anonymised. Two approaches can be adopted. First, coded data may not be considered personal data if some guarantees have been taken to avoid the identification of the participant by the researcher (user). Alternatively, a broad consent to the data transfer may be justified when a Research Ethics Board (REB) has evaluated the scientific interest to access to the data and sufficient guarantees have been adopted to protect the security of such data.

These legal mechanisms do not fit easily with the realities of biomedical research. The hypothetical possibility of re-identification does not automatically make data “personal” data. The possibility of re-identification (by reasonable means) can be seen as neutralized or minimized depending on the concrete aims of the research and the security measures adopted. \textsuperscript{19}If in this singular context, it seems unreasonable to think that coded data are going to be recoded, the data should not be considered as personal. So, the need to seek the consent of donor participants for each transfer of their data would not be compulsory. Where appropriate safeguard measures have been put in place throughout the process of data transfer, and there is sufficient oversight over the approval of researchers’ requests for access, participants’ rights may be sufficiently protected so do not require specific re-consent for every transfer of their data. Hence, the policy of all the stages of data processing, the ambit of the consent and the mechanisms to access and use the data, will determine if there is sufficient data protection to comply with privacy requirements.

2.3. Confidentiality

Database researchers and staff are ethically bound to respect participant confidentiality\textsuperscript{20}. This ethical duty may be included expressly in the

\textsuperscript{19} \textsc{The German Ethics Council}, Opinion: \textit{Human biobanks for Research}, 2010, p. 20.
\textsuperscript{20} UNESCO International Declaration on Human Genetic Data, article 14.
contract of employment of the researchers and staff, thus transforming it into a legal duty. The ethical and legal duties of confidentiality require staff and researchers to maintain the confidentiality of information acquired in the course of their work or research. Breaches of confidentiality are serious and can lead to dismissal from employment or other sanctions.\textsuperscript{21}

2.4. Privacy

The right to privacy is internationally recognized\textsuperscript{22} and involves a confidentiality commitment by all those who access participants’ data. If national researchers, data controllers / processors breach their duty in relation to their patients / participants, national law applies. At the national level, privacy law generally enables participants to access their records and files and know what personal data are being stored.

There is a question of who should be held responsible if a breach of security occurs in the context of an international database or mistakes are made in the authorization for access or transfer of data. There are also questions concerning the sanctions that could be invoked. Very little research and community debate has been undertaken on this important topic, an unfortunate situation given the need for the highest degree of standardization and transparency when designing controlled-access agreements. Researchers should undertake empirical studies on the clarity and accessibility of existing database access agreements and related policies in the near future.\textsuperscript{23}

2.5. International transfer

Consent should be sought for the transfer of personal data (identifiable data). According to general ethical and legal principles, a researcher should not transfer data to another research group without this consent, as well as the approval of an REB for the research.

\textsuperscript{21} Where a database is established by legislation, the Act may includes a statutory offence for unauthorized disclosure of information, see Estonia, Human Genes Research Act 2001.

\textsuperscript{22} E.g. The United Nations International Covenant on Civil and Political Rights, Article 17: “No one shall be subjected to arbitrary or unlawful interference with his privacy, family, home or correspondence, nor to unlawful attacks on his honour and reputation”.

\textsuperscript{23} \textit{See} Joly, Yann / ZePS, Nik / Knoopers, Bartha Maria, \textit{op. cit.}
Two approaches can be identified in this context: to ask for a broad consent or to enlarge the definition of “non identifiable data.” In both cases, implementation measures should limit re-identification. Another possible avenue is to replace either option by a “global governance strategy” designed to address the demands of data processing in the context of international genetic databases. Other aspects should also be considered. The governance tools adopted are of key importance when considering whether a broad initial consent is appropriate. In this context, other mechanisms exist to ensure participants’ control over their data (see 3 E. The ICGC controlled-access database model). This is especially important where the eventual recipients of the data are in other countries and sometimes their identities unknown from the participants.

As noted at 3E, all access to controlled-access data must conform to the ICGC policies and procedures and be approved and recorded by the DACO, which will provide an auditable tracking of data transfers. ICGC mechanisms also include monitoring systems and registries that inform the public of the use of data by researchers (with lay summaries), and finally, de-identification techniques.

