

Biobanks – Integration of Human Information to Improve Health



Stefan Nobel
Committee for Research Infrastructures and the Scientific Council for Medicine
Swedish Research Council

Cover Photo: Samples arriving at a biobank. Courtesy of the Swedish National Biobanking Program. Photographer: Kennet Ruona.

Preface

This report presents the results of an investigation commissioned by the Committee for Research Infrastructures and the Scientific Council for Medicine at the Swedish Research Council (VR). The study was conducted during two months ending November 23, 2007. I would like to thank everyone who gave their time for interviews and/or filled in the questionnaire or participated in other ways. Thanks also to Tove Andersson, Camilla Jakobsson, Magnus Friberg and Jan Larsson at VR for comments and practical assistance during the investigation. I would like to emphasise that I claim no expertise in the disciplines involved in biobank-related research, e.g. epidemiology, genetics, statistics, databases or clinical research. To bridge this knowledge gap, I have had the support of a reference group with expertise in some of the above-mentioned disciplines (Appendix 2). Naturally, my background in molecular biology may have had some influence on the investigation. The views expressed in this report do not reflect the official views of the Swedish Research Council, but represent mine alone. I take full responsibility for all facts and information presented in the report.

Stefan Nobel

Contents

1. Summary	3
2. Introduction – Methodology	5
3. Swedish Laws Governing Biobanks	7
4. Biobanks in Sweden	9
5. National Registries in Sweden – An Overview	11
6. International Outlook	12
7. Technologies for Global Analysis of Biobank Samples	13
8. Initiatives on Biobank Infrastructures in Sweden	16
9. Initiatives on Biobank Infrastructures Outside Sweden	18
10. Problems and Proposed Solutions	20
11. Conclusions – What more do we need to know?	29
Appendix 1: Objectives of Assignment	31
Appendix 2: Reference Group	32
Appendix 3: Summary of Questionnaire	33
Appendix 4: Participants	39
Appendix 5: National Quality Registries	44

Appendix 6: List of Abbreviations

46

1. Summary

Introduction

Sequencing of the human genome allows researchers to integrate new data on genetic risk factors with demographic and lifestyle data collected via modern communication technologies. The technical prerequisites now exist to merge large volumes of molecular genetic data obtained by using new high-throughput DNA analysis platforms with clinical, epidemiological and national health registry data. Transcriptomics and proteomics analyses of biobank samples offer completely new opportunities to develop new cures and diagnostics to address common diseases.

Aims

This study aims to provide an overview of the opportunities for biobank-based research in Sweden and identify the common resources needed to conduct scientific research of the highest quality. Information has been collected from a hearing in September, a Web-based questionnaire (184 respondents), a Web forum and interviews with 37 individuals from universities and agencies in Sweden and abroad.

Swedish Laws

Several laws regulate research on biobank samples, the most important being: The Biobank Act, The Ethics Review Act, The Secrecy Act and The Personal Data Act. Of these, the Biobank Act has attracted the most criticism in the field. This criticism can be summarised as: it increases bureaucracy, it does not include all biobanks, responsibilities are unclear and obtaining consent is unnecessarily complicated. Due to extensive criticism from all parties involved, the Biobank Act will be subject to revision. Hence, an investigation of the Biobank Act will be conducted during 2008.

Biobanks

A biobank is defined as long-term storage of human samples that are identifiable to a specific person and linked to personal data. Population-based research biobanks also collect environmental and lifestyle data to enable more powerful analyses. Health services manage most biobank samples since they are used in various screening programmes, in diagnostics and in quality improvement processes. Several population-based research biobanks have been established in Sweden during the past 20 years and contain samples from several hundred thousand subjects. The largest biobanks are located in Umeå and Lund/Malmö.

Researchers primarily experience problems with:

- High cost – long-term funding is unavailable to maintain existing or initiate new biobanks
- Access to biobanks for external researchers
- Information about existing biobanks
- Limited sample resources and fear of losing control over samples and data
- Linkage of biobank databases to databases in health care
- Lack of harmonisation between different biobanks
- High cost of genomic analyses – Swedish researchers fall behind
- Knowledge on how to use biobanks efficiently – epidemiology, genetics, statistics

These problems can lead to difficulties involving: reproducibility of research; performance of cross-disciplinary research; duplication of effort; loss or non-utilisation of valuable samples; and under-utilisation of full potential for modern global analysis.

Biobanks are used around the world, but one finds a predominance of biobanks in the Nordic countries, Europe and the United States. This, however, is about to change since large collections have now been initiated in China and Singapore.

Registries in Sweden

Sweden has an advantage given its many registries and databases on the population. National health registries in Sweden maintain detailed registers for epidemiological analyses of the Swedish population, comprising a valuable resource for biobank-related research. The health care system also maintains valuable national quality registries for the purpose of evaluating treatments. Several databases also exist outside of the health care system, e.g. the multigeneration registry, demographic databases and the twin registry.

Problems: Access to the different registries and databases varies widely. Efficient linkage is needed between these databases and biobank databases.

Infrastructure Initiatives

The National Biobank Programme (NBP), funded by the Knut and Alice Wallenberg Foundation, was active between 2002 and 2006. The programme included a series of work packages addressing quality issues, biostatistics, IT development in health care, technical automation (e.g. DNA extraction, tissue microarray), development of the multigeneration registry and research on ethical and legal issues.

Problems: Although considerable and important work was done to increase the value of biobanks, the goal of coherent, national coordination of Swedish biobanks was never achieved.

Three initiatives have been proposed for a future biobank infrastructure: *Biobank Sweden*, which is a continuation of NBP; *Life Gene*, which is a collection of a large new cohort of 500 000 subjects; and *BIMS*, which involves development of middleware software, enabling linkage between databases to serve as a federated database system.

Problems: There is no national consensus regarding the three proposals, and the different groups behind them have been polarised.

Proposals

Several issues must be addressed to achieve a coherent, valuable national resource for biobank research. Some local infrastructures already exist for biobanks, and several improvements and developments have been made in the NBP. To a large degree, therefore, steps are needed to build on *existing resources* and the infrastructure already in place, and to *efficiently coordinate* them on a national level. However, this must be complemented with new investment, making it possible to perform modern genomics and proteomics analyses in an efficient way.

Recommendation 1 – Biobank Act: VR should promote the revision of the Biobank Act into a more researcher-friendly law without interfering with the issues of personal integrity and safety. Harmonisation with Nordic and European laws should be considered.

Recommendation 2 – Coordination: VR should consider appointing a “Biobank InfraStructure Committee, BISC” that should have a national responsibility for coordinating Swedish biobanks and developing infrastructures to enable efficient use in research. BISC should have a strong interface with DISC, and coordination should involve all aspects, including a biobank federation, middleware solutions, biomolecular analysis and ELSI.

Recommendation 3 – Short-term: VR should consider establishing a joint call with other funding bodies for a genotyped national Swedish reference population, arrange an international review of the Life Gene project and establish a federated biobank solution and national biobank registry.

***Recommendation 4 – Internationalisation:** VR should consider establishing a joint call with other funding bodies for an efficient, flexible, middleware solution that should aim for international harmonisation. Three steps could be considered: national harmonisation, Nordic harmonisation and European harmonisation. VR should also support strong Swedish participation in the European Biobank Infrastructure project BBMRI. VR should promote common Nordic infrastructures, harmonisation of databases, biobanks and legislation, common global biomolecular analyses and competence centres.*

***Recommendation 5 – Coordinate Funding:** VR should coordinate funding with other funding bodies to promote availability of long-term funding for biobank infrastructures and evaluate incoming proposals on biobank infrastructures.*

***Recommendation 6 – Integration:** VR should participate in IT development in health care to achieve the most useful solutions possible for research purposes. VR should promote much closer general collaboration between health services and medical faculties. VR should promote collaboration between human biobanks and biobanks of other organisms to promote comparative research. VR should promote collaboration with the Swedish Environmental Protection Agency as regards biobanks.*

2. Introduction – Methodology

With the rapid progress in genomics research of humans and other organisms, biomedical and health research has expanded from the study of rare monogenic diseases to common, multifactorial diseases¹. High-throughput technologies allowing global analyses of biological systems are widely expected to enable better molecular dissection of these complex, causally heterogeneous diseases into more homogeneous subgroups – a requirement for the advancement of personalised medicine. A more accurate, biology-based definition of disease categories will enhance the development of more effective treatments, reduce undesired side effects of new treatments, improve success in clinical trial design, and will lead to new concepts of disease prevention. Elucidation of complex disease aetiology is challenging because causation arises from not one, but from many small, often additive effects, representing the outcome of genetic predisposition, lifestyle and the environment. Revealing these complex interactions will critically depend on the study of large sets of well-documented, up-to-date epidemiological, environmental, clinical, biological and molecular information and corresponding material from large numbers of patients and healthy persons, collected and made available through biobanks.²

The definition of biobanks in this investigation is similar to that offered by Wikipedia: Biobanks seek to integrate collections of bio-specimens (e.g. blood, DNA, tissue, biopsy specimens, etc) with corresponding patient (personal) data such as genetic profiles, medical histories (phenotype data) and life style information³. The bio-specimens are thus of human origin.

The above-mentioned term, *personalised medicine*, is not new per definition, but has attracted considerable attention as researchers learn more about how to predict disease risk for individual patients – in particular, how to identify people whose genes make them more likely to contract diseases such as diabetes, stroke and cancer⁴. Our knowledge of gene variants (person-to-person variation in the DNA sequence) has increased rapidly since the finalisation of sequencing in the Human Genome Project and the determination of the frequency of DNA variation (polymorphisms) in people of different backgrounds (The HapMap project). Furthermore, scientists have found that people vary not only by single nucleotide polymorphisms (SNPs), but that some people differ in large blocks of DNA, which

¹Collins F, Nature, vol. 429:475, 2004

² K. Zatloukal, M. Yuille, 2007, Biobanking and Biomolecular Resources Research Infrastructure

³http://www.informatics-review.com/wiki/index.php/Biobanking_Definition

⁴<http://www.bioscienceworld.ca/FunctionalGenomicsandProteomicsinPersonalizedMedicine21stCenturyApproachestoComplexDiseases>

are deleted or inserted. Until recently, the major focus was to determine how genetic polymorphisms affected protein structure and function (coding SNPs). However, approaches with global analysis utilising expression microarrays have demonstrated that small differences in an individual's DNA may affect disease risk by altering the regulation of gene expression, thus modifying the amount of protein produced in cells of the body (regulatory SNPs). These disease-associated polymorphisms provide a guide to possible molecular damage that causes disease. As we learn more about how these polymorphisms change the function of genes, proteins, cells and organs, we may be able to predict how small changes in the DNA sequence between different people cause illness, how to better predict how serious the illness may become and how to treat it most effectively. Personalised medicine is based on this new knowledge of genomics and proteomics and is widely believed to result in important changes in how we diagnose and treat many common and chronic diseases.

Improving diagnosis and treatment by further understanding the molecular and environmental basis of disease in humans is a top priority for both society and biomedical researchers. This, however, places a greater responsibility on funding bodies and researchers to improve our understanding as rapidly and efficiently as possible. Such improvement requires the efficient organisation of biological resources (biobanks) and related phenotypic and environmental data that are the objects of study⁵. This has been recognised collectively by the world's major economies who, via the Organisation for Economic Cooperation and Development (OECD), have stated⁶ unequivocally that "biological resource centres are an essential part of the infrastructure underpinning life sciences and biotechnology... essential for R&D in the life sciences, for advances in the quality of the environment, agriculture, and human health, and for the commercial development of biotechnology." In this investigation, however, the biological resource centres (BRCs) are those containing human samples and subsequently referred to as biobanks. This does not imply that BRCs of other organisms are not valuable. On the contrary, they may be even more so if the information they generate is linked to enable comparative studies at the molecular level, e.g. comparative genomics involving comparison on the DNA level.

Sweden has several structural advantages that facilitate research based on biobanks. These include our comprehensive national registries that can be linked by our Personal Identification Numbers (PIN), our open system of public health services and a population that has a positive attitude towards research. We also have a well-developed IT infrastructure, making it easier to communicate with participants in population studies and also legislation enabling participation on a global level. Furthermore, Sweden has a long tradition in biomedical research, including biobank samples. For instance, by utilising our biobank collections, Swedish researchers have contributed towards a better understanding of diabetes, cancer and rheumatoid arthritis. The benefits of conducting biobank-related research in Sweden are reflected by the willingness of foreign funding bodies to substantially fund research on Swedish biobanks⁷. Although we have a high level of expertise, much of the biobank research in Sweden has yet to utilise the modern technologies available in genomics and proteomics. Swedish scientists have also contributed substantially in the area of functional genomics, e.g. Swedish researchers have invented some of the new, high-throughput technologies.

Despite the fact that several Swedish biobanks are internationally competitive and we have the competence and expertise, this field of research suffers from low investment, fragmented structure and inadequate cross-disciplinary collaboration. If this could be changed, Sweden could be expected to achieve much greater success in modern molecular medicine.

