

## Genomic Databases and International Collaboration

Bartha Maria Knoppers, Ma'n H Abdul-Rahman and Karine Bédard\*

*Recent years have witnessed a key development within biomedicine—namely, the move from genetic to genomic research. Genomic research, which operates at the level of the whole genome rather than individual genes, requires a powerful new set of research tools, resources and supporting technologies. Having moved from DNA sequence mapping to the use of haplotypes, the next advances in our understanding of disease risk and health may well be achieved through the study of 'normal' genomic variation across whole populations. Such studies require not only samples and data, but also highly sophisticated, substantial database infrastructures to support them. Longitudinal and largely epidemiological in nature, these population-scale genomic database resources are designed to serve a multiplicity of specific research projects at both national and international levels. Current ethical guidance in the area of genetic research promotes the need for international collaboration. Yet, is international genomic research collaboration possible considering both the scientific and structural differences between national approaches to governing genomic databases and associated population biobanks? A review of existing norms at the international level—particularly with regard to benefit sharing and access to data—and their application in different countries, reveals areas of both convergence and divergence. But, most of all, it reveals the need for international harmonisation in order to secure interoperability and the public participation, trust and investment in such large initiatives that are crucial to their success.*

### A. INTRODUCTION

Recent years have witnessed a key development within biomedicine—namely, the move from *genetic* to *genomic* research. Traditionally, genetic researchers studying inherited

\* Bartha Maria Knoppers is Professor of Law, Faculté de Droit, and Canada Research Chair in Law and Medicine, Centre de Recherche en Droit Public (CRDP), Université de Montréal. Ma'n H Abdul-Rahman is a Research Assistant, and Karine Bédard a former Research Associate, at CRDP, Université de Montréal. Funds for the preparation of this article were obtained from the EU-funded PHOEBE project (Promoting Harmonization of Epidemiological Biobanks in Europe) and the Genomics and Public Health (GPH): Building Public 'Goods'? Project (Genome Canada and Génome Québec). We would like to thank Dr Jane Kaye and Dr Sue Gibbons for their comments on this paper, and Dr Sue Gibbons in particular for her excellent editorial skills.

disorders used a range of established tools and techniques, such as DNA sequence mapping, to identify and examine the genes thought to be implicated.<sup>1</sup> But while such tools and techniques have been pivotal for unlocking the secrets of monogenic diseases, the vast majority of the human disease burden results from common complex, multi-factorial disorders. These include cardiovascular diseases, diabetes and many cancers. Scientists generally believe that such common diseases result from complex interactions among three major elements: genetic variation, individual lifestyle behaviours, and environmental factors.

There is growing recognition amongst scientists that, in order fully to understand the complexities of the common disease risk and human health, it is necessary for studies of 'normal' genomic variation to be carried out across whole populations.<sup>2</sup> Rather than investigating single genes using traditional techniques, these studies must operate at the level of the whole genome (all of an individual's genes taken collectively). Moving up to the genomic level is necessary to facilitate systematic investigation into the complex patterns of multi-factorial interaction. Such studies therefore require a powerful new set of research tools, resources and supporting technologies. Over recent years, much progress has been made in mapping 'normal' similarities and differences within the human genome. Scientists have used genetic markers, such as single nucleotide polymorphisms (SNPs) and haplotypes (groups of SNPs that are commonly inherited together), to identify potentially significant genetic variations across whole genomes. This has enabled them to begin to investigate the associations between SNPs, haplotypes and the incidence of disease.

Having already progressed from DNA sequence mapping to the use of SNPs and haplotypes, the next advances in our understanding of disease risk and health may well be achieved through the study of 'normal' genomic variation across whole populations. But such studies require extremely large collections of biosamples and data to enable and support both longitudinal and epidemiological studies. Moreover, they require not only biosamples and data (including relevant medical history, genealogical, lifestyle and environmental information about participants), but also highly sophisticated, substantial database infrastructures to support them. Over the last decade, there has been considerable investment in what are often termed 'population biobanks',<sup>3</sup> as well as in large collaborative projects, and in the adding of genotypic data to well-established cohort

<sup>1</sup> For a succinct account of the move from genetic to genomic research see Susan MC Gibbons *et al*, 'Governing Genetic Databases: Challenges Facing Research Regulation and Practice' (2007) 34 *Journal of Law and Society* 163, 165–7.

<sup>2</sup> Muin J Khoury, 'The Case for a Global Human Genome Epidemiology Initiative' (2004) 36(10) *Nature Genetics* 1027.

<sup>3</sup> See, eg, Anne Cambon-Thomsen *et al*, 'An Empirical Survey on Biobanking of Human Genetic Material and Data in Six EU Countries' in Bartha M Knoppers (ed), *Populations and Genetics: Legal and Socio-Ethical Perspectives* (Martinus Nijhoff, Leiden 2003) 141; Nicole Palmour, 'A Survey of the Variability of DNA Banks Worldwide' in Knoppers, *ibid*, 123.

studies.<sup>4</sup> The development of such resources is a slow and painstaking process. It involves long-term commitment and substantial investments of time, money and expertise. No one study or resource will have the necessary size—and, therefore, statistical power—fully to address all of the complexities of the common disease risk by itself. Even if such a study ever were to exist, the results would not be realised for many years.

One solution to the problem of achieving the requisite scale of research materials needed for such initiatives is to establish frameworks, standards and norms by which existing or new database resources can be networked together. Ideally, such frameworks, standards and norms would enable datasets to be networked—and data and, where necessary, biosamples, to be transferred or shared—at both the national and international levels. Interoperability and international collaborative access to, and use of, the genomic data so created will be essential to the goal of achieving the rapid translation of research results into clinical knowledge and new therapies.

However, a major obstacle to achieving interoperability, international collaboration and database networking is the fact that few socio-ethical or legal norms exist at a global level to guide these endeavours. Moreover, there are many differences in the legal requirements that apply at the national level, which also militate against networking. Another factor against networking is the fact that there are significant differences in the way that scientific studies are designed, organised, regulated and carried out, even within individual countries. The purpose of this paper is to review the international and national frameworks that have been developed to date, to establish if, and how, those frameworks might be applied to facilitate the networking of population biobanks. Section B begins by mapping out the current international framework, such as it is, relevant to genomic research and population biobanks. It outlines the principal bodies and documents (international, regional and national), and considers their respective status, authority, scope, binding nature and enforceability. Section C examines the potential application of international norms, by focusing on two issues that are crucial to international collaboration: benefit sharing and determining access to genomic databases. Finally, section D discusses key roadblocks to interoperability, and identifies and assesses various current initiatives aimed at promoting international collaboration.