2.6. Participant’s ability to object

The inclusion of data obtained by researchers in the database will generally be a precondition of the research protocol and will be a requirement for participants when seeking their consent. Participants have to be informed about this circumstance as a basic policy of the research. If participants do not want their data stored in a national or international database, participation is likely to be declined.

---


2.7. Dissemination of research results

Research ethics’ principles recognise that outcomes of research should be accessible to participants and that participants should receive general information about the results of the research. They should normally be published and disseminated to contribute to the advancement of public knowledge. The question arises whether the results from secondary research, such as results obtained by researchers who accessed the database of the international project, should be given to participants. Return of individual research results remains a debatable issue, but if the results have clinical significance and utility, should participants not be notified? The ICGC has taken the position that policies on returning individual results should be decided at the local level when national and institutional rules are in place.

2.8. Ability to withdraw

The ability to withdraw from research is an essential and established principle of research ethics and medical law. It is a tool through which participants can exercise their right of control over their data. No justification is required for a decision to withdraw. Withdrawal means that the data will no longer either be used or be accessible unless it is maintained for legal purposes or has already been published / distributed and cannot be withdrawn. That said, the withdrawal of data does not necessarily imply their destruction, especially if already published or part of a dataset. In addition, the ability to withdraw may be limited by rules that require the conservation of information for different reasons.

In the context of biomedical research, individuals bring a wealth of personal information to a given project. On withdrawal, Article 9 of the

---

27 See for example, in Australia, National Statement on Ethical Conduct in Human Research 2007, Sec 1.5 “Research outcomes should be made accessible to research participants”. The UK Biobank will not provide “participants with information, genetic or otherwise, derived from examination of the database or samples by research undertaken after enrolment” Ethics and Governance Framework at 8.

28 Australia, National Statement on Ethical Conduct in Human Research, Sec 1.3(d) “disseminating and communicating, whether favourable or unfavourable, in ways which permit scrutiny and contribute to public knowledge”.

29 The ICGC is developing a further study about this issue and conducting an anonymous, voluntary survey to understand researchers’ experiences with finding and/or returning individual research results to research subjects participating in cancer genomics research. Available at: https://uncodum.qualtrics.com/SE/?SID=SV_bg75iepIrVLmScQ. [Accessed 2 Dec 2011].
UNESCO International Declaration on Genetic Data provides, “When a person withdraws consent, the person’s genetic data, proteomic data and biological samples should no longer be used unless they are irretrievably unlinked from the person in question.” However, as a result of the research, other data are obtained in the form of research findings. These findings may be used and published despite the withdrawal of an individual. This further use after withdrawal must be anonymised.

2.9. Ethical and legal issues of international transfer of genetic data. Concern in the ICGC

The ICGC recognizes that the above issues are basic requirements for the ethical conduct of research. It also recognizes that its member projects are required by their own national guidelines to conduct research in accordance with national laws and regulations, and with internationally accepted ethical standards. In order for the consortium to function, The Executive Committee of the ICGC has agreed that certain Core Bioethical Elements must be respected by all members as a precondition of membership. In addition to these Bioethical Elements, there are policies and guidelines that ICGC-member projects should consider in matters related to consent. The ICGC has been involved in defining minimum consent standards through the Ethics and Policy Committee (EPC).

Consistent with established international standards for research, the consent process emphasizes a participant’s informed and voluntary consent to research. The ICGC has made it a policy that participation is a member project is voluntary, that participants are informed that this is a cancer research project, and that they can withdraw at any time. Patients are informed that their care will not be affected by any decision they make regarding participation. Confidentiality and privacy are also key; participants are informed that their data will be placed in open and

\[\text{\underline{\text{\textsuperscript{[30]}} Art. 8b.}}\]

\[\text{\underline{\text{\textsuperscript{[31]}} Information on, inter alia, risks and benefit, if any.; samples and data to be collected and stored; nature of the research; research projects for the future; whether data from health-records required; ICGC policy on sharing samples and data with other research organizations; privacy security and confidentiality.; anonymisation; feedback of research results; data storage in national and international databases; right to withdraw; policy on benefit-sharing, IP, commercialization.}}\]
controlled data databases, depending on the information (see section 3) and that steps will be taken to protect the confidentiality of that information.