Aim and methodology

⁵ M Yuille et al, Briefings in Bioinformatics, in press

⁶ Biological Resource Centres: Underpinning the Future of Life Sciences and Biotechnology. OECD Paris 2001

⁷ Lernmark., Hearing on Biobanks at Swedish Research Council (www.vr.se) September 18, 2007.

This investigation aims to elucidate the potential of Swedish biobanks and their possibilities for development from a researcher perspective to identify the common resources and infrastructure needed to promote biobank-based research (Appendix 1). Information has been obtained from different sources; from the initial hearing at VR where the medical research community was invited, from a Web forum, from a questionnaire (Appendix 3) and by interviewing researchers and representatives from universities and agencies.

The investigation had the support of a reference group (Appendix 2) with expertise in research on biobanks and infrastructure. The group included representatives from Norway and Finland. Input from the group has been crucial in the investigation. However, the investigator is solely responsible for the facts and views presented.

A survey was developed basically for the purpose of structuring the input from a selected population. This did not require a sophisticated questionnaire design or statistical analysis. Furthermore, no information source is currently available that identifies the researcher population involved in biobank-related research in Sweden. The population that was selected to receive the questionnaire included individuals registered as managers of biobanks at the National Board of Health and Welfare, some of the persons attending the hearing and the review panel for the Scientific Council for Medicine. In total, 783 persons were contacted by email and asked to participate, of which 184 filled in the questionnaire. Most of the questions were derived from the assignment and were intentionally open-ended questions, giving the respondent freedom to answer without limitations or direction. These were complemented with several checkbox questions to simplify part of the analysis. Appendix 3 includes a summary of the questionnaire, and it is discussed further in Chapter 10.

Interviews were held with 37 persons in meetings or by telephone and in a few cases by correspondence (Appendix 4). Documents concerning other investigations and articles, information from different organisations, etc have also been collected. The short timeframe made it necessary to limit the investigation to only an overview of this large and complex area. Hence, the investigation does not claim to be comprehensive. The intent has been to invite as many as possible to express their views. In addition to the population given the questionnaire above, all university and college administrations in Sweden were contacted to inform their researchers of the investigation and the opportunities to participate through the website forum, a public version of the questionnaire on the website, or by direct contact.

3. Swedish Laws Governing Biobanks

Listed below are the most important laws in Sweden governing the use of biobanks in research:

- Biobank Act (SFS 2002:297, Biobankslagen)
- Ethic Review Act (SFS 2003:460, Etikprövningslagen)
- Secrecy Act (SFS 1980:100, Sekretesslagen)
- Personal Data Act (SFS 1998:204, Personuppgiftslagen)

The Biobank Act aims to protect donor integrity, while also promoting research on biobank samples⁸. The National Board of Health and Welfare is the central government authority commissioned to implement the Biobank Act. For this purpose, the board has developed several regulations and practical rules, namely SOSFS 2002:11, SOSFS 2004:2 and SOSFS 2006:19, which govern how to interpret the law.

Together, the law and the regulations can be summarised as follows (compiled from ref 8): they apply to biobanks formed in Sweden by public or private health services (primary biobanks), or to biobanks formed by using samples from a primary biobank (secondary biobanks).

Hence, the Act does not apply to biobanks that have been formed and assembled by any organisation other than a health care provider, e.g. a pharmaceutical company.

⁸ <http://www.bioethics.uu.se/biobanker/>

The Act applies only to samples that can be linked to the persons from which they are derived (by e.g. breaking a code-key) and only on samples stored for a long time.

The entity responsible for the biobank, i.e. the health care provider or the research institute, must determine the purposes for which the biobank can be used. The decision to form a biobank also needs to be registered at the National Board of Health and Welfare. The entity responsible also determines who will have access to the samples, and it cannot be forced to release samples.

Samples from a secondary biobank cannot be distributed further, with some exceptions, e.g. samples from a biobank used for research can be released to another unit for research purposes. Samples can be sent abroad, but need to be coded and to then returned or destroyed when the work has been performed.

The Biobank Act specifies that informed consent from the donor is required to store and use any human samples. Hence, donors should receive relevant information to enable them to decide whether or not they want to consent to the utilisation of their samples for biobank purposes, i.e. a specific research project. Consent may be withdrawn, completely or partially, at any time without giving motivation for the action. If this means withdrawal for any type of use, the sample shall be destroyed or be unidentified, i.e. the *linkage* between the ID and the sample of the donor shall be destroyed.

Hence, the Biobank Act regulates only the *physical samples*. Other laws apply to all other information on donors, e.g. health records, health registries and questionnaires.

If one wants to use samples in a biobank for any purpose other than that specified in the consent, then a new informed consent must be obtained for the new purpose. Exceptions to this condition apply regarding consent for research and clinical trials. In such cases an *ethical review board* will decide if new consent is needed. This is regulated by the *Ethic Review Act*. An ethical review board must approve all research using biobank samples.

The Secrecy Act regulates access to personal data and *the Personal Data Act* regulates how the data can be used. Hence, if a biobank is collected outside the health care system (e.g. by a pharmaceutical company), then these laws regulate its use for research, and the access to and handling of personal data.

The Biobank Act has received extensive criticism from several public authorities including: the National Board of Health and Welfare itself, the Swedish Association of Local Authorities and Regions, industrial organisations, researchers and institutions. The main points of criticism are:

- Increased bureaucracy to gain access to biobanks (especially when multiple biobanks are to be used in one study), or to form new biobanks.
- Unclear responsibilities – who can serve as the organisational entity responsible, and who can be responsible for access to biobank samples?
- The law does not include all biobank samples.
- No protection of biobank samples as regards use by the police.
- Complicated to obtain consent – why not use an opt-out system? Several researchers point to the situation in Denmark where biobanks are considered as special databases of human information (albeit the information may not be extracted from the samples yet). The default for donated samples is presumed consent. If the donor wants to withdraw consent he/she notifies a central registry at the Danish National Board of Health and Welfare – the opt-out registry. Researchers are obliged to routinely compare their biobank databases against the opt-out registry to withdraw any matching persons.

Although the investigator does not claim any expertise as regards legislation, and cannot completely judge the law, the criticism seems to be reasonable. However, the integrity of the individual is obviously crucial, and any changes in legislation must be carefully investigated. Ultimately, biobank research depends completely on participation by the public. Hence, good information and discourse with the public is highly important in these matters. Due to the criticism described above, and some reported inconsistencies, the Biobank Act will be

subject to revision. During 2008, a committee will investigate the revisions necessary in the legislation. As of this time it is unclear whether or not the Swedish Research Council (VR) will participate in this investigation. However, VR should at least be consulted about the proposal.

Utilising different governmental or county-controlled databases, the following laws may also apply to biobank-related research:

- Freedom of Press Act
- Archive Act (SFS 1990:782)

The Freedom of Press Act states the principle for public accessibility to any information handled by authorities, i.e. the right of Swedish citizens to access records or documents that have been sent to, or produced by, any public agency.

The Archive Act states that agencies are responsible for keeping material and documents for future use, e.g. for research. Documents to be archived are those that are of scientific value for the discipline in question or others, that are of value for cultural history or personal history or that are of major public interest.

***Recommendation - Action 1:** VR should attempt to become involved in the investigation of the Biobank Act. If this is not possible, then VR should closely monitor the investigation and provide it with important feedback. An opt-out system, similar to the one being used in Denmark, should be considered providing it is in concordance with the ethics and public awareness on consent. Several recent articles discuss the ethics concerning consent⁹. Furthermore, it is important to monitor developments in international legislation, primarily in the Nordic countries and Europe. Research on biobank samples increasingly involves international collaboration and harmonisation of legislation on biobanks, and is thus an important issue. The Norwegian Biobank Act is currently being revised, and the forthcoming (early 2008) new proposal should be closely monitored.*

4. Biobanks in Sweden

As of July 2007, Sweden had 651 biobanks registered with the National Board of Health and Welfare (Socialstyrelsen). Most of these biobanks are being used in health care, quality assurance, education, research and clinical trials¹⁰. Many different types of human tissue samples are being stored in biobanks, most of which are categorised as containing: tissues, cells/cell lines, genomic material (DNA), blood or blood-plasma or urine. However, the vast majority of *samples* are probably blood or different plasma fractions and paraffin blocks, although the registry does not provide this information.

With no claims of being comprehensive, the following list describes the major biobanks (in terms of number of subjects). The list is based on information from the National Biobank Programme (NBP)¹¹, the International Evaluation of Swedish Biobanks in 2005 and other sources.

- *Pathology biobanks* located at hospital departments of pathology over the past 50 years. They are estimated to contain approximately 50 million samples of formalin-fixed paraffin embedded blocks stored for pathological diagnosis and around 20 million cervical smears from the population-based, organised and invitational cervical screening programmes.

⁹ Hansson G et al, Lancet Oncol 2006;7:266-69. Helgesson G et al, Nature Biotechnol 2007;25:973

¹⁰ Biobank Registry at The National Board of Health and Welfare

¹¹ <http://www.biobanks.se/>

- *Biobanks at clinical virology* departments in hospitals have stored serum samples submitted for virological diagnosis going back as far as 30 years. An estimated 4 million samples are stored. Another 1 million samples are stored at the *Swedish Institute for Infectious Disease Control (SIIDC)* and date back to the 1950s.
- *PKU biobank/registry*. Population-based screening for metabolic diseases of newborns have stored all samples (dried blood on filter paper) from all infants born in Sweden since 1974, totalling an estimated 2.8 million samples.
- *The medical biobank in Umeå* contains around 265 000 blood samples from around 156 000 individuals collected in the Northern Sweden Maternity cohort and the Northern Sweden Health and Disease study (Västerbotten Intervention Project Cohort, the Northern Sweden Monica Cohort and the Västerbotten Mammary Screening Cohort). These cohorts involve up to 20 years of follow-up.
- *Malmö Preventive Medicine cohort* was a population-based, health-promotion project with blood samples from 33 000 subjects with 30 years of follow-up.
- *Malmö Diet and Cancer cohort* contains samples (plasma, buffy coat and lymphocytes) from 30 000 participants with 16 years of follow-up.
- *Cohort of Swedish Men* includes 46 000 recruited participants from Västmanland and Örebro Counties, coordinated from Karolinska Institutet.
- *All Babies in Southeast Sweden (ABIS)* contains 17 000 subjects, located at Linköping University.
- *KI (Karolinska Institutet) Biobank*. Includes samples from the *Swedish Twin Registry (STR)*. Contains around 40 000 subjects from different projects. Of these, 12 000 subjects are estimated to belong to STR in the beginning of 2008.
- *Fresh Frozen Tissue Biobank* at Clinical Pathology, Uppsala University has around 50 000 tissue and cell samples.
- *Tissue Biobank for Cancer Research* contains tumour tissue samples from 30 000 individuals located at Lund University.
- *The National Environmental Biobank* contains samples from nearly 10 000 individuals collected since 1990 (Ingvar Bergdahl, UmU). It is funded by the Swedish Environmental Protection Agency for monitoring environmental pollutants in humans and located at Umeå Medical Biobank.

Apart from the above, several smaller biobanks or collections located in clinics around Sweden are valuable from different perspectives, such as the longitudinal cohort of men (2000-3000) in Uppsala (ULSAM) with extensive follow-up (over 30 years) and diagnoses, e.g. dementia.

Lack of overview

The absence of a comprehensive national registry, listing all (larger) biobanks, makes it difficult for other researchers to efficiently utilise Sweden's biobanks. The registry at the National Board of Health and Welfare is not very useful and provides only limited information. The National Biobank Programme has made a good attempt to form a national registry by inviting all larger biobanks to be listed, but the list is incomplete and is not updated.

Lack of harmonisation

The National Biobank Programme (NBP) has information on participating biobanks. However, the information varies among biobanks, and researchers cannot access the biobank databases, which are probably stored in several different formats (some not even computerised). Annotation of information in the different databases is not harmonised since each biobank database was constructed individually without the intent to link to other biobank databases.

Access to data

The access to the different biobank databases and/or further detailed information about the samples within is completely dependent on the principal investigator (PI) of the biobank. As they, or the entity responsible, are required to assure the security of the biobank in compliance with the Biobank Act, they usually cannot release raw identifiable data for analysis and linkage to registries, and can only release processed, unidentified data. In some cases they might also be unwilling due to concern about losing control of their data, or about competitors using the data.

5. National Registries in Sweden – An Overview

National population registries in Sweden are kept primarily at Statistics Sweden (SCB)¹² and the National Board of Health and Welfare¹³. The latter keeps health-related national registries, such as:

- *Hospital Discharge Registry*, all diagnoses and medical treatments since 1961
- *Cancer Registry*, all cancer cases since 1958
- *Death Registry*, all immediate and underlying causes since 1952
- *Medical Birth Registry*, all births since 1973
- *Prescribed Drug Registry*, all prescriptions since 2005

Statistics Sweden houses useful registries for biobank-related research, for example:

- National Multigeneration Registry, can be used to identify first degree relatives to any person born in 1932 or later
- Longitudinal Individuals database (LINDA)
- Registry of living conditions (ULF)

Karolinska Institutet houses the Swedish Twin Registry (STR), containing phenotypic data on 85 000 twin pairs.

