

# Nature, nurture and exposure

Connecting biobank data with geographic data could yield public and individual health benefits, but risks to human rights need to be assessed

Jasper A Bovenberg<sup>1</sup>, Anna Hansell<sup>2</sup>, Kees de Hoogh<sup>3</sup> & Bartha Maria Knoppers<sup>4</sup>

In 1854, the Soho area of London was ravaged by cholera, with more than 600 people dying of the disease in less than a month. John Snow's maps of the cholera outbreak helped confirm the theory behind the most famous intervention in the history of public health—the removal of the handle of the Broad Street pump—and so demonstrated the utility of disease mapping. Over the past decade, scientists have developed more powerful systems that link geographical information with data on disease prevalence or severity to analyse the sources, distribution, spread and risk of diseases. Now, the combination of population biobanks with Geographic Information Systems (GIS) could take disease mapping to the next level by connecting a person's biological and health data with environmental exposure. This could expedite biomedical and public health research, but it could expose participants and the public to ethical and societal hazards.

Population biobanks hold great potential for population-based studies into the genetic and environmental causes of common complex diseases. A participant's health data, lifestyle data and family medical history are collected at enrolment, at repeat assessments and through web-based questionnaires. Genetic data are derived from blood samples, and clinical data are collected from medical and other health-related records, such as cancer registries and primary care records. What is missing in population biobanks to date, however, is data on a participant's environmental exposure, including data on air pollution from traffic or industrial sources, noise pollution from transport, climate data, or exposure to

contaminated soils or water. To study associations between phenotype, genotype and environment, the biobank data must somehow be linked to environmental data, which requires the analysis of spatial information and context [1]. The missing link here could be the use of Geographic Information Systems or GIS.

.....  
*“... the combination of population biobanks with Geographic Information Systems (GIS) could take disease mapping to the next level by connecting a person's biological and health data with environmental exposure”*  
 .....

Whereas John Snow had to walk the streets of London to link the cholera victims to the pump in Broad Street, today's epidemiologists can link individual phenotypes to environmental exposures with a mouse click, using GIS. The system enables epidemiologists to link individual addresses to estimates of past and present exposures to multiple environmental variables. It is thought that this combination of population biobanks and GIS tracking data will expedite the mapping and quantification of relationships between health and environment in different settings and should allow analysis at the individual level. Some even claim that these developments will pave the way not only for improved public health, but also for personal geomedicine.

There are clear possible benefits to public health in linking up these data sets, but there are also evident potential downsides if the data were to be used for unethical or harmful purposes. Publication of such findings could compromise the privacy of participants, for example, as the published information might enable their identification. To address this problem, various measures have been developed, including geo-masking and statistical disclosure control measures. However, the publication of these relationships could also compromise participants' social and economic human rights, including but not limited to the rights guaranteed under the European Social Charter (<http://conventions.coe.int/Treaty/Commun/QueVoulezVous.asp?NT=163&CM=8&CL=ENG>). Moreover, even the non-participating local population could be affected, as some of these publications enable anyone to zoom in and out on their local health hazards. This information could be used to deny or condition access to private or social services or could even have an impact on the value of real estate in a given area. This is all the more likely when rumours of local health hazards become scientifically validated. Even if, as a result of the successful removal of identifiers, this does not affect any identifiable participants individually, any of the past, present or future inhabitants of this area as a group could be at risk.

Admittedly, this risk may be a perceived risk rather than a real one. There may be no need to erect barriers to the publication of this kind of findings or data. And even if the risks are found to be real, the joint imperatives of scientific standards (publication)