## B. THE INTERNATIONAL FRAMEWORK

Most of the legally binding laws and regulatory bodies that exist for governing genomic research, population biobanks and genetic databases operate at the national level. As such, they do not possess any supra-national status, authority or enforceability. However, over

<sup>4</sup> In the UK, see, eg, the Avon Longitudinal Study of Parents and Children (also known as the 'Children of the 90s' study and ALSPAC): <http://www.alspac.bristol.ac.uk>.

the last 10 years, an increasing number of international bodies have developed relevant guidelines or statements of principle. Broadly speaking, those bodies can be grouped into four different types. There are: (1) bodies that are representative of all countries, such as the United Nations and its specialised agency UNESCO;<sup>5</sup> (2) the Council of Europe, a regional body that represents countries within Europe as well as other countries that are prepared to sign up to its conventions;<sup>6</sup> (3) international scientific organisations, notably the Human Genome Organisation (HUGO);<sup>7</sup> and (4) bodies that represent the industrialised nations, such as the OECD.<sup>8</sup>

Yet, none of these bodies (or, indeed, any other) has been specifically instituted to regulate or oversee population biobanks. Accordingly, no single actor at the international level possesses a clear mandate or authority to formulate or promulgate a global consensus position regarding population biobanking norms and standards, or to oversee the governance of large-scale, international collaborative genomic research. As well as there being no international *body* responsible for biobanking activities, no legally binding, international legal *instrument* applies specifically to biobanks either. Instead, what we currently have are multiple bodies, issuing various different kinds of documents, many of which do not concern biobanking specifically but merely address related or more general principles or activities, whose status, authority, content, definitions and enforceability mechanisms all differ. In sum, this leaves us with a piecemeal and incomplete international framework.

Turning to consider the four categories of international bodies, and looking first at UNESCO, its Universal Declaration on the Human Genome and Human Rights<sup>9</sup> of 1997 is one of several international instruments adopted over the past decade relating to the issues surrounding genetic testing and research. Prospective in nature, it outlines the basic ethical principles for the proper conduct of human genome research generally. Thus, it does not address biobanking specifically. Furthermore, while endorsed by the UN,<sup>10</sup> the Declaration is not legally binding. Indeed, in the case of both UNESCO and the Council of Europe, the declarations, conventions and treaties that they produce can only take effect once they have been signed and then implemented into national law by each country. As is well known, in order to become legally binding, international instruments such as treaties and conventions must be ratified by the countries that participate in their formulation, observance and enforcement. Consequently, the coverage and acceptance of the principles set out in the relevant documents produced by UNESCO and the

<sup>5</sup> United Nations Economic, Social and Cultural Organisation: <http://www.unesco.org>.

<sup>6</sup> See <http://www.coe.int>.

<sup>7</sup> See <http://www.hugo-international.org>.

<sup>8</sup> Organisation for Economic Co-operation and Development: <http://www.oecd.org>.

<sup>9</sup> UNESCO, Universal Declaration on the Human Genome and Human Rights (adopted 11 November 1997). Available at <http://unesdoc.unesco.org/images/0011/001102/110220e.pdf#page=47>.

<sup>10</sup> UN GA/RES/53/152 of 9 December 1998.

Council of Europe are somewhat piecemeal. They cannot be considered to provide a comprehensive regulatory system. Having said this, while, strictly speaking, UNESCO's Universal Declaration on the Human Genome and Human Rights is merely hortatory and proclamatory in nature, as Francioni has observed, it reflects<sup>11</sup>

... emerging principles of international law which, though expressed in the soft-law form of the Declaration, are designed to model the evolution of customary law and to eventually harden into more detailed and exacting standards. In any event, it is difficult to deny that the Declaration has already affected the *opinio iuris* of the international community.

Such influence is perhaps unsurprising given UNESCO's comparatively high profile among the four types of international actors. As a body embracing over 190 countries, statements and principles endorsed by the UN may be expected to attract serious international attention. In 2003, UNESCO provided more specific guidance in its International Declaration on Human Genetic Data.<sup>12</sup> Again, however, this non-binding declaration does not cover genomic databases *per se*; nor does it address the particular issues associated with population biobanks.

Turning to the Council of Europe, its Recommendation on Research on Biological Materials of Human Origin of 2006<sup>13</sup> was the first supra-national document to address population biobanks and associated data specifically. However, because it is merely a recommendation, it cannot acquire binding force. Moreover, it is a regional instrument only—reflecting the Council of Europe's limited status and authority as a regional organisation. In 1997, the Council of Europe adopted the Convention on Human Rights and Biomedicine.<sup>14</sup> That Convention (together with its Additional Protocol) covers the broader perspectives of the human rights implications of the applications of biology and medicine. As with Council of Europe conventions generally, countries that become signatories to the Convention must ratify it by introducing implementing legislation to bring their national laws into conformity with its principles. Thus, it can take on a binding and enforceable effect, at least at the national level within a particular group of States. However, several Member States (including the UK) have not signed up to the Convention; in part, due to its restrictions over stem cell research. This illustrates another

<sup>11</sup> Francesco Francioni, 'Genetic Resources, Biotechnology and Human Rights: The International Legal Framework' (European University Institute Working Papers, EUI LAW No 2006/17) 8. Available at <http://cadmus.iue.it/dspace/bitstream/1814/6070/1/LAW200617.pdf>.

<sup>12</sup> UNESCO, International Declaration on Human Genetic Data (adopted 16 October 2003). Available at <http://unesdoc.unesco.org/images/0013/001331/133171e.pdf#page=45>.

<sup>13</sup> Council of Europe, Recommendation Rec(2006)4 of the Committee of Ministers to Member States on Research on Biological Materials of Human Origin (adopted 15 March 2006). Available at <https://wcd.coe.int/ViewDoc.jsp?id=977859&BackColorInternet=9999CC&BackColorIntranet=FFBB55&BackColorLogged=FFAC75>.

<sup>14</sup> Council of Europe, 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 (Oviedo, 4 April 1997, ETS 164).

of the shortcomings we face when trying to rely on international instruments that cover a range of different areas to govern specific enterprises such as genetic and genomic databases—namely, the risk that disagreement over one unrelated aspect or set of provisions will lead to a rejection of the document as a whole.

The Council of Europe's Recommendation on Research on Biological Materials of Human Origin is a good illustration of another key problem that can result from having multiple bodies issuing partially overlapping documents in an unco-ordinated fashion: the problem of inconsistency or contradiction. As noted above, in the context of research involving whole populations, the databases created are often referred to as 'population biobanks'. Yet, this terminology is by no means universally settled. As well as using inconsistent terminology, current international documents also define population databases in quite different ways. For example, the Council of Europe's Recommendation defines 'population biobanks' as being collections of biological materials having the following characteristics:<sup>15</sup>

- i. the collection has a population basis;
- ii. it is established, or has been converted, to supply biological materials or data derived therefrom for *multiple future research projects*;
- iii. it contains biological materials and associated personal data, which may include or be linked to genealogical, medical and lifestyle data and which may be regularly updated;
- iv. it receives and supplies materials in an organised manner.