Sweden's health care system is regionally controlled and run by the different counties, which in turn are coordinated by the Swedish Association of Local Authorities and Regions (SALAR). SALAR holds responsibility for the National Healthcare Quality Registries, to date 56 registers on different diseases (Appendix 5)¹⁴. Together, the National Board of Health and Welfare and SALAR finance the formation and maintenance of the registries. Three competence centres are established for this purpose:

- UCR – Uppsala Clinical Research and Registry Centre
- NKO – National Competence Centre for Musculoskeletal Disorders
- Eye-Net Sweden

There are several additional databases in social sciences and social medicine that could be of interest for medical research on biobank materials. Traditionally, these databases have been located at various universities and institutes, with no overview and little effort towards harmonising¹⁵. This has been known for some time, and to promote efficient research on these types of databases a special Database Infra-Structure Committee (DISC) was formed at the Swedish Research Council. DISC plans and coordinates investments to make research databases accessible in social sciences, social medicine (epidemiology/public health research) and research in the humanities. Examples of databases that are of interest to biobanks that DISC now evaluates for funding on behalf of VR include: the Swedish Twin Registry, Cohort of Swedish Men, the Demographic Database in Skåne, and the Longitudinal Database on Familial Cancer. Another database of interest is the Demographic Database

¹² <http://www.scb.se/>

¹³ <http://www.socialstyrelsen.se/>

¹⁴ Nationella kvalitetsregister 2007, ISBN-13:978-91-7164-280-6

¹⁵ Vetenskapsrådet, Rapport "Strategi och infrastruktur för världsledande forskning på svenska register" 2005

located at Umeå University. The plan is to link this database to both the multigeneration registry and the biobank database at Umeå University.

One initiative taken from DISC has been to fund the MONA project at Statistics Sweden (SCB), aiming to increase researcher access to the registries at SCB. As SCB holds the multigeneration registry, this is a promising development for biobank research in Sweden. However, further development is needed to connect to biobank samples and molecular data to the registries.

6. International Outlook

The use of biobanks is well distributed throughout the world, but the Nordic countries have disproportionately large numbers of collected human samples as reported by the international consortium, Public Population Projects in Genomics (P3G)¹⁶. One of the most successful and controversial biobanks, both globally and within the Nordic countries, is the biobank controlled by the company deCODE¹⁷ (Iceland), which covers nearly half of the population, i.e. approximately 100 000 subjects. The strengths of this biobank are its high population coverage (aims to cover the entire population of Iceland), its ability to efficiently utilise health care records and its coupling to an outstanding genealogical database going back as far as 1000 years. However, the fact that it is commercially controlled has raised concerns.

The Norwegian Institute for Public Health (NIPH) coordinates large population cohorts from different regions in Norway, sampled outside the health care system through the “Biohealth Norway” programme (3). Biohealth Norway is funded by a substantial grant from the Norwegian Functional Genomics Research Program, FUGE (at the Norwegian Research Council). Currently, it includes around 400 000 subjects, consisting of the two cohorts CONOR (185 000 subjects) and the Norwegian Mother and Child cohort study MoBa (210 000 subjects). The collections are ongoing with a target of 500 000 individuals for all the participating biobanks. This large number makes the combined cohorts the largest in the Nordic countries and one of the larger biobank collections in the world.

Denmark also has several large research cohorts, e.g. the National Birth Cohort “Better health for mother and child” with blood from 100 000 pregnant women, the Nutrition, Cancer and Health biobank with around 60 000 subjects and the Greenland Biobank with blood samples from > 16% of the Greenlandic population. Denmark has constructed a national pathology registry of all pathology biobanks and is famous for its large number of registries in all aspects of life (up to 200 databases) and the well-developed interface between them, which together form an extremely important national asset.

In Finland, the National Institute for Public Health (KTL) coordinates several of the research biobanks. A large biobank harbouring DNA from over 200 000 individuals has been constructed and includes an automated collection service and automated DNA extraction facilities. Collections include the Finnish Twin Cohort (42 000 subjects), the ATBC study (29 000 subjects) and the Finnrisk study (23 000 subjects). KTL also houses internationally renowned research in human genetics, e.g. KTL is the coordinator of several joint research efforts such as the GenomeEUtwin project and the Nordic Centre of Excellence in Disease Genetics (NCOEDG). KTL also serves as the Finnish node in the recently established Nordic EMBL networks in molecular medicine.

Looking at the large (>10 000 subjects) registered cohorts at P3Gs observatory, apart from the Nordic countries, one finds that Great Britain and the Netherlands stand out with several large cohorts each including UK Million Women Studies (1 300 000 subjects) and the forthcoming UK biobank (target of 500 000 subjects), LifeLines (target 165 000 Dutch subjects, in planning phase) and the Netherlands Twin Registry (75 000 subjects). Estonia and Germany (the KORA-Life cohort) also have several larger biobanks and ambitions to

¹⁶ <http://www.p3gconsortium.org/about.cfm>

¹⁷ <http://www.decode.com/>

start new ones. Although not listed at P3G, the Biobank in Graz (Austria) founded by Kurt Zatloukal has apparently collected a large number of samples (from 800 000 subjects). Several attempts have been made to pool samples and data to obtain a sufficient number of cases in Pan-European projects, one of the largest being EPIC (European Prospective Investigation into Cancer and Nutrition) that includes 520 000 subjects in the joint program. Another successful Pan-European collaboration has been GenomEUTwin, including as many as 600 000 twins from Europe and Australia. Australia has clear ambitions both in biobanking and research expertise utilising them, e.g. the Melbourne Collaborative Cohort Study (41 000 subjects) and the cross-disciplinary Laboratory for Genetic Epidemiology at Western Australian Institute for Medical Research with their extensive population registries.

In Asia, large cohorts are found in Japan, China and Singapore. To date, the Kadoorie Study of Chronic Disease in China encompasses 415 000 subjects (target 500 000), the Japan Public Health Centre-Based Study contains 140 000 participants and the Singapore Consortium of Cohort Studies targets 250 000 individuals (only 3000 recruited so far).

Internationally, the United States has the largest number of research cohorts, exemplified by biobanks such as the Nurses' Health Study (original cohort) of 122 000 subjects with very long follow-up and the newly established Women's Health Study (1 750 000 targeted, 40 000 recruited).

Globally, comparing 82 large, single-country, research cohort studies (>10 000 participants per study) one striking result is that of the 7 800 000 subjects included, 3 500 000 are European, whereof as many as 2 000 000 individuals come from the Scandinavian countries. In other words, 25% of the subjects currently enrolled worldwide belong to our minor population in the North, reflecting a disproportionate global sample¹⁸. This probably relates to the similar national health care structures in the Nordic countries that offer favourable conditions for biobank sampling. These figures were obtained from the P3G website so they include only the studies registered there.

7. Technologies for Global Analysis of Biobank Samples

The development of new 'omics' technologies have placed high expectations on the wealth of information that seems to be extractable from biobank samples. The 6 FP programme MolPAGE (Molecular Phenotyping to Accelerate Genomic Epidemiology) has addressed this through a comprehensive approach, although the first run focuses focussing on diabetes¹⁹. Presented below is an overview of their work packages, covering basically all types of global analyses existing today. As shown in Figure 1, operationally the work packages fall into four main areas; 1) sample-related, 2) technology-related, 3) informatics and analysis and 4) training and management. This reflects the complexity of the ideal situation, i.e. where one aims to compare data from all sets of technologies, since sample handling varies between analyses, the technologies vary in terms of their maturity and global scope of their analyses and bioinformatics vary between technologies, with the merging of datasets imposing yet another challenge.

¹⁸ P3G and Pedersen, N, Hearing on Biobanks at Swedish Research Council (www.vr.se) September 18, 2007

¹⁹ <http://molpage.org/index.asp>

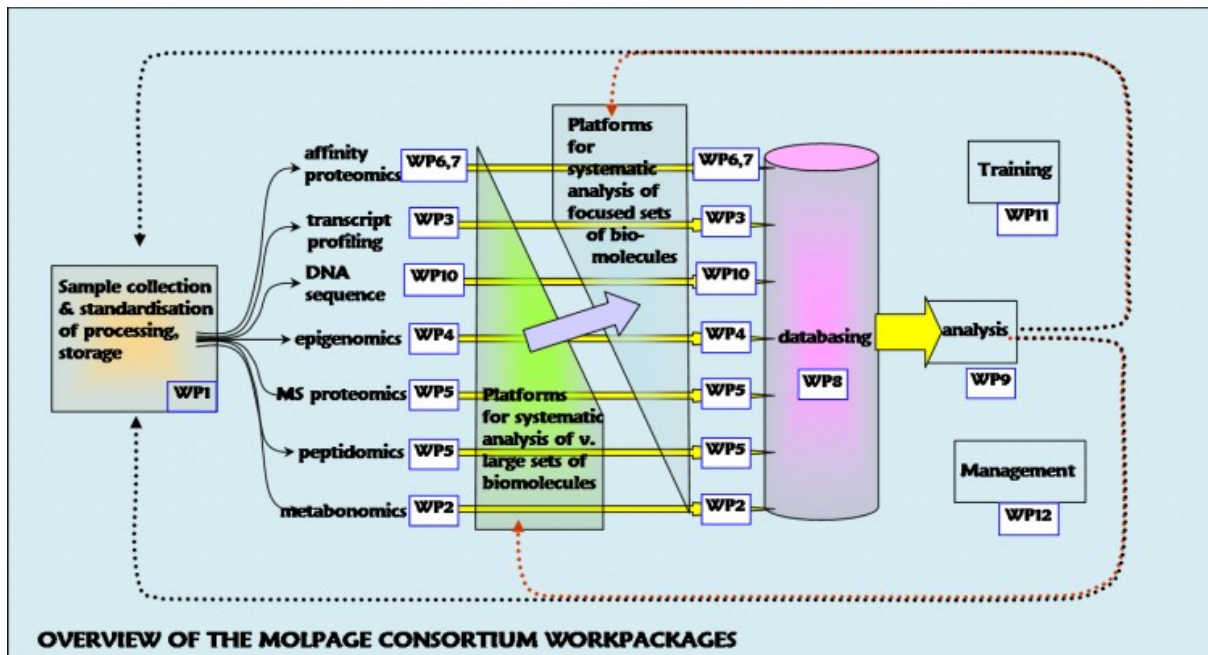


Fig. 1 Overview of the 6 FP MolPAGE.

The meeting, *Standards and Norms in Population Genomics*,²⁰ addressed issues concerning biobank-related research and pointed to the need for standardisation and harmonisation. Participants emphasised the importance of pre-analysis (sample processing) and how analyte measurements are performed (technologies), illustrating the need for international standardisation and formation of reference laboratories. Also, IARC has addressed standardisation in cancer research²¹. The challenge varies between technologies. Some global analyses are described briefly below:

- Genomics – Genotyping. Remarkable progress has been made on technology for SNP analysis (Single Nucleotide Polymorphism), and robust technology now enables genome-wide analysis of large amounts of SNP. Commercial array chips are available with up to 1000 000 SNPs with genome-wide coverage for humans. *Processing of DNA samples for genomics is not a major issue since DNA is quite stable. Facilities for these analyses are found at UU, KI and KTH.*
- Genomics – next generation sequencing²². With the advent of new massively parallel sequencing technologies, science can now (as of October, 2007) sequence the entire human genome in 2 months for 20% of the cost when compared to traditional technologies, and costs are dramatically decreasing. A goal has been set to sequence a human genome for as little as \$1000. This is widely believed to be achievable in the next 5 years if single molecule sequencing is successful. When this happens, sequencing will probably replace today's array technologies for measuring SNPs and RNA expression, yielding much more information. For example, the newly published sequencing of an individual genome indicated many new and unknown SNPs, and that most of the observed genetic variation came from deletion and insertion of blocks of DNA rather than SNPs²³. This suggests that SNP analysis can

²⁰ https://molpage.c2.hostexcellence.com/File_System/Linked_files_webpage/Standards&NormsReport_SP.pdf

²¹ Common Minimum Technical Standards and Protocols for Biological Resource Centres dedicated to Cancer Research / editors, E. Caboux, A. Plymoth, P. Hainaut, 2007 (IARC Working Group Reports ; 2), International Agency for Research on Cancer, WHO

²² Muken and Cherny, 10 September 2007 Healthcare Services & Technology Next-Gen Sequencing 101

²³ Levy et al, PLOS Biology, vol.5 : e254, 2007

show only a smaller part of the total genetic variation. Recently established facilities are found at UU, KTH and soon at LU.