1 Legal Pathways, Aerdenhout, The Netherlands. E-mail: jabovenberg@xs4all.nl

2 Imperial College London, London, UK

3 University of Basel, Basel, Switzerland

4 McGill University, Montreal, QC, Canada

DOI 10.15252/embr.201439902 | Published online 5 March 2015

and public health (alert the public) might override concerns over social and economic human rights in some circumstances. Nevertheless, we argue that population biobanks should at least work to identify, assess and mitigate any pertinent risks that relate to the use of GIS to ensure that their impact is appropriately addressed. At this stage, we simply do not know whether the risks are real and whether and to what extent studies that link biobank and GIS data qualify as “dual use research of concern”. Once the data are published, they cannot be recalled, and once the information is public, there is no way to control it. Prior to publishing their findings or data, those involved will almost certainly need to conduct a risk-benefit assessment to weigh the various interests of the researchers, public health and human rights. This cannot be done without a proper human rights impact assessment. Also, participants of biobanks or studies must be fully informed about the risks of their participation, including any risks triggered by the publication of their local health hazards. In some jurisdictions, statutes, regulations and guidelines address risks to social and economic human rights—for example, by seeking to ban discrimination for life insurance or in the workplace.

However, such legislation is typically limited in scope. Whether these protections also apply in the specific context of biobank data using GIS remains to be verified. Even if there already may be a practice of discrimination on the basis of zip codes, population biobank studies and data should, as a matter of scientific responsibility, not, however unwillingly, provide additional underpinning for such practices and should consider its impact on social and economic human rights before publishing. This consideration would, as an additional benefit, provide a mechanism to consult and engage those whose social and economic human rights might be impacted.

We therefore propose a model for oversight of population biobank-GIS research of concern to social and economic human rights. It is adapted from the US government Policy for Institutional Oversight of Dual Use Research of Concern to address risks to biosecurity (<http://www.phe.gov/s3/dualuse/Documents/oversight-durc.pdf>). We are aware that the range of real (not hypothetical) threats is

quite different and have attempted to balance these (Fig 1).

There is no need to set up a separate oversight body to implement the policy. Funders and IRBs could simply incorporate the policy in their existing policies and apply them as needed. Generally, the purpose of oversight is to preserve the benefits of research while minimizing the risk that the knowledge, information and data generated could be used in a manner that compromises social and economic human rights. Oversight includes the identification of population biobank-GIS research that raises dual use concerns, as well as the implementation of measures to mitigate these concerns. Such measures should be applied in a manner that minimizes, to the maximum extent possible, adverse impacts on legitimate research and its attendant health benefits, is commensurate with the risk.

“... population biobanks should at least work to identify, assess and mitigate any pertinent risks that relate to the use of GIS to ensure that their impact is appropriately addressed”

Under this Policy, research that uses sensitive biobank-GIS data and produces, aims to produce or can be reasonably anticipated to produce one or more of the effects listed below will be evaluated for potential abuse or unintended consequences. This can involve data that reveal racial or ethnic origin, or concerning health or sex life [2], and studies that associate genotype, phenotype