But the ICGC also recognizes that there are differences between projects and these can be reflected in local policies. The ICGC needs to be flexible and has provided guidelines in certain areas. For example, the ICGC has taken a position that it will not, as a consortium, return research results to individuals. However, it is agreed that member projects may have local requirements. Therefore, the ICGC has a guideline that states,

Provided it is agreed at recruitment, if clinically important and validated findings emerge during the initial recruitment and screening phase, or in the early research, attempts will be made to pass this information back via the clinician, by whatever mechanism may be agreed at the local level.

3. International Cancer Genome Consortium (ICGC): policy governing management and transfer of data

In order to fulfill the ambitious scientific objectives of the ICGC, mechanisms have been developed to enable researchers outside the core groups creating the data to have access to scientific results. This permits greater research efforts, wider dissemination of the information to the entire scientific community and maximises the opportunity for discoveries to benefit the community.

It should be emphasized that each project using ICGC genetic databases operates at two different governance levels: the national level, where results may be held locally, and the international level where data is stored in a federated genetic database and shared with researchers around the world.

3.1. National institutions

Each participating research team operates within a research institution that will have established governance and ethical standards for the creation and use of a genetic database. Institutional level standards will comply with national legislation and national ethics codes, guidelines and policies. Many managers and directors of databases have devel-
oped their own database guidelines that supplement national guidelines.  

Each institution will have access to a Research Ethics Board (REB) or a similar committee to review research project applications, within the traditional REB role of protection of the interests of research participants, to make sure that the use of the data is consistent with the original purpose they were established for and to review the quality of the aimed research.

3.2. The national research team and data controllers

The research team must inform patients and participants about all aspects of the research and their participation in a project, including the fact that their project is part of an ICGC collaborative project. The research team is responsible for obtaining REB approval and participants’ consent.

The person who creates the project database is the data controller, who is any person, public administration, body, association or other entity that is competent to determine the purposes and methods for the processing of personal data. This person has responsibility for ensuring they have the relevant means to maintain the database including security matters (such as processes for coding personal data). Coding requires removing personal identifiers and separating them from the data and keeping them secure. The non-identifying data is given a code, which links it back to its personal identifiers. Only the coded data is transferred to other researchers.

__________


34 Article 2 d) Directive 95/46/EC of the European Parliament and of the Council of 24 October 1995 on the protection of individuals with regard to the processing of personal data and on the free movement of such data.
The determination of the nature of the data (whether personal or not) is essential to the implementation of principles of personal data protection. The definition of an “identifiable subject” is also a key point. The tendency is to consider that codified data are not personal if the user or the data controller has no access to the code. But security measures have to be put into place so the possibility of re-identification would be unreasonable for a third party that has access to the data (see 2.B Personal data, above).

The data processor is distinct from the data controller and is any person, public administrative body, association or other agency that processes personal data on the controller’s behalf. The legal relationship between the data controller and the data processor requires the data processor to follow the instructions of the data controller. This will be specified in a contractual agreement.

3.3. International persons or institutions

At the international level, data shall not be considered personal data if in the particular context they cannot be linked to the subject by reasonable means (see 2.B above). The controller of an international database, such as that of the ICGC, coordinates input from the national databases (it is a database of federated databases). All researchers (i.e. depositors, managers, users) must agree to abide by the terms set by the international database concerning data management.

3.4. International data transfer

The integration of data from the members’ national projects for the purposes of research use is a major step towards the dissemination of the project’s scientific benefit. Therefore, one of the first aims of international research projects like ICGC is to create an international database containing the results of the various national projects.

35 Article 2 e) Directive 95/46/EC.
36 For further details of the ICGC model, visit: http://www.icgc.org/icgc/goals-structure-policies-guidelines/e9-data-management
International data transfer is the movement of identifiable personal data that are transmitted from a national to an international database\textsuperscript{37}. Special processes may be required; for example, in the European legal framework, there are specific provisions for data transferred outside the European Union\textsuperscript{38}. In the case of international transfers, the \textit{Data Exporter} is the person or entity that transfers data to a third country and the \textit{Data Importer} is the person or entity receiving the data (whether the Controller, the Data Processor or an authorised third party)\textsuperscript{39}.