- Genomics – epigenetics. Inherited polymorphism is studied in terms of DNA-methylation sequencing combined with certain chemical modifications of the DNA. Genome-wide data on epigenetic variation are available²⁴, but will be even more detailed if combined with next-generation sequencing. *The importance of sample handling for analysis of DNA methylation does not appear to be well studied.*
- Transcriptomics. DNA microarray-based technology has been utilised for more than a decade to analyse RNA expression, and the technology is mature and robust. Competence in bioinformatics and bioinformatic analysis software developed in Sweden were noted. RNA sample handling is more of an issue than DNA, but there are standard ways to minimise these risks. *NBP has also studied the importance of biobank sample handling for RNA quality. International standards exist for microarray experiments (MIAME, MAQC). Facilities are found at several universities: UmU, UU, KTH and LU.*
- Proteomics. With recent developments in mass spectrometry (MS) instrumentation, highly parallel detection and identification of proteins or peptides is now a reality. In minute volumes as many as 30 000 peptide sequences can be identified. In theory, perhaps 1/3 of the proteome can now be determined simultaneously in a single sample. However, there is wide variation between runs and between similar samples, pointing to sample handling as critical in these massive parallel analysis. *NBP has studied the importance of biobank sample handling for protein quality. International standards for handling samples in these modern types of analyses have yet to be developed*²⁵. Facilities for MS-based proteomics are found at most, if not all, universities: UmU, UU, SU, KTH, KI, and KI. The Swedish Human Proteome Project (HPR) is bringing affinity-based proteomics closer²⁶, and antibody arrays are being constructed with these new antibodies. Facilities for developing immunoarray technologies are found at KTH and LU. At UU, the development of new antibody-based detection technology combined with DNA amplification makes this method highly sensitive compared to traditional immunoarrays²⁷.

According to the Swedish Proteomic Society²⁸ a race is under way to develop technology, increasing the possibility for massive parallel protein determination. Current problems concern the bioinformatic analysis of all generated data and also the biological samples; proteins are inherently much more labile than DNA, making the treatment of samples an issue. Furthermore, biological variation is naturally much higher on the protein level so biological standardisation also becomes crucial. The ultimate goal for understanding disease mechanisms in humans comes down to understanding the mechanisms on the protein level. In terms of the large population-based biobanks, the primary aim of proteomics is to find biomarkers in blood. Regarding this goal, MS-proteomic analysis has not yet been particularly successful. The problems lie in the current dynamic range of the technology, which is much lower than the biological dynamic range, hindering the possibility to find new scarce biomarkers leaking out into the bloodstream. Affinity-based technologies might be more powerful for this purpose, provided that antibodies or other affinity agents for the proteins are present. The ultimate goal of the HPR project is to systematically produce antibodies against 22 000 proteins (one per gene locus). To date, 3000 have been produced. Similar projects are under way at NIH where researchers are trying to produce monoclonal antibodies for the whole proteome. Detailed understanding of protein functions may also be helped by metabolomics, global analysis of low molecular weight metabolites, indicating the

²⁴ Shen et al, PLOS Genetics, vol. 3:e181, 2007

²⁵ <http://www.sps.se/>

²⁶ http://researchprojects.kth.se/index.php/kb_1/io_8632/io.html

²⁷ http://www.uu.se/Adresser/X39_59.html

²⁸ <http://www.sps.se/>

metabolic status in the sample. In Sweden, only a few laboratories conduct such analyses (UmU probably has the largest core facility).

8. Initiatives on Biobank Infrastructures in Sweden

The National Biobanking Programme (NBP)

NBP was the first initiative on coordinating biobanks and building infrastructures for utilising them more efficiently. It was funded by the Knut and Alice Wallenberg foundation through the WCN and Swegene programmes²⁹.

The first aim was to improve the *overview and knowledge* about Swedish biobanks. A national programme with comprehensive participation of the largest biobanks was the objective, and several of the large biobanks in Sweden were recruited to the programme.

The next aim was to develop *national quality standards* for biobanking, which was led by the Umeå Medical Biobank. The quality issues concerned were:

Organisation and process documentation (safety and integrity matters); sample handling methodology and characterisation of sample quality; and documentation of the usefulness of different type of samples (what analyses can be performed). This work is described to some degree in “Good Biobanking Practice” (www.biobanks.se) and a forthcoming book “Methods in Biobanking” (Humana Press, 2007).

The third aim was to make the *Regional Biobank Registries* (RBR, now formed in each county) *as scientifically useful as possible*. These registries are needed for health services to manage patient consent requirements for specific uses linked to the samples. They are also needed to trace samples for destruction if individuals want to remove their samples. This is required by the Biobank Act discussed above. The NBP attempted to incorporate several other tasks, such as:

- comprehensive overview of all biobank samples stored (for the public, researchers and the health care system)
- Implement biobank quality standards
- Provide information about biobanks, definition of study base by linkage to health registries
- Act as an external code-keeping agency

Development of new computer software, called the Biobank Information Management System (BIMS), was visualised and planned to help fulfill the above tasks for RBR.

The fourth aim was to enhance the *usefulness and accessibility* of existing biobanks. This was addressed by setting up the National Tissue Array Centre in Malmö to promote spotting of samples from pathology biobanks and enhance, e.g. immunohistochemical analyses. This centre appears to have been well utilised and serves, e.g. the Human Proteome Resource (HPR) at Uppsala University and also 10 other universities. Researchers have also arrayed, e.g. all breast cancer cases in the Malmö Diet and Cancer cohort together with linked clinical data. Another example involves systematic DNA extraction from large research biobanks by funding DNA extraction facilities and utilising them in Malmö and Stockholm (KI Biobank). An automated pipetting and dispensing robot to enhance delivery of samples has been funded at the Umeå Medical Biobank.

The fifth aim has been to improve parts of the Swedish Registry Infrastructure; funding (together with SCB) was provided to achieve full coverage of the multigeneration registry (located at SCB).

NBP has funded *ethical and legal studies on biobanks* at the Centre of Bioethics (Karolinska Institute and Uppsala University) to investigate the consequences of the Biobank Act and to identify the most important ethical guidelines for biobank-related research. Other important legal issues studied have included the potential copyright aspects of biobanks and the extent to which biobanks can be commercialised.

An international review panel *evaluated NBP in 2005* (www.biobanks.se). The panel emphasised that considerable progress had been made through the activities in the program.

²⁹ <http://www.biobanks.se/>

However, the panel also stressed the need for better informatic linkage to retrieve samples and the information concerning them. Furthermore, the strength of the current biobanks was viewed to be insufficient to study common multifactor diseases. Clear, common rules regarding access were also called for. Generally, the view was that NBP represented a good initial step, but it was not achieving a coherent and coordinated Swedish biobank program.

After the Committee for Research Infrastructures was formed at VR, several applications for grants to plan biobank-related infrastructures were submitted to the committee in 2005:

Biobank Sweden

Biobank Sweden is more or less a direct continuation of NBP and involves the same leading researchers. They aim to build a national resource from existing research cohorts in NBP together with a new collection from a nationally representative control group (reference population) of 50 000 individuals. This population plus other complementary collections would total over 400 000 subjects. Another addition is the formation of working groups around common diseases, e.g. cardiovascular diseases, diabetes mellitus, cancer and other diseases, to promote advanced research on the cohorts. They also envisage further development of the RBR as above. The main argument for Biobank Sweden is that it already has large cohorts with adequate numbers of cases. Hence, research can start at once to investigate the cases compared to other large planned cohorts.

Life Gene

Life Gene is a collection of a new and very large prospective Swedish cohort of 500 000 subjects with detailed information on lifestyle and phenotypic factors. The samples will be open to all researchers, thus constituting a common resource. The aim is to build on the twin registry and the multi-generation registry for recruitment. To complement other new prospective cohorts in other countries (e.g. UK Biobank) the idea is to collect from *below* 50 years, starting with the twins and expand to include their households and other subjects in the same age range. Utilising modern IT communication, e.g. Internet and mobile phones, should dramatically cut the cost of collecting environmental and phenotype data. Collection of such data will be a continuous process, ensuring that the collected environmental data can be changed over time as appropriate. This will increase the likelihood of maintaining up-to-date data to meet research needs in the future. The main arguments are the following: First, to reach significant power to understand common multifactor diseases we need to pool samples and data from several cohorts of this size worldwide. We currently have a window of opportunity to harmonise with the other large cohorts being planned or about to start. Second, no other cohorts have been, or are being, sampled in this age interval, focussing on diseases that establishing themselves at earlier ages. Third, the ambitious goals on lifestyle data collection do not have any comparison. The main criticisms are directed at the high cost of the project, estimated at around 1 billion SEK and the long time horizon, i.e. it will take many years (10-15 years) before the biobank can be used for research.

BIMS

Biobank Information Management System (BIMS) is a proposal by Jan-Eric Litton at KI. Although it has the same name as the software described by NBP, it is somewhat different as it concentrates on the ability to link phenotypic data from clinical records and registries to the samples through the Personal Identification Number (PIN). NBP-BIMS also includes subject consent and sample tracing. However, the KI-BIMS proposal is much more developed, especially since the initial proposal. It can be described as “middleware”, or a hub linking data from the local biobank LIMS (Laboratory Information Management System) to other LIMS at other biobanks. It also links data from various databases with both phenotypic data (registries) and molecular data (e.g. genotype databases). The system has been tested and is described in several publications³⁰. The BIMS developed by Jan-Eric Litton and

³⁰ Ölund G et al, 2007, IBM Systems Journal; 46:171. Muilu et al, 2007, Eur J Hum Genet.; 15:718

colleagues is now in use by several of the biobank projects that are stored at the KI Biobank. Jan-Eric Litton is also coordinating several international database harmonisation efforts, e.g. COGENE, GenomEUTwin and P3G.

Existing vs. new biobanks

The different initiatives reflect two different views concerning the type of biobank structure needed today:

On one hand we find the notion that we already have large biobank collections but they are underutilised. If they are not large enough on their own to generate sufficient statistical power, then they should join with ongoing collaborative projects in Europe and globally to pool subjects and increase the statistical power of the analyses. In terms of prospective cohorts, new collections take a long time before they can be used, i.e. until they have collected enough cases to study of the disease in question. Hence, the need is for long-term funding to maintain existing biobanks, developing them further with high-throughput facilities, and for research projects utilising them.

On the other hand we find the notion that existing biobanks in Sweden are not large enough to provide statistical power to study particularly complex multifactorial diseases. Efforts to remedy this through collaborative international projects face substantial barriers, e.g. harmonisation of data, some of which are considered nearly impossible to resolve. Furthermore, the quality of many existing biobanks is insufficient, both in terms of their collected phenotype data and in terms of their sample quality (e.g. storage, handling and sample type might be inadequate for the desired analyses). Hence, it might be more efficient to collect new biobanks – either case-control collections in the short term, or new large cohorts in the long term.

9. Initiatives on Biobank Infrastructures Outside Sweden

Nordic countries collaborate on biobank-related research through different constellations. One is the Nordic Biological Specimen Bank's working group on Cancer Causes and Control (NBSBCCC). This constellation utilises the strength of pooling some of the Nordic biobanks, totalling up to 2 million subjects, which creates a need for harmonisation efforts. Nordic biobanks are also involved in several European networks, e.g. EPIC, COGENE and GenomEUTwin aimed at harmonising databases and sample data.

The Nordic Centre of Excellence Programme in Molecular Medicine is a more recent initiative with connections to biobank research. The *Centre* is a joint venture between the Joint Committee of Medical Research Councils (NOS-M), the Nordic Council of Ministers and the Nordic Research Board (NordForsk). One of the three new Nordic Centres of Excellence (NCoEs) is the Centre for Disease Genetics, NCOEDG. This programme involves research groups from Denmark, Sweden and Finland, and involves the use of human biobank materials at several sites. One initiative in the programme is to assemble existing genotyped samples (Genome Wide Association, GWA) to form a *Nordic reference population* (control group) with the aim of reaching at least 5000 individuals. This population would serve as a common resource for several disease studies in the programme.

Given the Nordic countries' strengths in molecular medicine, the Nordic EMBL Partnership for Molecular Medicine was recently launched³¹. It consists of three nodes with different areas of expertise: one at Oslo University – The Centre for Molecular Medicine Norway; one at Umeå University - Laboratory for Molecular Infection Medicine Sweden; and one at the University of Helsinki - The Institute for Molecular Medicine Finland. The node primarily utilising biobanks is the one in Finland, which will focus on disease genetics. Over time, however, we can anticipate closer linkage to biobank-related research in the other two nodes as well.

As mentioned above, several pan-European networks have been initiated with EU funding. One such example is GenomEUTwin (FP5-6 integrated project) where 7 European

³¹ <http://www.embl.de/aboutus/news/press/press07/03oct07/>

countries plus Australia collaborated in trying to pool up to 600 000 twin pairs. One of the lasting outcomes was the harmonisation of databases of the twin biobanks and connected phenotypic databases. Jan-Eric Litton at KI coordinated the *database harmonisation*³².

PHOEBE, Promoting Harmonisation of Epidemiological Biobanks in Europe³³ is a FP6 project with the general aim of exploring the key issues to be resolved to efficiently utilise the large cohorts in Europe and to harmonise the newly initiated large prospective studies on the continent. To some extent this continues the work of programmes such as COGENE and GenomEUtwin. The work packages are: 1) Future Biobanking in Europe, 2) Databases and Biobank Information Management Systems, 3) Strategies for Genotyping in Large Scale Biobanks and 4) Ethical and Societal Issues.

Other forthcoming projects in Europe include: ENGAGE (European Network for Genetic and Genomic Epidemiology) for molecular epidemiological studies in existing, well-characterised European and other population cohorts; GEN2PHEN (Genotype to Phenotype Databases: A Holistic Solution) for unifying human and model organism genetic variation databases; and USING-Biobanks for coordinating biobanks and creating common guidelines (in application phase).