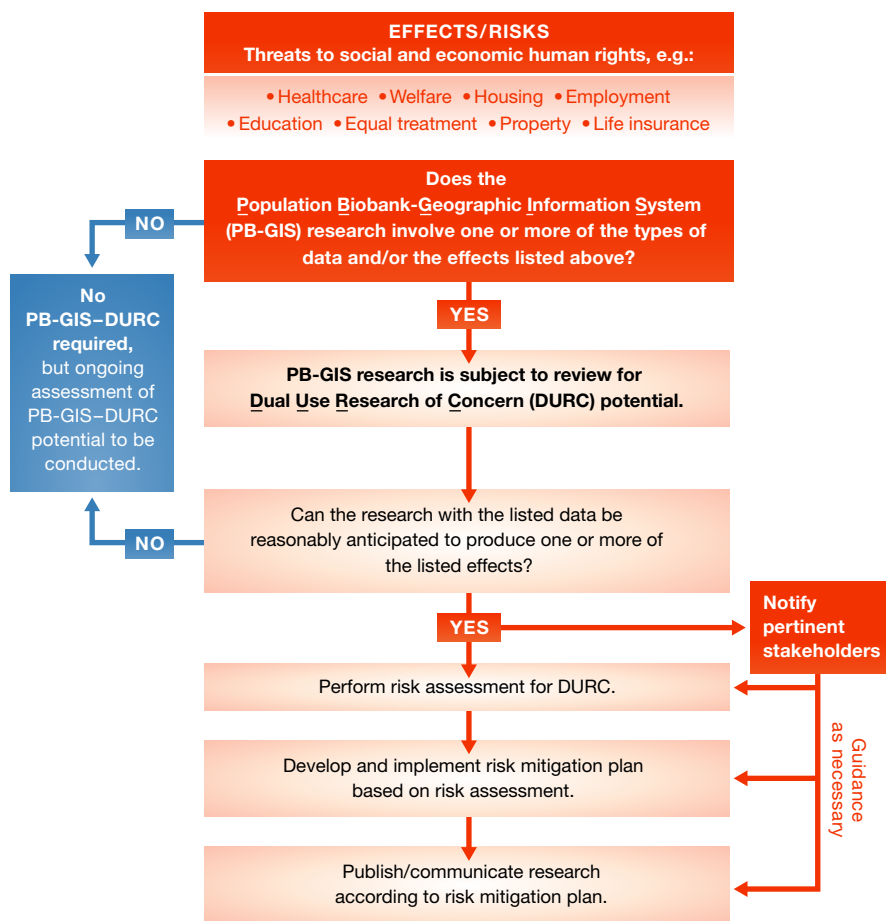


Figure 1. A potential oversight process for biobanks that link biological and health data with geographical information.

**Sidebar A**

The following principles serve as a general guide for oversight of Biobank-GIS research:

- (i) Biobank-GIS research makes possible advances in both public and personal health.
- (ii) Biobank-GIS research has the potential to produce beneficial knowledge, information and data that can also be used in a manner that results in harm to social and economic human rights of both research participants and the public. Therefore, it is appropriate to have in place a framework and tools for the responsible oversight, conduct and balanced communication of such research, on an ongoing basis for dual use potential.
- (iii) Oversight of Biobank-GIS research must recognize (1) the need to protect the social and economic human rights of both biobank participants and non-participants, (2) the need for research progress *per se* and (3) the need to inform private individuals and the public about health benefits and risks. Hence, the degree of oversight should be consistent and proportionate with the possible consequences of both beneficial use and misuse.
- (iv) Effective oversight helps maintain public trust in the Biobank-GIS research enterprise by demonstrating that the population biobank community recognizes the implications of Biobank-GIS research and is acting responsibly to protect social and economic human rights of both biobank participants and non-participants.
- (v) Agencies that fund Biobank-GIS, the recipients of those public funds and individuals who conduct PBGIS research share the oversight responsibility.
- (vi) Any oversight process for Biobank-GIS should be periodically evaluated for both effectiveness and impact on the Biobank-GIS research enterprise.
- (vii) Educating the scientific community about the dual use potential of Biobank-GIS research and cultivating a sense of responsibility for dual use research among Biobank-GIS scientists is essential for promoting responsible research behaviour.

and environmental exposures using GIS. The risks could include threats to social and economic human rights, including but not limited to the social rights to health care, welfare, housing, employment and education and equal treatment listed in the European Social Charter. A risk-benefit assessment should consider the nature of the data or findings and their local specificity; reasonably expectable impact on social and economic human rights of both biobank

participants and non-participants; the existence of statutes, regulations and guidelines that address risks to social and economic human rights; and to what extent these statutes apply to the pertinent findings.

The above only reflects a number of core elements of a potential Biobank-GIS Policy. Multiple issues remain to be worked out, such as the exact applicability and scope of the policy, the distribution of roles and responsibilities over institutions, principal

investigators and mechanisms for oversight, monitoring and enforcement. In any case, population biobanks are starting to use GIS to connect their participants' nature and nurture data to environmental exposures, which could yield benefits for both research and public health. However, the social and economic risks—even if they appear putative—need to be addressed and properly managed early on. Once results from such research are published, it is too late to mitigate any social or economic hazards.

**Acknowledgements**

The research leading to these results has received funding from the European Union Seventh Framework Programme (FP7/2007-2013) under grant agreement no. 261433.

**Conflict of interest**

The authors declare that they have no conflict of interest.

**References**

1. Nuckols JR, Ward MH, Jarup L (2004) Using geographic information systems for exposure assessment in environmental epidemiology studies. *Environ Health Perspect* 112: 1007–1015
2. European Commission (1995) 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. *Official J Eur Uni* 281: 0031–0050, Article 8(1)