By contrast, in its Statement on Human Genomic Databases of 2002, the HUGO Ethics Committee defined a 'genomic database' as being simply 'a collection of data arranged in a systematic way so as to be searchable'.<sup>16</sup> Even limiting ourselves to international organisations that have guided the specific domain of genomic research since 1997, it is immediately evident from such discrepancies that there is much confusion over the very nature of population research. Drawing on the issues surrounding traditional genetic testing for monogenic conditions, and fears about possible discrimination against individuals—to say nothing of eugenics—many international documents reflect either a total misunderstanding of the very specific and often basic epidemiological nature of population genomics, or suggest norms that are 'hybrid' in nature, thus creating confusion as to the actual risks and benefits specific to population studies. This is most unfortunate since, as already mentioned, most genomic database projects envision the creation of *infrastructures*—that is, of scientific *resources*—rather than the study of a specific genetic condition or drug.<sup>17</sup>

<sup>15</sup> Council of Europe (n 13), art 17 (emphasis added).

<sup>16</sup> HUGO Ethics Committee, *Statement on Human Genomic Databases* (adopted December 2002). Available at <http://www.hugo-international.org/PDFs/Statement%20on%20Human%20Genomic%20Databases%202002.pdf>.

<sup>17</sup> Bartha M Knoppers and Alastair Kent, 'Ethics Watch: Policy Barriers in Coherent Population-based Research' (2006) 7 *Nature Reviews Genetics* 8.

In the search for, or construction of, international norms, it is also imperative to distinguish population genomic biobanks and their associated databases from residual tissue collections<sup>18</sup> or biosamples collected during the course of clinical trials. Again, each kind of collection has its own particular normative profile. Having said this, helpful lessons can be learned from the approach taken to governing residual tissues, such as stored tumour tissues for example. Recently, European cancer researchers involved in TuBaFrost<sup>19</sup> (a central European database of information derived from frozen tumour samples stored in numerous different countries) proposed a Code of Conduct for the use of residual tissue for research. That code supports the ‘co-ordinating principle’, whereby ‘the regulations of the country where the tissue was taken from the patient and was stored decide whether the tissue may be used in another country with possibly different regulations’.<sup>20</sup> Moreover, the same would apply to data; although, in the case of population biobanks, all data would be coded and so not be identifiable by the researcher seeking access.<sup>21</sup>

To give just one more illustration of inconsistency before moving on, across the board the international documents use terms to describe the required confidentiality mechanisms for protecting genetic data that are confusing and contradictory<sup>22</sup>—although international clarification may soon be forthcoming.<sup>23</sup> In part, this can be explained by the fear surrounding the possible misuse of genetic data. But the underlying causes may well be traceable to a deeper uncertainty surrounding the concepts of ‘anonymisation’ and ‘identifiability’ themselves, as well as the tendency towards genetic exceptionalism. With regard to anonymisation and individual identifiability, subject to laws and guidelines, the dual requirements of reasonable manpower and practicality provide the degree of data security necessary to ‘anonymise’ individual biosamples and data sufficiently. Fully anonymised data is no longer subject to data protection legislation as the data can no longer be linked to an individual. Genetic exceptionalism, however, further complicates this issue by fostering a distinction between genetic and other medical or personal data, leading to a requirement of heightened security mechanisms for genetic

18 Evert-Ben van Veen *et al*, ‘TuBaFrost 3: Regulatory and Ethical Issues on the Exchange of Residual Tissue for Research Across Europe’ (2006) 42 *European Journal of Cancer* 2914.

19 See <http://www.tubafrost.org>.

20 van Veen *et al* (n 18) 2920.

21 *Ibid*, 2921.

22 Bartha M Knoppers and Madeleine Saginur, ‘The Babel of Genetic Data Terminology’ (2005) 23 *Nature Biotechnology* 925.

23 International Conference on Harmonization of Technical Requirements for Registration of Pharmaceuticals for Human Use, ‘Draft Consensus Guideline: Terminology in Pharmacogenomics E15’ (Current Step 2 version, 25 October 2006). Available at [http://www.hc-sc.gc.ca/dhp-mps/alt\\_formats/hpfb-dgpsa/pdf/prodpharma/e15\\_step2\\_etape2\\_e.pdf](http://www.hc-sc.gc.ca/dhp-mps/alt_formats/hpfb-dgpsa/pdf/prodpharma/e15_step2_etape2_e.pdf).

data. Obviously, in large population studies it is necessary to follow participants over time. So then, data necessarily are coded.<sup>24</sup>

These and many other issues surrounding genomic data also have been discussed by various scientific organisations, at both the international and national levels. A leading example representing this third type of international actor is HUGO. For example, the HUGO Ethics Committee's Statement on Human Genomic Databases, discussed above, maintains that such databases are 'global public goods'.<sup>25</sup> International bodies like HUGO have a significant influence over the scientific community. By declaring that such a principle forms the basis for its recommendations, the HUGO Statement sends out a powerful and influential normative signal—at least, within the scientific community—that human genomic databases should be seen as public resources; that all humans should share in and have access to their benefits; and that the free flow of, access to, and exchange of data are essential.<sup>26</sup> The HUGO guidelines are not, however, binding on countries.

Nor, indeed, are guidelines issued by the fourth category of international actors—namely, bodies that represent the industrialised nations, such as the OECD. The OECD has issued numerous publications, guidelines, statistics and working papers in the areas of biotechnology, health, science, innovation, transborder personal data protection, and access to digital research data from publicly funded projects.<sup>27</sup> Notably, in 2006 the OECD published a report detailing the findings of a working party convened specifically to examine the management and governance of human genetic research databases. That report explores numerous options and possible ways forward.<sup>28</sup> While the OECD's guidance documents and reports lack enforceability, and its formal membership is limited in scope,<sup>29</sup> the OECD's role as a forum for reaching multilateral policy agreement and achieving national and international policy co-ordination should not be underestimated. Through mechanisms such as intergovernmental peer pressure and issuing 'soft law',<sup>30</sup> the OECD can perform a potentially powerful, global norm-setting role. Occasionally, its soft-law instruments do result in hard law in the form of binding treaties.

Overall, then, it is clear that, to achieve interoperability and collaboration, it is at the international, multilateral level that a sound and comprehensive legal framework is needed for governing genomic databases. Not least, this is because the networking of population biobanks and other databases will be carried out at a global level, crossing

<sup>24</sup> For example, CARTaGENE uses multiple codes for each institutional partner and separate codes for each research project seeking to access the database. CARTaGENE itself cannot identify participants. See <http://www.cartagene.qc.ca>.