When the European Strategy Forum on Research Infrastructures (ESFRI) presented the European Roadmap for Research Infrastructures in 2006 (where the presented projects participated in the 7 FP of Capacities Specific Programme), one of the 35 proposed projects was the Biobanking and Biomolecular Resources Research Infrastructure (BBMRI), which is the first attempt to form a pan-European research infrastructure for biobank-based research. It builds on much of the work from the EU projects presented above. The general aims are stated as follows: “A pan-European and broadly accessible network of existing and *de novo* biobanks and biomolecular resources. The infrastructure will include samples from patients and healthy persons, molecular genomic resources and bioinformatic tools to optimally exploit this resource for global biomedical research”. One finds that Europe has a specific strength in its existing biobanks, constituting an essential resource for: 1) Discovery of gene function; 2) Identification of disease relevance of genes; 3) Exploration of gene-environment interactions; 4) Identification of new targets for drug discovery; and 5) Identification of biomarkers for individualised therapy. Biomolecular resources and analysis tools are required to deliver this potential. The project is divided into seven work packages (WPs): Project Management (WP1), Population-based Biobanks (WP2), Disease-oriented Biobanks (WP3), Biomolecular Resources and Technologies (WP4), Databases and Biocomputing (WP5), Ethical, Legal and Societal Issues, ELSI (WP6) and Funding and Financing (WP7)³⁴. WP4 and WP5 are coordinated by Swedish researchers (Ulf Landegren, UU and Jan-Eric Litton, KI respectively). Kurt Zatloukal (Austria) is the overall coordinator for the infrastructure. BBMRI is receiving funding of 5 million Euros for the preparatory phase, which will be initiated in January 2008. The overall budget will roughly be equally divided over the WPs. The first important deliverables will be: 1) Inventory of resources; 2) Inventory of solutions; 3) Incentive systems for existing biobanks to participate (new merit systems are discussed); and 4) a budget for the construction phase (after 2008). A major problem will be the heterogeneous situation in the different participating countries. Funding is also a major issue since EU has no central budget for the ESFRI projects – it must be decided at the national level. However, having over 50 organisations as co-applicants is an advantage, and among the possible sources (apart from the different member states) is the European Investment Bank. Different types of commitments are also envisaged: a) The National Biobank Program couples directly to BBMRI; b) Commitment analogous to EMBL membership; and c) Pay per access, usage-time.

Apart from the different national infrastructures for biobanking outlined in Chapter 6, a few *de novo* biobanks are also planned – new, large prospective cohorts that are considered

³² Muilu et al, Eur. J. Hum. Gen, 2007

³³ <http://www.phoebe-eu.org/eway/?pid=271>

³⁴ <http://www.biobanks.eu/>

as infrastructures. UK Biobank³⁵ is a prominent example with the objective: “To create a research resource comprising a cohort study of 500,000 participants aged 40-69 for the future investigation of the separate and combined effects of genetic, environmental and lifestyle factors on major morbidity, mortality and health.” Participants have been invited to health care centres where consent, questionnaires and blood and urine samples are being collected. To date, there have been problems with a low rate of response to the invitation. In November 2006, the expectation was to accomplish the collections by 2010. Other prospective cohorts in Europe include Generation Scotland³⁶ of around 50 000 subjects and Estonian Genome Project³⁷ aiming at 100 000 participants in a national gene bank. An interesting new biobank infrastructure is the one being planned in the Netherlands, called The String of Pearls³⁸. This infrastructure aims to coordinate prospective collection of biobank samples through the health care system, more specifically at the University Medical Centres (UMCs), which are integrated organisations of the respective medical faculties and their corresponding university hospitals. The total 5-year budget is 67 million Euros, with co-financing between the state and the universities. The biobanks will be at interuniversity levels, and the initial focus is on eight different diseases (eight pearls).

The Public Population Project in Genomics (P³G) mentioned above is an international consortium for the development and management of a multidisciplinary infrastructure to compare and merge results from population genomic studies. It aims for international harmonisation of databases, datasets and study designs to fulfil the goals. The members are leading public organisations partaking in large-scale genetic epidemiological studies and biobanks from the USA, Canada, Mexico, Europe, Asia and Australia.

Since 2003, an ongoing discussion in the USA has focused on the formation of a new, large national prospective cohort to complement existing ones, described by Francis Collins³⁹. Apparently, decisions and agreements have yet to be reached on the matter.

10. Problems and Proposed Solutions

Questionnaire:

Of the 784 persons contacted 184 (23%) responded to the questionnaire. Sixty-eight percent (68%) of these were currently using biobanks for research. Several of the other 32% indicated that they had stopped, were retired, were only administrators of biobanks or were working in the industry. As mentioned earlier, there is no defined group of biobank researchers. Hence, this survey can be viewed only as unique sampling and not necessarily representative of the whole group of biobank researchers.

As only 68% were currently biobank researchers, the responses from this group were compared to the group as a whole. This comparison did not reveal any major differences (Appendix 3, checkbox questions). Hence, the answers and conclusions below were considered from the whole group (even retired researchers can have interesting views). Questions about goals and plans were often answered in terms of making biobank research easier. The responses could be categorised, e.g. as increased accessibility, simpler regulations, national standards, better link to registries and better utilisation of existing biobanks (Appendix 3). A significant group, but not the majority (40%), said they could foresee problems in reaching their goals with biobank research. Problems that were considered were connected to the goals in the second question, i.e. unclear laws/regulations, little information on biobanks (access). Emphasised here was the funding issue, e.g. difficult to obtain long-term funding to maintain biobanks. Suggested solutions to the problems included: revision of the Biobank Act; more open biobanks/better access; national coordination of biobanks; more resources; need for common infrastructures as IT development; development of common standards and expensive physical equipment for

³⁵ <http://www.ukbiobank.ac.uk/>

³⁶ <http://129.215.140.49/gs/history.htm>

³⁷ <http://www.geenivaramu.ee/>

³⁸ <http://www.hightechconnections.org/index.php?pageid=1>

³⁹ Collins F, Nature, vol. 429:475, 2004

storage and pre-analysis (to maintain quality); open access to biobanks vs. the need of the PI to be in control; and linkage to registries. Even if the majority did not observe problems *that hinder their progress*, most respondents mentioned “problem-related” answers to the question regarding their long-term goals. Hence, this leaves the impression that the majority *do* recognise problems related to their future/current biobank research.

Questions concerning the existing biobanks (Q6 to Q11) revealed an overweight for using clinical vs. population-based biobanks (67% vs. 46%), reflecting the fact that researchers using cohorts also want to use disease-based biobanks either to confirm cases or for case-control studies. This is in agreement with the overlap (respondents could choose both) and the motivations in Q7. Pharmaceutical companies reported that clinical trials and diagnostic development were the motivating reasons for clinical biobanks. However, users emphasise that cohort-based biobanks are the most valuable in terms of their superior study design. A majority (68%, here 81% in the researcher group) preferred to use the existing biobanks. However, some overlap is found between existing and new since many biobanks are ongoing collections and researchers want to use the old, but also start new collections to answer new questions. Regarding the organisation of biobanks, 56% preferred them to be nationally centralised, and 42% preferred local/regional sites. Centralisation is motivated by lower overall costs and the possibilities to automate sample handling, increase visibility and enhance access to the biobank. This contrasts with the fear of increased bureaucracy and loss of control of the local biobank. Some overlap was found (not reflected in the figures), and some respondents mentioned that biobanks could be both centralised and local as long as they are federated, or in a network with national coordination.

Regarding the question on how international initiatives influence organisation and research (Q12), respondents often pointed to the need for widespread international pooling of biobanks to achieve enough statistical power and the need for common regulations and laws to permit this. In answering the two questions on the importance of their biobank research, respondents highlighted the benefits to society, e.g. increasing knowledge (to improve health) and decreasing health care costs (via more efficient therapies). Improved health care was also related to discoveries of new biomarkers for different diseases and molecular diagnosis, both of which improve diagnosis. Personalised medicine was a common hope, relating to much better differentiation of diagnoses (above) along with more differentiated therapies.

Only a minority of those surveyed gave a positive response to the final questions on the possibility of commercialisation, whereof a significant proportion were from drug companies and other smaller enterprises.

Some of the issues from the questionnaire are discussed below along with the information obtained from other sources (interviews, documents etc).

The problem with biobank access

The issue of incentives:

The larger biobanks are partners in several international collaborations, and are utilised by many external researchers, as reflected by the many publications derived from them. The larger biobanks claim to have well-organised structures with dedicated neutral expert committees evaluating incoming proposals for access on scientific grounds only. However, the final decision rests with the PI of the specific collection. The incentive to contribute samples seems to be the aspect of collaboration, yielding publications, which is the common way of dividing merits for work in the scientific community. In some cases the funding bodies (VR for example) require the biobank to be open to anyone, and priority should be based solely on scientific grounds. This appears to be a reasonable working model. However, when someone (i.e. PI) controls a valuable resource, that person has a highly influential position. In some cases, collaboration may not be appropriate (the external scientist and PI may be competitors), creating an unbalanced situation between the two parties. The imbalance could also be just the opposite; the external scientist might be very prestigious, and if he/she competes with a similar external project represented by a lesser known researcher the

situation could easily favour the former since it would bring greater merit to the biobank PI. By not having to “pay” with publications the relationships would be more neutral. If the valuable resource (i.e. biobank) receives public funding, then the requirement for open access is obvious, as put forward by VR above. The requirement for open access could actually be taken a bit further: the biobank could be considered a service organisation, and as long as the customer (external scientist) pays for the service no other strings would be attached. The user fees should then cover only the actual costs of the biobank, not driving the commercialisation of biobanks. However, it is hardly possible for an external scientist to cover the actual cost of delivering samples, especially in larger studies, so funding bodies need to subsidise the cost. The major benefit of considering biobanks as service organisations would be the truly open access.

The problem is the loss of control for the PI, which should not be taken lightly considering the substantial effort often invested by the PI to collect the biobank in the first place, and the need to keep it as a source for professional merits/publications. Historically, biobanks have emerged mainly from the hard work and visions of individual PIs. Today, larger collections tend to be handled on an organisational level rather than on a personal level. Nevertheless, we should not forget the importance (even in the future) of individual drive behind the collection of biobanks to investigate specific hypothesis and research questions. We should avoid actions that would take away this motivation from the individual researcher. One approach might be to give the PIs of the biobanks a couple of years with complete control after the establishment of a mature biobank. This would allow time for PIs to publish their own work before the biobank becomes openly accessible.

Initiatives for increasing incentives:

Smaller biobanks are often completely controlled by single PIs, with only the larger biobanks having neutral expert committees that judge access to samples and exercise some influence over the PI. The European initiative (BBMRI) has addressed the incentives for biobank PIs to join open access networks of biobanks with the possibility of acquiring merits *aside* from the normal merit system (publications). Ideas have been proposed on impact factors for biobanks – a quality stamp that one has a certified open access biobank, which then should be a recognised merit in seeking infrastructure funding.

Overview of available biobank samples:

A need exists for a comprehensive national biobank registry with as detailed information as possible (discussed in Chapter 4).

Study design: Several researchers have mentioned the benefits of having assistance in designing studies, both on a superficial level and on a more detailed level. For instance, there is no public place where researchers can view and search phenotype data sets (health registries), genotype datasets and sample availability data for a given disease or condition. Yet, this is essential to the design of further experiments on a given subset of samples. For an overview, one would need details on case numbers, quality of samples, etc freely available on Internet. Presentation of the data generated from a biobank would also be of great importance, but is often missing. One option might be to have two levels of access (similar as NIH has for GWAS data); the first level with freely available processed data describing the biobank in detail and a second level with secured access where only approved scientists could access the raw data, enabling necessary linkages to design experiments in detail. Linkage of phenotype databases (health registries and other databases) to biobanks is addressed in a special work package in the BBMRI initiative, reflecting its importance (Chapter 9). It is headed by Swedish researchers (BIMS, Chapter 8).

Utilisation:

Although biobanks are such a key element in modern medical research, they are not utilised as extensively as they could be. This relates not only to the lack of information, but also to the lack of funding for utilisation of biobank samples. Simply withdrawing a large number of samples is costly – and, of course, biochemical analyses add further to that cost. Performing modern genomic analyses on biobank samples is particularly expensive, and on a scale that normally is not funded by individual grants in Sweden. However, given the recent success of genome-wide association studies (GWAS) on a reasonable number of subjects⁴⁰, these types of studies would also be valuable to perform on Swedish biobanks. A commonly encountered resource problem for many clinical-molecular and epidemiological case-control studies based on biobank samples is the need to collect and analyse a reference population/control group to compare cases against. In the investigation, the reference group put forth a proposal to establish a common resource infrastructure: to fund a national reference population to be used as a general reference for many types of diseases, and which should involve the determination of genetic polymorphism on a genome-wide basis. For efficiency, this might involve existing samples of good quality, complemented perhaps by new collections. The data could then be made accessible to Swedish researchers. This would cut future GWAS costs in half. The proposal received positive feedback from most (but not all) researchers. Several researchers commented that the Swedish twin biobank would be valuable to use as a reference group.