The integration of the data for later research use is the first step towards the dissemination of the project’s scientific outcomes. To provide data with additional safeguards, ICGC data have been split into two categories, \textit{open} and \textit{controlled-access data}. The first includes only non-identifiable data. Special measures and attention have been paid to the latter data category (controlled access data).

\textbf{3.5. Oversight: the ICGC controlled-access database model}

The ICGC national projects will develop their own project-specific data models, workflows, and databases. Each data producer at the national level manages its own workflow and is responsible for primary quality control, data integrity and the protection of confidential information. A common core of ICGC data intended for integration into the international database and redistribution is organised within designated “franchise databases” which share a common data model and struc-


\textsuperscript{38} Prior to transfer, the equivalency of data protection must be verified. The EU has published a list of countries which provide the requisite safeguards (all the Commission decisions on the adequacy of the protection of personal data in third countries are in the Commission web: http://ec.europa.eu/justice/policies/privacy/thridcountries/index_en.htm). For EU members, the National Database Registry must be informed of all data transfers (article 19 e9 of the Directive 95/46/EC). Reports must be submitted by each researcher regarding how the transfer will take place and how the method complies with national legislation.

\textsuperscript{39} According to Data Protection Directive 95/46/EC,
The franchise database software (schema, integrity-checking utilities, and load and dump utilities) are written by the data coordination centre (DCC) and managed by the data producers.

At regular intervals, a subset of the information contained in the project-specific databases will be exported to a local ICGC franchise database, which implements a uniform simplified data model that captures the essential data elements needed to implement ICGC-wide policies on data release, quality control and milestones. The franchise also includes a set of standardized validation and quality control tools developed and deployed by the DCC to verify that the information placed in the franchise database is complete and internally consistent.

In order to provide the research community with a single portal to ICGC data, a coordination backend database acts as the amalgamation of all the franchise databases. As a result, from the user’s perspective, all the ICGC data are presented in one place. All requests for transfers of controlled access data to third parties (i.e. to “secondary” researchers), under the ICGC policies, must be approved by the Data Access Compliance Office (DACO). These exchanges between ICGC and collaborating researchers and their institutions are recorded to provide a continuous “chain of responsibility”\textsuperscript{41} and an auditable tracking of the data transfer. Secondary researchers must apply for the access through the consortium website and sign an agreement including, for example, the conditions of the use of data, security measures, an obligation of confidentiality, and a publication moratorium.

\textsuperscript{40} The Franchise Database is a database containing data from National Databases which implements a uniform simplified data model that captures the essential data elements needed to implement ICGC-wide policies.

4. Conclusions

The scientific value of international genetic databases has been recognized by the scientific community. Identification of the legal regime applicable to the data (personal or not) included in the database is a key step in order to determine the operations that should be implemented concerning these data.

Although from a legal perspective there are still some doubts concerning the concrete rules governing an international database, the ICGC has developed policies and procedures to protect participants taking into account some international adopted general principles, while facilitating the ethical conduct of this collaborative research. The ICGC controlled-access database model proposes a useful solution to reconcile all the interests involved by ensuring that the data stored in the central database cannot be traced back to an individual participant without procedural checks. The participants give their consent after being provided general information of the policy, procedures and purposes of ICGC.
The structure implemented is on the one hand useful and operative and, on the other hand, fits with those national systems that consider a broad consent (including future data transfers) as the most adequate tool to guarantee the rights of research participants, as well as with those that consider that coded data may not be considered personal data if some guarantees have been taken to avoid the identification of the participant (as a matter of fact there has been a favourable evaluation of the ICGC model by the Research Ethics Boards of all 14 countries in the Consortium.

Finally, it should be underlined that facing the uncertain international legal framework, a transnational ethical assessment and policy, integrated in each international consortium or project in the field of human genetic sequencing, is a indispensable requisite. This assessment and policy should be the way to find the more adequate solutions in each particular case to the eventual conflicts, and will facilitate the national positive evaluation and/ or authorisation of the research.