<sup>25</sup> HUGO Ethics Committee (n 16), recommendation 1.

<sup>26</sup> *Ibid*, recommendations 2 and 3.

<sup>27</sup> See, eg, OECD draft guidelines, 'Principles and Guidelines for Access to Digital Research Data from Public Funding: Promoting International Co-operation in the Use of Scientific Data Resources' (December 2005).

<sup>28</sup> OECD, 'Creation and Governance of Human Genetic Research Databases' (OECD Publishing, Paris 2006).

<sup>29</sup> The OECD has 30 member countries. But it also has active relationships with many more countries, as well as non-governmental organisations.

<sup>30</sup> Within the OECD, 'soft law' refers to recommendations for action by national governments.

national boundaries. Yet from the foregoing analysis it is obvious that, as the OECD's report noted above puts it,<sup>31</sup>

there is currently no international, comprehensive framework setting forth global consensus on the issues of ownership, commercialization, exclusive licensing, access for researchers, benefit sharing and other issues, as these pertain to population databases.

As a solution to this legal vacuum at the international level, some countries—for example, Iceland and Estonia—have adopted national, legally binding and purpose-specific legislation to govern their population database projects. Estonia's Human Genes Research Act 2000<sup>32</sup> contains provisions covering a variety of issues such as data protection, prohibition against discrimination, the right of ownership of tissue samples and oversight. Likewise, Iceland's Act on a Health Sector Database, No 139/1998,<sup>33</sup> together with its supporting Regulation,<sup>34</sup> protects confidentiality, access to data from health records, rights of patients, transfer of medical data and intellectual property. But such national approaches to governing the creation and maintenance of genomic databases are far from uniform. In particular, they offer a variety of positions—or no position at all—on issues such as data protection, access to and release of database materials, and possible commercialisation. Nor do existing national legal instruments cover the issues associated with the *networking* of biobanks. Furthermore, they are only enforceable in the national jurisdiction. They too, then, have limited value as a tool at an international level.

Some national funding bodies—notably the National Institutes of Health in the USA—wield considerable influence over genomic research practice. This is because their policies have an enormous effect over the way in which research in this area is carried out. Thus, one example drawn from the USA is the National Cancer Institute's First Generation Guidelines for NCI-Supported Biorepositories.<sup>35</sup> This document serves as an illustration of a national disease-specific initiative attempting to formulate common guidelines and standards. But, as the OECD report discussed above points out, while such 'policies may play a significant role in shaping the debate and offering guidance they do not provide a complete global framework.'<sup>36</sup> In part, this is because the influence of funding bodies is informal rather than strictly legally binding. It is also largely confined in scope to the scientific projects and initiatives that they fund, rather than being universally applicable.

<sup>31</sup> OECD (n 28) 61.

<sup>32</sup> *Inimgeeniuringute seadus*, RT I 2000, 104, 685.

<sup>33</sup> *Lög um gagnagrunn á heilbrigðissviði*.

<sup>34</sup> Regulation on a Health Sector Database, No 32/2000 (*Reglugerð um gagnagrunn á heilbrigðissviði*).

<sup>35</sup> National Cancer Institute, *First Generation Guidelines for NCI-Supported Biorepositories* (April 2006). Available at <http://biospecimens.cancer.gov/biorepositories/First%20Generation%20Guidelines%20042006.pdf>.

<sup>36</sup> OECD (n 28) 61.

Despite the obvious problems of enforceability and applicability that they face, international organisations are still keen to promote the networking and sharing of resources worldwide. Thus, UNESCO's Universal Declaration on the Human Genome and Human Rights upholds the need for States 'to continue fostering the international dissemination of scientific knowledge concerning the human genome, human diversity and genetic research and, in that regard, to foster scientific and cultural co-operation, particularly between industrialized and developing countries'.<sup>37</sup> This was reinforced by UNESCO's International Declaration on Human Genetic Data, which states that researchers 'should encourage the free circulation of human genetic data ... in order to foster the sharing of scientific knowledge'.<sup>38</sup> Perhaps the most explicit, comprehensive and clear guidance on the issue of genomic databases is to be found in the Council of Europe's Recommendation on Research on Biological Materials of Human Origin. It maintains that '[m]ember states should take appropriate measures to facilitate access by researchers to biological materials and associated data stored in population biobanks'.<sup>39</sup> In the same vein, the OECD's report on human genetic research databases notes that 'increased collaboration seems inevitable and desirable'.<sup>40</sup>

Collaboration using population genomic databases will, however, be largely dependent on their ability to harmonise in many crucial areas—not least with regard to normative principles, oversight, governance mechanisms, technical and security standards, and scientific practices.<sup>41</sup> Whether such harmonisation will ever be forthcoming depends, in turn, not only on agreeing an appropriate set of international norms relevant to the creation and use of genomic databases, but also on the degree of influence that such norms have on national approaches. Therefore, the application of international norms is the next issue that we must examine.

### C. THE APPLICATION OF INTERNATIONAL NORMS

For the purposes of examining trends in the application of international norms to date, it is helpful to focus the analysis on specific issues about which there is some degree of general normative agreement. Two crucial issues for international collaboration in genomic research in respect of which international norms have emerged are: (1) how benefits should be shared; and (2) how access to population biobank resources and materials should be determined and supervised.

<sup>37</sup> UNESCO (n 9), art 18.

<sup>38</sup> UNESCO (n 12), art 18(c).

<sup>39</sup> Council of Europe (n 13), art 20(1).

<sup>40</sup> OECD (n 28) 134.

<sup>41</sup> See below, section D.

There is an obvious tension between these two issues, even though data sharing is a form of benefit sharing. This is because, for certain stakeholders at least, there may well be potential financial or other benefits to be made—for example, through commercialisation of research results, or enjoying exclusive rights to control or exploit database resources—that conflict with open access policies and anti-proprietary principles. Yet, at the same time, it is obvious that the scientific value and usefulness of genomic databases to researchers or companies wishing to access them would be greatly increased if, at a minimum, certain biochemical measurements collected and key data were to be entered in such a way as to acquire the meaningful, statistical significance necessary for eventual clinical utility. As the OECD has observed, '[t]he main challenge in the discussion of access to genetic research databases is to strike an appropriate balance between the freedom of researchers and the interests of the participants and the public'.<sup>42</sup>

Benefit sharing is an international norm. It seeks to offset or to mitigate the effects of State sovereignty over bioresources—or, at a lower level, of proprietary interests (or, at a minimum, personal rights of control) over biological tissues and the data derived therefrom.<sup>43</sup> According to UNESCO's Universal Declaration on the Human Genome and Human Rights, the human genome in its natural state 'shall not give rise to financial gains'.<sup>44</sup> Yet, eventual commercial exploitation of biosamples and the data derived therefrom, such as in the form of new tests, drugs, or other intellectual property, is not precluded by any international body. How, then, is it possible to safeguard benefit sharing, and ensure that population genomic databases are recognised and treated as the 'global public goods' that HUGO has declared them to be?<sup>45</sup> A closer examination of the concepts of benefit sharing and data access as applied at the national level reveals a hybrid approach. Under this approach, there is no State or individual 'sovereignty' over biosamples or data. But there is no open access either, in that researchers must respect limits over the use of biosamples and data dictated at the national level, and must return research findings and data derived from their use of the biosamples and data so as to enrich the population database.