Having access to expertise and technologies in functional genomics is also important to validate findings, e.g. from GWAS. These include transcriptomics, proteomics and model organisms such as facilities for transgenic mice. In Sweden, a network in functional genomics exists for access to such expertise. This network is frequently used by scientists working with different disease models in simpler organisms. Thus, the field of functional genomics is a cross-disciplinary field that could generate many fruitful connections between biobank researchers and other researchers. A powerful example of such connections would be comparative genomics taking advantage of the fact that on a molecular level many functions are conserved between species. Such collaboration should be encouraged.

National coordination:

As NBP was the first to undertake national coordination of biobanks in Sweden it has made a considerable effort to increase quality, usefulness, efficiency and accessibility for researchers. NBP also addressed important ethical and legal issues by funding research in the area. Nevertheless, the international evaluation did not consider NBP to have reached its goals as a truly national program. Several researchers have also confirmed this. A possible explanation could be that NBP was probably considered to be an individual initiative since it is not part of an institution with formal national responsibility, and funding comes from a private foundation – not national sources. Furthermore, different research disciplines are involved, promoting fragmentation of the research field.

Federated vs. centralised biobanks:

The NBP programme involved a loose coupling between the participating biobanks, promoting participation via the incentive to fund certain functions to be developed. Not all biobanks had the expertise and tended to have no incentive to participate. NBP could not really be considered a federation, but a collaborative project working to form common standards and principles. IARC (WHO) and OECD have outlined the importance of large, centralised robotised facilities capable of handling the enormous number of samples needed for genomics and proteomics, i.e. Biological Resource Centres (BRC). Such ambitions also exist in Sweden. Current biobanks are not easily transformed into such centralised structures. Hence, this issue also relates to the issue on existing vs. new (central) biobanks.

⁴⁰ The Wellcome Trust Case Control Consortium, *Nature*, vol 447:661, 2007; McPherson et al., *Science*, vol. 316:1488, 2007; Helgadottir et al., *Science*, vol. 316:1491, 2007; Sarmani et al., *NEJM*, vol. 357:443, 2007.

To accommodate both existing and new (centralised) biobanks, a realistic approach towards coordination might be to form a true federation of Swedish biobanks where all adhere to certain common regulations and standards demanded by the funding bodies.

Current initiatives:

Our neighbouring Nordic countries, at least Iceland, Norway and Finland, have much better national coordination. Iceland, by virtue of its small size, is easier to coordinate, and Iceland's noteworthy success in biobank research comes from the fact that a single organisation (deCODE) manages the biobank and has the expertise and tools to explore this resource. Norway and Finland have their national institutes of public health (NIPH) as national coordinators. In Norway, the NIPH coordinates the BioHealth programme with funding from the national research council, and NIPH also has national responsibility for these issues even if it does not directly control all participating biobanks. In Finland, the NIPH does not coordinate all biobanks. However, it does control or collaborate on several of the population-based biobanks. The Finnish NIPH (KTL) includes the central national biobank facility for robotized DNA extraction and storage. Furthermore, as in Iceland, KTL has the expertise and tools to explore the biobanks as witnessed by their highly successful genetic research on biobank material.

Recommendations – Action 2 – Coordination:

The Swedish Research Council (VR) should consider appointing a committee “Biobank InfraStructure Committee, BISC” that should have a national responsibility for coordinating Swedish biobanks and developing infrastructures to enable efficient use of biobanks for research. Issues that need to be addressed are detailed below. Several of these are shared with research databases collected in the social and medical sciences. Hence, it is important to consider possible ways to integrate BISC with the VR supported Database InfraStructure Committee (DISC). However, several other issues are not related directly to databases, which calls for careful investigation of the organisation of biobank infrastructures. A rough proposal is discussed below:

Organisation:

A. One way to ensure integration of the efforts of biobanks and databases would be to include biobanks as a separate function under the DISC committee in parallel with the newly established SND (Swedish National Database Resource) and the planned climate/environmental SND.

Advantages

- Close collaboration with other database expertise in DISC
- Synchronised common efforts (e.g. ELSI)
- Easier to manage a single unit

Disadvantages

- Not all ELSI issues shared
- Biobank issues are not only database-related; physical biobank infrastructure, technology/analysis, education
- In contrast to DISC, several biobank issues are often related to integration with health care; physical infrastructure, ELSI, sample logistics, health care databases, medical record systems
- Visibility problem if hidden in a large organisation (DISC)
- Responsibility problems if organised under DISC (which does not share all issues)

B. Another way to integrate the shared functions/issues of DISC and a biobank infrastructure would be to build in some overlap between them (Figure 2). Consideration should be given to placing the BISC committee at a neutral site, where the nodes/functions below could be distributed out to different universities after a competitive call.

Advantages

- Clear, national responsibility for all biobank-related issues by BISC
- Covering all related tasks and issues with distributed nodes
- Integration with DISC expertise (middleware, shared databases)
- Synchronisation with DISC efforts (e.g. ELSI)
- Responsibility of BISC to promote further integration with the health care system

Disadvantages

- Unclear responsibilities of the shared nodes between DISC and BISC
- Less natural contact between BISC and DISC
- More difficult to manage one committee with distributed nodes

Several nodes or functions would be included in BISC. Attention should be paid to the discussions and ideas on organisation presented in the BBMRI project. The following proposed nodes are analogous to the BBMRI proposal:

ELSI = Ethical Legal Societal Issues

Middleware = Development of middleware solutions to link databases (biobanks, registries, medical records). Collaborate with SNIC on computer issues. Collaborate with Bioinformatics expertise.

Biobanks= Coordinating the Federation of Swedish Biobanks, coordinating physical infrastructure investments, development of quality standards (from NBP and international standards)

Biomol Analysis= Expert node in biomolecular analysis (global analysis/omics), standardisation issues, statistics, model organisms (functional verification of candidate genes). Biostatistics, educational issues.

BISC

Proposed responsibilities:

Overall national responsibility

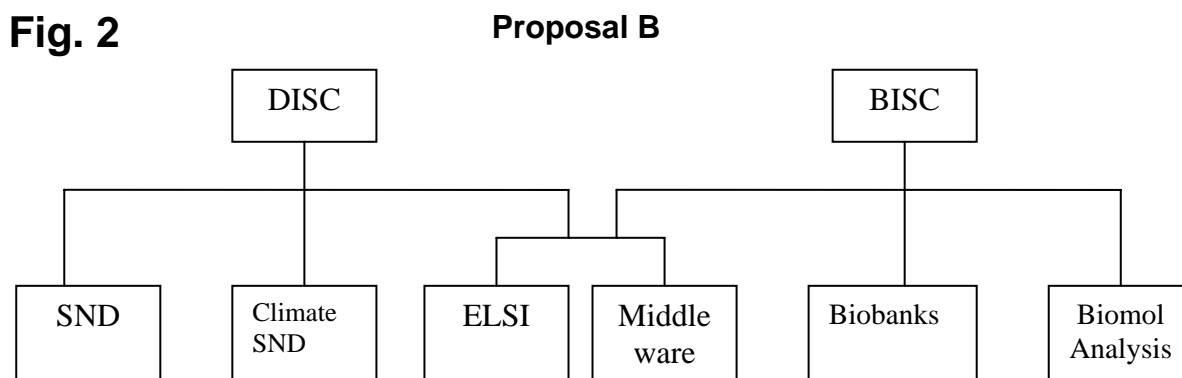
Establish common regulations and standards for participating biobanks, continue work by NBP

Contact point for international collaboration

Contact point for collaboration and integration w SALAR/health care

Coordinate funding

Evaluate grant proposals (infrastructure projects)



Recommendations – Action 3 – Short-term:

In line with the discussions above, issues to be handled immediately include:

- *To establish a national biobank registry with a common meta-data nomenclature for biobank samples. Work from the NBP programme should be continued and expanded. An important issue would be to promote feedback and to link generated data back to the biobank samples. Funding bodies could agree to make this a requirement for funding.*
- *To promote a federated biobank solution, where existing biobanks can remain at their current locations (universities/hospitals), with access opened through middleware software. Participating biobanks need to adhere to common rules and standards developed by BISC. The requirements on funding bodies could offer incentives to participate, and further incentive systems should be developed (e.g. new merit systems). As a complement, larger centralised biobanks integrated with the health care systems would be supported, as exemplified by the plans developed at Karolinska Institute-Stockholm County Council (KI-SLL) and at Lund University and the county region of Skåne (RSKC; Region Skåne Competence Centre).*
- *To retain Sweden's competitive edge, VR and other major Swedish funding bodies (e.g. VINNOVA, KAW and SSF) should jointly establish a call for proposals on a national Swedish reference population. This could be used (i) to investigate the Swedish genomic population substructure and (ii) as a common control group for Swedish Genome Wide Association Scans (GWAS). Such a reference group would cut future GWAS genotyping costs in half. The Wellcome Trust Case Control Consortium (WTCCC) has shown that such a common control group is highly efficient. A proposed and quick solution would be to carry out a genome-wide scan on already available and banked DNA from the Swedish Twin Registry. An international review panel should evaluate the proposals.*
- *International review of Life Gene: Arrange for an international scientific review panel to evaluate the Life Gene proposal. Depending on the outcome, VR could then support Life Gene on scientific grounds, and with the demand that such a large cohort should serve as a truly national resource open to all interested researchers. In particular, a new prospective cohort could open for large-scale proteomics/biomarker studies on humans – for which no infrastructure exists today. The larger biobanks are not specifically designed for this type of research. Any funding of Life Gene should be outside VR's responsibilities.*

Harmonisation problems:

The ability to pool samples from different biobanks to find enough cases and controls for the study in question is a valuable strategy utilised in all of the large collaborative studies. This, however, is associated with problems in that the databases are not constructed in the same way, the information is not annotated in similar fashion and issues of personal integrity arise. Biobank databases need to be harmonised to allow such studies.

Current harmonisation initiatives:

Harmonisation has been addressed in several large collaborations (e.g. EPIC, GenomeTwin, and P3G) and in the recent BBMRI infrastructure project. Where databases have been harmonised, this has been done in a project-driven fashion, depending on the specific information needed for each study. It appears to the investigator that general harmonisation solutions have not been constructed. In Sweden, the BIMS project described in Chapter 8 is the most developed solution for these harmonisation problems and has been used in, e.g. GenomeTwin. The project leader of BIMS is also coordinating database harmonisation efforts at the P3G level. According to BIMS, harmonisation between different biobank databases needs to be performed on a project basis and may not yet be feasible as a general interface. Development towards a fixed interface solution that does not need to be continuously adjusted for each individual project would be important. Hence, harmonisation

and linkage of databases is an important aspect of the international infrastructure for biobanks.

Legal issues:

One of the issues mentioned repeatedly in responses to the questionnaire and the interviews concerned the problems experienced with the Biobank Act and related regulations. To some extent, this may be due to a lack of knowledge on how to interpret the law by the researchers and at the National Board of Health and Welfare. (This criticism is discussed in Chapter 3.) As pointed out in the questionnaire, one of the concerns focuses on international harmonisation to enable the required collaboration between biobanks around the world. According to several researchers, laws in the other Nordic Countries are viewed to be more research-friendly than laws in Sweden, with the exception of Norway. However, Norway is currently revising its Biobank Act. When the Swedish law is investigated in 2008, it will be important to harmonise with the other Nordic countries since we have similar situations in terms of health care, public attitudes towards research and opportunities for biobank research.

Recommendations – Action 4 - Internationalisation:

- *VR should, after further investigation, consider issuing a jointly call for proposals with other funding bodies to establish an efficient, flexible middleware system as an important part of health informatics. A powerful middleware system should be able to link databases from various sources, e.g. biobanks, health registries, health care quality registries and possibly medical records and the MONA system at SCB. It should promote the general formation of federated databases and – to make it as versatile as possible – include experiences from developments in health care IT and other national (e.g. SNIC) and international (e.g. Denmark⁴¹) expertise. First phase: make this a functional national informatics system. Second phase: expand it to the Nordic countries. Third phase: integrate into BBMRI.*
- *Promote collaboration at the Nordic level on common infrastructures; harmonisation of databases, biobanks, legislation and competence centres. The Nordic countries have a competitive niche compared to the rest of the world so this is probably an efficient level for expanded international collaboration.*
- *Establish Sweden as a strong participant in the BBMRI project.*
- *Work with DISC on Ethical, Legal and Societal Issues (ELSI). This includes revision of the Biobank Act and establishment of a national helpdesk to aid in questions related to ELSI. The National Biobank Council at SALAR is a useful resource in this context. ELSI work also includes communication with politicians and authorities on issues related to health registries and biobanks and to promote research and public awareness in the area.*
- *Promote education connected with biobanks; e.g. database usage, statistics, genetic/molecular epidemiology, modern ‘omics’ technologies*

Recommendations – Action 5 – Coordinate Funding:

⁴¹ Dybkaer R. An Ontology on Property for Physical, Chemical, and Biological Systems. APMIS. 2004;Suppl. 117(112):1-210

- *VR should coordinate funding with other funding bodies such as VINNOVA, SSF, private foundations and the pharmaceutical industry to promote availability of long-term funding for biobank infrastructures.*
- *Evaluate incoming proposals on biobank infrastructures*
- *Evaluate if funding should be given for new case-control collections where the existing biobank cases are deficient in quality or number.*

Integration with health care:

This final paragraph does not exclusively address issues surrounding biobank research, but concerns all clinical research. The Swedish Association of Local Authorities and Regions (SALAR) has driven important efforts to meet the conditions of the Biobank Act. This work has been conducted through a specially appointed group, i.e. the National Biobank Council⁴², including representatives from SALAR, the Swedish Association of the Pharmaceutical Industry, the universities and the National Board of Health and Welfare. This group has developed further guidelines and directions on how to utilise the biobanks in health care. SALAR and all universities with a medical faculty have also reached a principal agreement on access to biobank samples for research⁴³. In the future, even if the universities and industry also have their own biobanks, the biobanks in health care will continue to be important for research. Accordingly, continued close collaboration on these matters is necessary, and BISC (or similar body) should be the natural interface between universities/researchers and the health care system on the biobank matters. To make full use of biobank research and other medical research, it is necessary to efficiently reach all medical information in the health care system databases in a *secure* way that ensures individual integrity. Today, it is primarily the health registries, and to some extent the quality registries, that are utilised for research. To access medical information in medical records and other databases one needs to have an informatics system (e.g. middleware) that is integrated with the IT systems in health care. Furthermore, the terminology must be standardised to enable the import and collection of data in a meaningful way. For this purpose, Sweden has recently joined the International Terminology Standards Development Organisation (IHTSDO), which develops the ontology system SNOMED CT. The National Board of Health and Welfare is leading the implementation of SNOMED CT within the health care system. These developments are important to be able to utilise medical databases. Since the counties have sovereign responsibility for managing health care in their counties (total: 18 counties and 2 individual regions), communication problems arise at several levels, e.g. between different IT solutions. To overcome this, SALAR is working through a collaborative project, Carelink, to harmonise the development of IT systems in health care. VR's involvement in this area of development is important to encourage the use of health care's IT systems for research purposes. It is beneficial to have an interface from researchers, e.g. representation by BISC (or similar body) in Carelink and other IT initiatives. Funding research projects in this area could also be considered. Collaboration is ongoing between the universities and the different counties. However, by tradition, this is not always smooth and simple since health services and universities have distinct and separate immediate objectives. The long-term goals are however the same. Although it is not a direct responsibility of VR, one should consider how VR/BISC could initiate discussions with SALAR and the different counties on how to promote further integration through organisational change. VR/BISC could act on behalf of the scientific community. This would help pursue the goals of integration on physical biobank facilities and integration of medical information between the two principal organisations. One foreign, but nearby, example is found in the Netherlands where for the past 10 to 15 years fully integrated organisations called University Medical Centres (UMCs) have brought together the medical faculties and

⁴² <http://www.biobanksverige.se/>

⁴³ Sveriges Kommuner och Landsting, 2005/1964

the university hospitals. The UMCs in Holland have recently initiated an interesting initiative on a common national biobank infrastructure called the String of Pearls (Chapter 9).

Recommendations – Action 6 - Integration:

Health care

- *Participate in the development of harmonised national medical record systems (Carelink) and promote them to be as scientifically useful as possible. Monitor and partake in other IT developments in health care to enhance research expertise and perspectives on IT systems. Fund and promote research in the area.*
- *Promote much closer collaboration between the university hospitals and medical faculties. Initiate discussions with SALAR and the universities concerning integration models such as the UMCs in the Netherlands.*

Other disciplines

- *Collaborate with the Swedish Environmental Protection Agency on their experience of handling human samples for analyses on low molecular weight compounds.*
- *Promote interaction with biobanks and databases on other organisms to enhance the possibilities of comparative genomics, e.g. along the lines of the EU project Gen2Phen, and to share experiences on sample logistics and handling. Collaborate, e.g. with expertise at SLU, regarding biobanks on other organisms.*

11. Conclusions – What more do we need to know?

Due to the time limit, the scope of this investigation does not cover all of the different issues facing biobank-related research. The following are some of the more important structural and technical issues needing further investigation.

Structural problems concern:

a) The proposed organisation (BISC) for national coordination of Swedish biobanks is discussed in Chapter 10. BISC should be independent, but tightly linked to the existing organisation for databases at VR, DISC. Independent nodes perform the important functions needed and are derived from the outlined structure for the European infrastructure project, BBMRI. The details of the BISC organisation and its functions require further development. This effort should consider drawing on the organisational experience of the Swedish National Infrastructure for Computing (SNIC), apparently a functional organisation at VR with a similar structure.

b) Incentives should be explored for sharing the biological resources of biobanks and the data generated from them. As this is a general problem in science, ideas on how to promote openness and transparency could be drawn from several sources. Chapter 10 discusses this issue and presents ideas and suggestions. The problems are similar to those found in a recent investigation on databases for climate and environmental research⁴⁴. National incentive systems, harmonised with international systems, need to be developed.

c) Biobanks have a longitudinal structure and require long-term financing. Currently, financing is short-term and fragmented, placing the future use of biobank resources at risk. Maintaining valuable biobanks, and securing resources to utilise them, demands considerable funding over the long term. This requires a collaborative effort involving several sources, e.g. health care systems, universities, VR, private foundations etc. Analyses are needed on this issue.

d) To better understand current biobank-based research, we need to identify the population of researchers involved in biobank-related research in Sweden. This is important to better understand the importance of the area for medical research. There is a need to determine the resources that biobank-research utilises today, the amount of funding used, the impact of the research and the profiles of research projects.

⁴⁴ Eklund, L, "Data för svensk klimat- och miljöforskning", investigation done for VR 2007 (www.vr.se).

e) As discussed in Chapter 10, deeper integration is needed between the universities and the health care system. Different traditions and different primary objectives have created barriers between the universities and the health services. As biobank-based research involves several resources from the health care system in need of further development (i.e. major investments in construction and maintenance), there is an absolute need to share costs and expertise. University medical centres (UMCs) in the Netherlands offer a good example, as they represent a completely integrated organisation between the medical faculties and corresponding university hospitals.

Technical problems concern:

f) Linking and harmonising databases. Several international initiatives have addressed this issue, as discussed in chapters 8 through 10. Currently, the Biobank Information Management System (BIMS) developed by Swedish researchers appears to be the most advanced solution for this type of middleware system. However, it would probably be beneficial if the experts behind BIMS could work together with other computer and database experts from other faculties, DISC and experts working on IT solutions in health care. As discussed in Chapter 10, harmonisation of the ontology used in health care databases also needs to be addressed. As middleware and database harmonisation is crucial in both national and international collaboration on biobanks, the solutions must be developed on an international level. The P3G and BBMRI initiatives appear to coordinate in this regard.

g) Environmental data. Apart from the need for high-quality biomolecular analyses outlined in Chapter 7, a similar need exists for high-quality analyses of exposure data with which to correlate to enable high-quality epidemiological research. This has been addressed by NIH in the GEI (Genes, Environment and Health Initiative)⁴⁵ programme, which funds both genetic studies and Environmental Monitoring Technologies. This issue should be further investigated together with the Swedish Environmental Protection Agency, the National Food Administration and other analytical expertise at the universities.

An effort for national coordination has already taken place with the National Biobanking Program. Several of the universities with a medical faculty view biobank-based research as a future research area to investment in. Hence, the materials and the regional ambitions exist, presenting a good opportunity to establish and sustain national coordination. This is possible if all parties participate with good will and contribute in a transparent and open joint effort. National collaboration would decrease duplication of effort and make the best use of society's resources.

Having one national interface is important for effective international collaboration. Extensive international collaboration would probably be most effective if first extended on a Nordic level, as we share many features in the area. Several leading scientists believe that the Nordic countries would achieve a competitive niche in the global perspective if we were to work together more closely. Combined, the Nordic countries have a population around 24 million people, which is considered to be sufficiently large to acquire resources and form adequate sample populations. Collaboration is ongoing, but needs to be expanded – good examples to draw on include NCOEDG and the newly established EMBL nodes in molecular medicine. Any Nordic collaboration should not, however, be isolated from other collaborative initiatives, and could fit naturally in the framework of BBMRI.

The extensive public funding targeted at health care infrastructure should be utilised in the best possible way for research, ultimately to promote better health care. If the Swedish health care system happens to have an international advantage in pursuing biobank-related research, then we have a responsibility to develop this further to promote better research and health.

⁴⁵ <http://www.genome.gov/26022424>

Appendix 1 – Objectives of the Assignment



Vetenskapsrådet

Datum

2007-07-23

Diarienummer

811-2007-7597

Handläggare

Tove Andersson

Investigation of biobanks as a national resource for biomedical research

The investigation aims to illuminate, from a research perspective, the development potential and opportunities for Swedish biobanks now and in the future. The following questions shall be addressed:

- What are the important goals for biobank-related research in the next 5 to 10 years?
- How can the organisation of Swedish biobanks best support these goals?
- How important are large international initiatives in biobank-related research?
- How can the results of biobank-related research be utilised to achieve the greatest possible benefits?

The investigation is not intended to propose solutions, but to address the needs and potential problems. However, the investigator is free to propose any solutions that reveal themselves. The investigation should not be characterised as a survey.

The investigation should build on three sources:

- Presentations and input from the hearing held on September 18, 2007.
- Interviews with active researchers and agency representatives who are biobank users and/or manage biobanks.
- Viewpoints that submitted to the investigator through November 1, 2007.

Viewpoints may have been submitted directly to the investigator, to the appointed research officer at the Swedish Research Council, or via the Web forum designed for this purpose.

The investigator should assure that representatives of all relevant higher education institutions, agencies and research fields are given the opportunity to submit their viewpoints. The investigator shall determine who should be contacted, but this must be done in consultation with the appointed research officer and the reference group for the investigation.

The findings of this assignment shall be submitted as a report, written in English, to the Committee for Research Infrastructures and the Scientific Council for Medicine no later than November 23, 2007. The report shall specify those who were contacted and those who submitted viewpoints. The report should be formulated so as to clearly indicate that its findings are not the official views of the Swedish Research Council.

Tove Andersson
Research Officer

Appendix 2 – Reference Group

Åke Lernmark, University of Washington and Lund University

Juni Palmgren, Karolinska Institutet and Stockholm University

Dan Holmberg, Umeå University

Taina Pihlajaniemi, University of Oulu

Kristian Hveem, Norwegian University of Science and Technology

Appendix 3 – Summary of Questionnaire

Several of the questions were open questions. Hence, the responses are presented by dividing them into categories and presenting some descriptive answers/comments. The responses have been translated and condensed. They are not citations.

Q1 Do you currently perform research on samples from biobanks with human material?

68% Yes, 32% No

Q2 What are your hopes/goals/plans within biobank-related research on a time scale of 5-10 years?

Accessibility: Increase the information on biobanks and their contents, develop a national biobank registry, make biobanks available for uses other than the ones they were intended for, easier to reach larger bodies on material.

Laws and regulations: Simpler laws and regulations, fewer bureaucratic barriers against usage, revise the Biobank Act, international harmonisation.

Funding issues: Need for long-term funding to maintain biobanks and to allow larger new collections.

Standardisation and Quality: Standardise the design of collections to allow pooling and “superstudies”, national quality standards.

Biobanks as resources: Increased awareness of the major resource that biobanks represent, increased usage by most disciplines, use in clinical trials, confirm experimental data on biobanks, use to allow personalised medicine.

Registries: Allow the coordinated use of different health registries (linking and matching files), increase the accessibility of health registries.

International collaborations: Better coordination between biobanks internationally, simplify regulations in this context.

Advance with my research: A fantastic resource, continue with my research.

Q3 Do you see any problems that make your plans/goals in Q2 difficult? If not continue to Q 6.

40% Yes (of all who answered), 44% Yes (of biobank researchers, who answered yes in Q1).

60% No (of all). 54% No (of biobank researchers, who answered yes in Q1).

Q4 Which hinders/problems do you see (apart from too small project grants) in Q3? Give detailed answer!

Lack of resources at clinics: Clinics keeping biobanks (e.g. pathology) need more resources to handle requests, deliver samples, analyse tissues (pathologists are limited). Resources to collect samples are missing.

Laws, regulations: Who has the access rights to biobanks? Complicated and unclear regulations on biobank usage, obtaining consent. Hard to get approval from ethics review committee. Complicated to move samples between biobanks. Hard to use old biobanks for new research questions. Our regulations make it hard for international collaborations.

Funding issues: Long-term funding lacking. Short-sighted funding through project grants and by the health services (clinics). Uncertain funding for maintaining biobanks. Impossible for a single research group to raise money to collect biobanks due to the increased requirements on traceability, data management, biostatistics etc. Expensive analysis with modern genomics means that Swedish biobanks serve only as sample and data deliverers with all credit/knowledge going to USA and UK. Need to “refine” our biobanks before it’s too late. Lack funding for a Swedish-owned reference population for genome-wide scans – would be valuable for genetics in complex diseases.

Information: More information on current biobanks and their linked phenotypic data. What analysis can the samples be used for? PKU biobank is not computerised – very expensive to conduct investigations.