Limiting ourselves to countries that either already have or that are building large genomic databases—such as Iceland (Icelandic Health Sector Database), Estonia (Estonian Genome Project Gene Bank), the UK (UK Biobank), the USA (National Cancer Institute Biorepository), Scotland (the Generation Scotland projects) and Canada (the CARTaGENE Project in Québec)—it is evident that a variety of approaches have been adopted, with respect to both benefit sharing and data access.

<sup>42</sup> OECD (n 28) 113.

<sup>43</sup> Lori Sheremeta and Bartha M Knoppers, 'Beyond the Rhetoric: Population Genetics and Benefit-Sharing' in Peter Phillips and Chica Onwuekwue (eds), *Accessing and Sharing the Benefits of the Genomics Revolution* (Springer Kluwer, in press).

<sup>44</sup> UNESCO (n 9), art 4.

<sup>45</sup> HUGO Ethics Committee (n 16), recommendation 1.

### 1. Benefit Sharing

Benefit sharing, as applied to human genetic research, has its origins in HUGO's 1996 Statement on the Principled Conduct of Genetic Research,<sup>46</sup> later developed more fully in another HUGO Statement on Benefit-Sharing in 2000.<sup>47</sup> This concept was incorporated by UNESCO into its 2003 International Declaration on Human Genetic Data, which holds that 'benefits resulting from the use of human genetic data, human proteomic data or biological samples collected for medical and scientific research should be shared with the society as a whole and the international community'.<sup>48</sup> Benefit sharing can take many forms. These include technology transfer, capacity building, and access to medical care or drugs given by way of recognition for those who have participated in research.

At a national level, this principle is not explicitly mentioned in any of the six countries under study, except for Canada within its CARTaGENE project.<sup>49</sup> Common to all six population biobanks, however, is the fact that potential participants are told that their involvement in the building of such infrastructures will not create or confer any property rights in their biosamples or data.<sup>50</sup> While this has become standard language in biomedical research, it was necessary to reassure scientists in both the public and private sectors and to be more transparent for participants. Most of these resources—while open to access requests from both commercial and non-commercial entities—will impose terms of use compatible with their general objectives. Yet, none has *excluded* the possibility that some research may lead to biomedical products that return a profit. To that end, both the fees charged for access and the obligation to return data to the resource in accordance with the terms of agreements signed with industry, academic or charitable organisations can function as sources of equitable re-investment in the databases.

In Iceland, neither the Act on a Health Sector Database, No 139/1998 nor its supporting Regulation establishes clear guidance on whether intellectual property rights associated with, or generated through using, the database belong to the licensee (deCODE Genetics, a private biopharmaceutical company), to the Icelandic Health Sector Database, or to both. Yet, the Act maintains that '[t]he licensee shall ensure that after the expiry of the period of the licence, the Minister of Health and Social Security, or the party assigned by the Minister to operate the database, shall receive indefinite use of all software and

<sup>46</sup> HUGO Ethics Committee, *Statement on the Principled Conduct of Genetic Research* (adopted 21 March 1996). Text published in (1995) 6 *Eubios Journal of Asian and International Bioethics* 59.

<sup>47</sup> HUGO Ethics Committee, *Statement on Benefit-Sharing* (adopted 9 April 2000). Available at [http://www.hugo-international.org/Statement\\_on\\_Benefit\\_Sharing.htm](http://www.hugo-international.org/Statement_on_Benefit_Sharing.htm).

<sup>48</sup> UNESCO (n 12), art 19(a).

<sup>49</sup> CARTaGENE, *Statement of Principles on the Ethical Conduct of Human Genetic Research Involving Populations* (2003). Available at [http://www.cartagene.qc.ca/docs/enonce2003\\_1.pdf](http://www.cartagene.qc.ca/docs/enonce2003_1.pdf).

<sup>50</sup> Amelia M Uranga *et al*, *Outstanding Legal and Ethical Issues on Biobanks: An Overview on the Regulations of Member States of the EuroBioBank Project* (Instituto de Salud Carlos III, Madrid 2005) 62.

rights required for the maintenance and operation of the database.<sup>51</sup> In Estonia, the Human Gene Research Act 2000 is brief, but it at least indicates that 'an authorized processor or gene researcher shall unconditionally deliver descriptions of DNA or parts thereof to the chief processor ... with or without charge'.<sup>52</sup> Furthermore, it stipulates that, as a matter of law, the results of genetic research and the related intellectual property rights should be provided. Until 2006, EGeen, a private commercial company, had exclusive access to the Estonian Genome Project Gene Bank. However, since then, the project has become fully publicly funded.

The approach of UK Biobank to benefit sharing and intellectual property is more detailed in its specific policy document on intellectual property and access; although, unlike in Iceland and Estonia, that document has no binding legal force, and UK Biobank may amend it at will. UK Biobank's policy affirms as one of its 'core principles' that UK Biobank 'is a managed research resource for the public good'. Thus, it continues, 'UK Biobank will encourage and provide access to the Resource and the results that flow from it as widely and openly as possible in order to maximise its use and value for research'.<sup>53</sup> Moreover, in addition to disseminating research results generally, users of 'protected material' will be required to provide UK Biobank with 'a copy of all of the results of their research based on this material, including negative findings and supporting data, for incorporation into the Resource'. Additionally, 'users who have had access to samples will be required to provide sufficient details of the assay techniques used so that other researchers will be able to comprehend the results'.<sup>54</sup> This same policy of return of enriched results is followed by CARTaGENE and Generation Scotland.

Overall, in the absence of royalties, profits or patents (which, at first glance and for a number of years to come, would seem to be inapplicable to such infrastructures) it is difficult to apply the international principle of benefit sharing beyond the obligation to generate and widely disseminate new knowledge, and to return research results to the population databases. Rapid publication is part of a growing trend (and pressure) to make research results public, whether they be positive or negative. In the case of genomic databases, this is all the more important considering the high levels of public participation and investment.<sup>55</sup> The application of the concept of benefit sharing to genomic databases will continue to be problematic to say the least. For, as noted already, such databases are not primarily disease-based or part of clinical trials, but rather research infrastructures serving as a resource for multiple, more specific protocols. Participants provide access to

<sup>51</sup> Article 5.11.

<sup>52</sup> Section 19(1).

<sup>53</sup> UK Biobank, *Policy on Intellectual Property ('IP') and Access* (Draft, 11 January 2005), p 2, para A.3. Available at <http://www.ukbiobank.ac.uk/docs/UKBiobankIPandAccesspolicyfirstpublicdraft11.1.5final2.pdf>.

<sup>54</sup> *Ibid.*, p 8, para C.9.3. 'Protected material' includes data (in anonymised form) relating to individual participants' health, lifestyle and environment, biological samples and data derived from sample analyses: *ibid.*, p 5, para C.6.3.

<sup>55</sup> Bartha M Knoppers and Yann Joly, 'The Social Genome?' (2007) *Trends in Biotechnology* (submitted).

their DNA and socio-demographic and medical information over extended periods of time. But they do so for the future benefit of others, not for themselves. Accordingly, their contributions are founded on the principles of solidarity and equity,<sup>56</sup> as there are unlikely to be any immediate products or profits forthcoming that can be provided by way of return benefits, other than the datasets created.

## 2. Determining Access

The second crucial issue for the interoperability of genomic databases and fostering international collaboration in genomic research is how access to database resources and materials by third parties should be controlled and determined. While oversight through ethical review and ongoing monitoring are major tenets of biomedical research, the form that these control measures take must be somewhat different in the case of genomic databases. Their governance is influenced by the fact that specific regulatory bodies will often be created, which need to uphold the principles of transparency and accountability over time.

The OECD Global Forum on the Knowledge Economy has provided some guidance for the operation of biological research centres. It suggests that audit programmes and quality review are essential, particularly when dealing with users, suppliers and outside bodies.<sup>57</sup> Furthermore, according to the World Health Organization, one model for ensuring the accountability of genetic database creators, managers and users would be through 'the establishment of a regulatory body with a power to grant licences to create and operate databases ... This body might also oversee and monitor the process and outcome of research activities involving genetic databases'.<sup>58</sup>

Yet, beyond such statements, at the international level, at least, very little has been said about the *type* of ethical review and oversight that is specifically required for such long-term infrastructures.<sup>59</sup> Population databases hold the public's interest in trust for future generations. However, neither the level or type of independent ethical review—including the expertise needed to judge the scientific worthiness and public acceptability of would-be users' access requests—nor the mechanisms for continuing oversight of such structures have been properly addressed.<sup>60</sup> In fact, it is at the national level that more

<sup>56</sup> Ruth Chadwick and Kåre Berg, 'Solidarity and Equity: New Ethical Frameworks for Genetic Databases' (2001) 2(4) *Nature* 318.

<sup>57</sup> OECD, *Guidance for the Operation of Biological Research Centres Part 1: General Requirements for all BRCs*, (OECD Publications, Paris 2004), p 6, para 4.2.2.

<sup>58</sup> World Health Organization, *Genetic Databases: Assessing the Benefits and the Impact on Human & Patient Rights*, 2003, para 8.3. Available at <http://www.law.ed.ac.uk/ahrb/publications/online/whofinalreport.doc>.

<sup>59</sup> Shawn HE Harmon, 'The Recommendation on Research on Biological Materials of Human Origin: Another Brick in the Wall' (2006) 13 *European Journal of Health Law* 293, 300.

<sup>60</sup> Mylène Deschênes and Clémentine Sallée, 'Accountability in Population Biobanking: Comparative Approaches' (2005) *Journal of Law, Medicine and Ethics* 40.

specificity is found. Even there, however, while there are important areas of convergence, there are also significant gaps and marked differences of approach. It is illuminating to consider three specific topics in turn: (1) what kind of consent is required to obtain participants' data; (2) data access policies and practices relating to various different categories of users; and (3) ethical approval and oversight provisions.

First, in relation to consent to obtain participants' data, there is no doubt that Iceland's Act on a Health Sector Database, No 139/1998 and its Act on Biobanks, No 110/2000<sup>61</sup> were not only controversial but attracted worldwide attention with regard to the issue of consent in respect of population biobanking. By way of background, the former statute sought to ensure automatic, one-way downloading of encrypted health information on all Icelandic citizens into deCODE Genetics, the private genomics company that was granted an exclusive 12-year licence to set up and run the Icelandic Health Sector Database and to mine the data. Consent to the transfer of health information was 'presumed' under the Act, although eventually opting-out was provided for.<sup>62</sup> In 2003, such 'presumed' consent to the transfer and use of health data was declared unconstitutional by the Icelandic Supreme Court.<sup>63</sup>

Iceland's approach sparked a heated international debate over the nature of the consent required for entry into such databases. By virtue of their longitudinal, largely epidemiological, research nature, and the fact that data security is such that participants cannot be individually identified by researchers who are granted access to the database resources, they are distinct from other kinds of medical research, as well as from drug trials with definite endpoints and interventions. To be true to their nature—as repositories designed to be used by many different researchers, over many decades, for many different sorts of studies, including those that are currently unforeseeable—these resources necessarily require a broader, more general or 'generic' form of consent; albeit not a presumed one. It is unsurprising, then, that the other national databases, whether mandated by legislation (for example, Estonia)<sup>64</sup> or by the project initiators (for example, UK Biobank), all require an explicit, broad consent particular to the specific nature of such databases.

In terms of re-contacting participants to obtain additional data or biosamples over time, some databases incorporate permission to re-contact participants into their initial consents. Thus, UK Biobank and CARTaGENE ask potential participants upon

<sup>61</sup> *Lög um lífsýnasöfn*.

<sup>62</sup> Note that deCODE obtained separate, informed consent from participants for its collection of DNA samples. Furthermore, demographic and genealogical data are largely within the public domain in Iceland. So then, the controversy was focused on the presumed consent basis for the transfer and use of medical records.

<sup>63</sup> *Ragnhildur Guðmundsdóttir v The State of Iceland*, No 151/2003. The Icelandic Supreme Court ruled that art 7 of the Act on a Health Sector Database, No 139/1998 was unconstitutional because it did not give adequate protection to personal privacy.

<sup>64</sup> Human Genes Research Act 2000 (n 32).

recruitment for their permission to re-contact them in the future, should new samples or data be needed. Generation Scotland adopts the same practice by explaining that certain uses will necessitate re-contact.<sup>65</sup> But this practice is not entirely uniform.

Secondly, in terms of data access by different categories of potential user, it is interesting that Estonia is the only genomic database of the six under study to offer participants a right to access their data contained in the Gene Bank. Under the Human Genes Research Act 2000, gene donors and their general practitioners have a statutory right to access their decoded genetic data and descriptions of state of health held in the database at any time, free of charge.<sup>66</sup> This entitlement stands in sharp contrast to the other five databases. But this may be partly explained by the fact that the aims of the databases are not identical. While they all aim to promote genetic and genomic research in order to improve public health, the Estonian project has an additional core objective: to be used to improve the medical treatment of gene donors.<sup>67</sup> Further, the Estonian project is run through general practitioners, who can provide participants with access to their individual data and explain and interpret it for them in a meaningful way within a clinical setting.

With the exception of Estonia, then, which is unusual in this respect, all the genomic databases inform potential participants of the long-term goals of the database, the open-ended nature of potential uses of it, and the fact that all access requests are subject to prior ethical review. But they maintain the position of no return or feedback of individual results. This latter position reflects the general nature of these databases;<sup>68</sup> although in the case of UK Biobank its 'no feedback' policy stance has proven somewhat contentious, particularly because declining to provide individual feedback is by no means an inevitable policy position to take.<sup>69</sup>

Data access policies regarding other potential users are less uniform, although all of the six databases retain full control over access to, and uses of, their resources. Generation Scotland takes a very strict line with regard to access. It allows only its own investigators to have direct access to its database.<sup>70</sup> Elsewhere, when access by the public, charitable or private sector is permitted, this is predicated on future research results eventually being

<sup>65</sup> See, eg, Graeme Laurie and Johanna Gibson, *Generation Scotland—Legal and Ethical Aspects*, AHRB Centre for Studies in Intellectual Property and Technology Law, September 2003, 50. Available at [http://www.law.ed.ac.uk/ahrc/files/71\\_gslawandethicsfullreport03.pdf](http://www.law.ed.ac.uk/ahrc/files/71_gslawandethicsfullreport03.pdf).

<sup>66</sup> Sections 11 and 16(2).

<sup>67</sup> Susan MC Gibbons, 'Are UK Genetic Databases Governed Adequately? A Comparative Legal Analysis' (2007) 27 *Legal Studies* (forthcoming).

<sup>68</sup> Bartha M Knoppers, 'Biobanking: International Norms' (2005) 33(1) *Journal of Law, Medicine and Ethics* 7.

<sup>69</sup> Gibbons (n 67). As she notes, participants in another population database project planned by the US National Human Genome Research Institute will be given the option to be told about findings that affect their health, such as whether they are HIV positive or unknowingly developing cancer: A Coghlan, 'One Million People, One Medical Gamble' *New Scientist*, 19 January 2006.

<sup>70</sup> See Laurie (n 65).

returned to the databases. Due to the possible combination of a number of datafields, the absence of unique identifiers is not a guarantee of anonymity for individual participants where data are released to third parties. Accordingly, safeguarding security is a primary concern. In practice, several different mechanisms are used to ensure security. For example, data access may be limited by the amount and type of data sought, or through mechanisms such as indirect data queries handled 'in-house' by the database operator. Under the latter approach, scientists employed by the database itself undertake data analysis at the request of external researchers, and release only the (aggregated) results of the analyses, not the original data. Other than in Iceland, there is no exclusive right of access by any single party to a genomic database.

With respect to international collaborations and access by users located outside the jurisdiction, national practices and policies diverge even more markedly. Some databases stipulate that biosamples themselves may not leave the country (Iceland and Estonia); or that an aliquot must remain in the country (CARTaGENE). For its part, UK Biobank is reluctant to release biosamples at all, and it requires special justification and approvals before it will agree to do so. Meantime, in order for researchers outside Québec to access CARTaGENE, they must collaborate with researchers in Québec.<sup>71</sup> Once again, we can see that the requirements concerning transborder flows of biosamples and data—which are closely linked with deeper concerns to protect the databases' integrity, their longevity (especially against the depletion of physical biosamples) and participants' privacy—lack consistency.

Thirdly and finally, in terms of ethical approval and oversight requirements and mechanisms, across the board data access is subject to proof of prior ethics approval of research protocols by the appropriate local ethics approval bodies, and agreement to respect the database's chosen security methods (for example, double-coding, use of a data officer or privacy commission, and so forth). This is entirely in keeping with the long-standing international consensus that biomedical research involving human subjects requires prior ethical (and scientific) approval. All six countries under study have specific committees, entities or designated data protection managers to handle access requests. For example, CARTaGENE has a specific data and sample access committee; while the National Cancer Institute's Biorepository uses a system of data access with defined levels of access privileges, which are approved by the institutional review board and/or the scientific advisory board.<sup>72</sup> Requests for access to UK Biobank will be reviewed by a specially designated NHS Research Ethics Committee, as well as being vetted by UK Biobank itself.<sup>73</sup> The position is not so clear-cut, however, in respect of *ongoing* oversight of specific individual projects once data or biosamples have been released to external

<sup>71</sup> See <http://www.cartagene.qc.ca>.

<sup>72</sup> National Cancer Institute (n 35) 27.

<sup>73</sup> UK Biobank Ethics and Governance Framework (version 2.0, July 2006), available at <http://www.ukbiobank.ac.uk/ethics/efg.php>.

users. Nor is it consistent in relation to the ongoing oversight of the operations and management of the genomic databases themselves. National provisions and practices, and the powers, roles and responsibilities of key oversight bodies, tend to vary quite markedly.<sup>74</sup> Once again, this is an area where greater international agreement, harmonisation and standard-setting would be invaluable.

#### D. CONCLUSION: KEY ROADBLOCKS AND INITIATIVES TO PROMOTE COLLABORATION

The above analysis demonstrates that there appears to be some symmetry, both internationally and nationally, on at least two key normative issues (benefit sharing and data access) that are central to achieving successful international collaboration using population genomic databases. There is also widespread global support for such collaboration as a matter of principle, as expressed by several influential international bodies. Yet, as we have seen, there remain significant gaps in the existing framework (especially at the international level) and areas where policies and their implementation diverge between different countries. On top of this, there are several additional roadblocks to interoperability.

In terms of roadblocks, first, we can see that governance is a common priority for individual countries. But not only do their oversight mechanisms differ, the countries are not consistent—or they have no position at all—on either the acceptability of the exchange of data or samples, or on the equivalent recognition of other countries' ethics approval systems. The reality of this issue as a potential obstacle to interoperability and collaboration is best illustrated by the recent opinion of New Zealand's Health Research Council Ethics Committee (HRCEC), which was asked to give a ruling on the ethical acceptability of sending tissue samples (and data) outside its jurisdiction.<sup>75</sup> The researchers (members of an international research group) wanted to send samples abroad, so that the samples could be stored and used overseas for future, unspecified research uses. The proposal was that explicit consent would not be sought from participants for future uses. Nor would new approval be sought from any New Zealand ethics committee. Instead, future researchers would simply seek ethical approval from their relevant local (overseas) ethical committee. The HRCEC concluded that the proposal was ethical, and that separate approvals for future uses abroad would not be required from a New Zealand ethics committee. However, this process would *only* be permissible if 'there is satisfactory

<sup>74</sup> See, eg, Gibbons (n 67), which compares the ethical oversight frameworks applicable to four population biobank projects, including those in the UK, Estonia and Iceland.

<sup>75</sup> Note, 'New Zealand Health Research Council Ethics Committee Case Study: The Ethics of Storing of Tissue for Future Unspecified Research' (2005) 2. Available at <http://www.hrc.govt.nz/assets/pdfs/publications/Case%20study%20Tissue%20Banking.pdf>.

evidence of a robust approval process carried out by a responsible ethics committee and that committee is accredited within a system which meets the standards held in New Zealand'.<sup>76</sup>

Short of establishing a central clearing house of 'ethically' approved international protocols, however, the New Zealand example illustrates at the macro-level what is already a chronic problem in multi-site studies at the national level. Practices, policies and principles applied by different approval bodies, even within a single country, can differ markedly. In the international context, problems of verifying overseas practices and standards, and satisfying multiple bodies in various countries, are even more acute. In recent years, several initiatives have been launched in an attempt to encourage standardisation of ethical norms and practices with regard to population genomics so as to ease international research collaboration. For example, UNESCO has created an Ethics Observatory.<sup>77</sup> While certainly a welcome step, the Ethics Observatory is more of an inventory of resources and centres of expertise than a means for actually co-ordinating ethical approval mechanisms in order to facilitate international collaboration. Similarly, the international HumGen database contains regrouped information on policies, laws and literature specific to the socio-ethical and legal issues surrounding population genomics.<sup>78</sup> Such resources can serve to guide researchers. But examples of actual consents, confidentiality clauses, security mechanisms and intellectual property policies would be equally if not more instructive. Ideal as it may seem, there are examples—such as the Pediatric Oncology Group (POG)<sup>79</sup>—that have managed to use consents and policies that cross international borders, and whose main components and language are standardised (and, therefore, immune from the 'linguistic tinkering' of local ethics review).

Secondly, as mentioned above, even if an international collaboration were to be ethically approved and overseen under the auspices of some 'equivalency recognition' process, its realisation may well be thwarted by the maze of terminology describing the applicable confidentiality and security mechanisms. Nowhere is the issue of semantic interoperability more acute than in the arena of data security. While standardisation is not

<sup>76</sup> *Ibid.*, 2.

<sup>77</sup> The GEObs (Global Ethics Observatory) is a resource hub that links databases of information about ethics activities around the world. It aspires to provide a platform to support the development of bioethical policies and practices. See [http://portal.unesco.org/shs/es/ev.php-URL\\_ID=6200&URL\\_DO=DO\\_TOPIC&URL\\_SECTION=201.html](http://portal.unesco.org/shs/es/ev.php-URL_ID=6200&URL_DO=DO_TOPIC&URL_SECTION=201.html).

<sup>78</sup> See <http://www.humgen.umontreal.ca> and <http://www.popgen.info>. The PopGen site lists the key international, regional and national norms relevant to population genetic research, and details selected literature relevant to each.

<sup>79</sup> POG was a US and Canadian clinical trial co-operative group, formed to undertake childhood cancer studies. All of its protocol-driven cases were reviewed centrally at the Quality Assurance Review Center, a non-profit healthcare organisation based in Providence, Rhode Island. In 2000, POG merged with several other paediatric co-operative groups to form the Children's Oncology Group (COG).

possible due to cultural and legal diversity, at a minimum a lexicon or harmonisation of approaches is necessary. How else, for example, are researchers who receive or access samples from multiple databases to know if a 'reversibly anonymised' sample is the equivalent of a 'de-identified', 'coded', 'pseudonymised' or 'unlinked' one? This issue has an impact well beyond the obvious difficulties of international collaboration when confronted with incomprehensible and differing terminology used to describe the degree of identifiability of data. Certainly, it could thwart the validity of the initial consents given by participants. Most participants will not understand terms such as 'proportional anonymity' or 'pseudonymised' data, for example.<sup>80</sup> However, there will probably also be inconsistent interpretation by regulatory authorities, ethics committees and sponsor companies. At a minimum, a concordance of language—if not the use of simpler language such as 'coded' (single or double) and 'anonymised' (irreversibly stripped of identifiers)—would encourage and foster international collaboration.

Thirdly—and perhaps most importantly—international collaboration must be ensured through a more co-ordinated effort to build, foster and sustain the necessary mechanisms and tools for interoperability. Here, an umbrella organisation, called the 'Public Population Project in Genomics', or P<sup>3</sup>G Consortium, is leading this effort.<sup>81</sup> The objectives of P<sup>3</sup>G are to:

- harmonise biological, medical, demographic and social data collected from participants to allow effective comparisons of datasets, and pooled/combined studies in order to increase statistical power of gene/environment analyses;
- share approaches to ethics, public engagement, governance and intellectual property issues;
- develop policies and strategies for effective translation of genetic data to health care systems in developed and developing countries; and
- promote multidisciplinary training in initiatives in population genomics research (to include biomedical, social, ethical, and public health aspects).

P<sup>3</sup>G is neither a research project nor a database. Rather, it is a network of the major biobanks and population studies around the world. For its members, collaboration in P<sup>3</sup>G substantially reduces the time needed to achieve the size of cohort of persons with a particular disease needed to undertake large-scale genomic and epidemiological research. It thereby enables an understanding of disease risk to be achieved more rapidly, and discoveries to be translated more quickly into health care.<sup>82</sup> The tools necessary for such

<sup>80</sup> See Knoppers and Saginur (n 22); International Conference on Harmonization of Technical Requirements for Registration of Pharmaceuticals for Human Use (n 23) 32.

<sup>81</sup> See <http://www.p3gconsortium.org>.

<sup>82</sup> Sharon A Savage, 'Genetic Association Studies: Where Are We Now?' (2006) 3 *Personalized Medicine* 371.

international efficiency in population genomics are being built and shared through the P<sup>3</sup>G Observatory.<sup>83</sup>

The crucial and pressing question that remains is this: can the same achievements as the P<sup>3</sup>G initiative is bringing about, especially in terms of scientific and technological harmonisation, be accomplished with respect to international socio-ethical and legal interoperability in a way that still respects diversity between genomic databases?<sup>84</sup> In the absence of such common tools, norms, laws and approaches within a properly harmonised international framework, international collaboration will remain an empty platitude.

<sup>83</sup> See <http://www.p3gobservatory.org>.

<sup>84</sup> Jane Kaye, 'Do We Need a Uniform Regulatory System for Biobanks Across Europe?' (2006) 14 *European Journal of Human Genetics* 245.