Q5 What solutions do you see on the problems you gave in Q4? Give a detailed answer!

Resources: Increase the resources at pathology departments to serve researchers with biobanks samples. Create incentives for suppliers at pathology clinics in form of collaborations.

Laws, regulations: Revise the Biobank Act. Harmonise first with European laws. Make scope of ethics review approvals wider to allow different studies without new permission. Start a central helpdesk having the knowledge and resources to guide researchers on all related matters.

Infrastructures: Need to make common infrastructure investments in IT development, quality documentation (to meet international standards), investments for standardised pre-analysis handling (robotised etc). Consider the biobank sample as part of the infrastructure! We are collaborating to collect genotyped control samples to build a Nordic reference population.

National coordination: Let SRC lead the national coordination of biobanks. Create a versatile national biobank registry. Feed back research data to the biobank registry.

Openness: Create incentives for the PI to share his/her biobank samples. Discuss the access-rights – time limited? The PI should have the control – have much work invested. Should be common rules for access to biobanks in all Nordic countries.

Linkage to registries: Construct a simple IT/software system to link biobank samples to clinical registries.

Q6 What type of biobank resource would you like to use in your future research?

1. Clinical biobanks in the health care system

67% (of all), 77% (of biobank researchers)

2. Population-based research biobanks

46% (of all), 58% (of biobank researchers)

Q7 Motivate your answer in Q6!

Complementary: Different biobanks reach different parts of the whole population. Research in personalised medicine needs both types (CB and RB). You need both: RB are needed to evaluate the significance in findings from CB. Common study design compares findings from RB with clinical cases from CB.

Clinical Biobanks (CB): Needed for clinical trials by drug companies. I need pathology samples. The studied phenomena do not occur in a healthy population. No experience of biobanks other than CB. Easier to communicate with other clinics than research institutions. CB needed for diagnostic development, pharmacogenetics, and research on treatments. CB needed for tissue-dependent research. Research on transplants needs CB. I need CB since I need tumour tissue + control tissue from the same patient. I want to correlate my preclinical research findings to clinical patient samples.

Population-based Research Biobanks (RB): RB is more important to understand the importance of environmental factors for disease. Factors observed *before* disease develops very important to know and found only in RB. RB necessary to be able to correlate genotype – exposure – disease – possibilities of prevention. RB needed to compare disease groups with healthy groups. Samples are seldom taken in the psychiatric clinic – RB necessary as source for DNA. RB necessary for epidemiologic research. You miss co-morbidity if you *only* select for one disease as in CB, therefore is better with RB.

Q8 Will you, in your future research, primarily use:**1. Existing biobanks**

68% (of all), 81% (of biobank researchers)

2. Development of new

40% (of all), 45% (of biobank researchers)

Q9 Motivate your answer in Q8!

Both: We have ongoing collection at our existing biobank. We are building a reference population from existing biobanks. Pharmaceutical industry needs both EB and NB. Valuable archive material exists, but I plan to collect new. I follow up earlier cohorts, but want to start new. Advantage if you can use samples in EB to minimise collection of NB. Analysis of EB yields new questions – NB needs to be collected.

Existing biobanks (EB): Don't have time to make NB. Hard to start NB; low frequency of participants also in Sweden yields no representation of the population and a long time before they can be used (until cases appear). Difficult to start NB! Material already exists. The EB are Sweden's big advantage. An old EB has the necessary clinical follow-up. Extensive material collected over a long time is what I need.

Develop new biobanks (NB): Material for my research does not exist. Important documentation on routines for sample handling is missing in most EB, especially important for protein detection. Necessary to have prospective biobanks with repeated sampling for biomarker research. Larger biobanks with fresh frozen tissue do not exist. Very few biobanks in psychiatry.

Q10 How do you think one should organise Swedish biobanks to make the best use of them for research?**1. Centrally on a national level**

56% (of all), 55% (of biobank researchers)

2. Locally on the regional level

42% (of all), 50% (of biobank researchers)

Q11 Motivate your answer in Q10!

Both: Different types of biobanks at different levels; epidemiological as CNB and clinical as LB. Biobanks for specific research projects as LB and larger comprehensive projects as CNB. Start to organise biobanks as LB, continue over time as CNB. Organise biobanks as LB, connect them in a national network. For tissue biobanks needed LB to avoid variation, but central standardisation.

Central National Biobank (CNB): Our reference population must be representative for all of Sweden and the Nordic countries. Need to have CNB to be internationally competitive. CNB needed to have common standards for quality and accessibility. Move some LB to CNB for increased efficiency. Samples will be accessible for a larger group. Centralisation will save resources, less duplication. Saves administration. CNB for larger projects, linked LB for smaller projects. CNB can easier give support with more resources and knowledge. Genetic research demands CNB. We need CNB infrastructure with a common ontology. Major resources needed to keep a biobank, centralisation will increase the quality. Local organisational responsibility with central coordination to give same standards everywhere.

Local Biobanks (LB): Problem with CNB is risk for increased power to bureaucrats and agencies. CNBs will give advantage to certain researchers. LB better – closer to the researcher. Better accessibility! Increased bureaucracy with CNB. LB important for tissue biobanks. Better for logistic reasons to have LB. Close to your biobank means better

service! Current research questions do not justify a national principle. CNB has risk for increased influence from other actors (politicians, authorities).

Q12 What importance do (larger) international initiatives have for biobank-related research and for how one organises biobanks in Sweden?

Large studies: Greater statistical power for projects. Large cooperative studies can also be sources for error in the data as biobanks are constructed differently. More important for influential researchers than for researchers who created the biobanks. Important to reach right expertise in other countries. Large new cohorts (UK biobank etc) are questionable; low frequency of participants means no representation of underlying population. Large international studies important to understand the impact of (genetic) variations between different populations and the importance of different environmental factors. Large international studies needed to study rare diseases. Very important for my research to access international samples. Makes it possible to become part of large and expensive studies, e.g. GWAS.

Common regulations/laws: International standards needed for quality and harmonisation to enable large studies. Different legislation makes exchange of biobank samples difficult. Our national access must be good to allow international collaboration. International demands on quality, ontology and generic datasets is necessary for us to adjust to. Difficult to handle different systems for regulation and research cultures. Ethical problems when Swedish samples are part of new, large international collaborations. NIH's demand to make individual data accessible for other researchers is a problem in collaborating with NIH.

Structure in Sweden: Internationalisation will speed up centralisation and specialisation in Swedish biobanking. Good quality biobanks demand large resources, which demand international collaboration and structure. International collaborations works fine without national coordination, already several examples.

Unique opportunities: Sweden has the personal ID number and good order, making my research interesting for funding from USA. International agreements should not hinder Swedish initiatives – e.g. many countries do not have possibility to link biobanks to health registries.

Q13 What importance could your future biobank-related research have for society?

Economy: Stopping the obesity epidemic should improve the economy of society and quality of life. Important for individualised medicine, which also is important for biotech industry. Our research strengthens biomedical research in Sweden and applications in the biotech/pharma industry. Decreased costs for health care and decreased suffering if hypertonic and diabetes related to pregnancy is diagnosed and treated earlier. Better prognosis of consequences of diseases and of response to therapy both has health economic effects. Saves costs of long-term care due to infections and inflammations if we can find new therapies.

Knowledge: Our biobank research is the only way to solve the genetic background of complex common diseases. Special biobank with breast and colorectal tumours – prerequisite for knowledge of the pathology of these diseases and development of treatments in radiology, surgery and oncology. Knowledge on the aetiology of common neuropsychiatric diseases requires access to a neuronal tissue biobank. Important for generation of knowledge on risk factors for psychiatric diseases and abuse. Important for biomarker research. Important for research on arthritis. Important to explain common multifactor diseases. Gives important new knowledge for cancer treatment. Faster verification of animal models. Verification on human samples of mechanisms behind brain damage derived from animal models is extremely important in development of neuroprotective drugs. Population cohorts together with the demographic database lead to more knowledge on welfare-related diseases.

Improved health care: Biobanks are critical for developing new diagnostics and new drugs. Development of personalised medicine. Prediction of risk factors leading to orthopaedic surgery. Preventive actions against asthma and allergy. Biomarkers for mental illness and stress-related problems. New principles for cancer treatment by using biobanks together with experimental systems. Research on the PKU biobank important for developing new screening methods on children with inherited diseases, which can lead to treatments to stop progression of the disease.

New treatments: Personalised medicine. Individualised cancer treatments. Clinical trials yield new drugs. To decrease and prevent tumour formation in the liver, also important in countries where liver cancer is more common. More effective treatments against Alzheimer's disease. Analysis of efficiency of drug treatments for diabetes.

Q14 Could the results from this type of research give improved health care? If so, how?

New therapies: Individualised medicine leads to less suffering and costs. Decrease unnecessary medication. Speeds up the drug development process. Utilising human samples early will speed the development of new and safer drug candidates. The studied defect is gene-regulatory by nature; if we could understand it, it could lead to the expression of the patients' own genes and thus production of endogenous antibodies. Individual treatments: large variation in the onset of disease. The importance of the genotype for psychiatric treatments is important. Neuroprotective drugs missing for traumatic brain damage. Avoid side effects. Molecular knowledge on disease processes helps in the long run to develop new drugs.

Diagnosis-prognosis: Better possibilities for prognosis. New factors valuable for diagnosis and prognosis of stress-related illness. We have identified new disease markers, new antibodies. Molecular diagnostics yield a better-differentiated diagnosis, which guides therapy better. Improved diagnosis. Better diagnosis and therapy of preclampsy and coronary heart disease.

Prevention: Identification of patients at risk for parodontitis and coronary heart disease. Map the causes for type 1 diabetes and learn how to prevent it.

Knowledge: Biobank research makes important contributions to international knowledge, but translational research is the true bottleneck for progression. Knowledge on factors behind "Developmental origins of health and disease".

Q15 Is it conceivable that the results from your future biobank-related research can be commercialised? If so, how?

New therapies: From our research we already have several patents being commercialised. Our goal is to deliver innovative drugs and diagnostics. Our developed methods are used already by several large biotech companies, we hope to collaborate more with pharmaceutical companies in the future. We are commercial already, have only company-sponsored clinical trials. Within the area of IVF this has been done since the 1970s. Possible to search after anti-obesity drugs. Yes, new anti-microbiology, anti-inflammatory and immunosuppressive drugs are conceivable.

Diagnostics – biomarkers: Biomarkers can be commercialised leading to new diagnosis. We have developed new methods for isotype-specific diagnosis of autoantibodies. New cancer markers to aid in choice of therapies. Development of biomarkers for neoplasia and preneoplasia as risk factors.

Technologies: Develop new diagnostic tests including both sampling and analysis, utilising nanotechnology and computer science. I am developing new technology for surveillance of dialysis in real time, which could be used to develop new dialysis machines.

Commercialisation of biobanks: Unethical to commercialise biobanks. Attempts have been made in commercial sampling of cord blood, which has met hard resistance and stopped. To sell samples from a biobank to a company might be OK if intended for well-defined and acceptable use. I hope that my biobank will not be commercialised.

No: Cannot be judged today! Public health research has few commercial interests. Not possible as there are many established, well-working drugs today – mainly a question of dosage.

Appendix 4 – Participants

Interviewed

Hans-Olov	Adami	Department of Epidemiology, Harvard School of Public Health
Kjell	Asplund	The National Board of Health and Welfare
Ingvar	Bergdahl	Dept. Environmental Medicine, Umeå University
Åke	Borg	Department of Oncology, Lund University
Johan	Botling	Department of Genetics and Pathology, Uppsala University Hospital
Ingemar	Carlstedt	Assistant Dean, Faculty of Medicine, Lund University
Jan	Carlstedt-Duke	Dean for Medical Research, Karolinska Institutet
Joakim	Dillner	Dept. Laboratory Medicine, Lund University
Göran	Elinder	National Biobank Council, SALAR
Lennart	Eriksson	Dept. of Laboratory Medicine, Karolinska Insitutet
Henrik	Grönberg	Department of Medical Epidemiology and Biostatistics Karolinska Institutet
Göran	Hallmans	Dept. of Public Health, Umeå University
Mats	Hammar	Dean, Faculty of Health Sciences, Linköpings University
Britta	Hedlund	Division for Environmental Analysis, Swedish Environmental Protection Agency
Sophia	Hober	Dept of Proteomics KTH/AlbaNova University Centre
Ellen	Hyttsten	Swedish Association of Local Authorities and Regions
Peter	James	Protein technology, Lund University
Gerd	Johansson	Dept. of Public Health, Umeå University
Juha	Kere	Dept. of Biosciences, Karolinska Institutet
Ulf	Landegren	Department of Genetics and Pathology, Uppsala University
Lars	Lannfelt	Department of Public Health and Caring Sciences, Uppsala University
Olle	Larkö	Dean, Sahlgrenska Academy, Göteborg University
Thomas	Laurell	Department of Electrical Measurements, Lund University
Jan-Eric	Litton	Department of Medical Epidemiology and Biostatistics Karolinska Institutet
Joakim	Lundeberg	School of Biotechnology, Royal Institute of Technology
György	Marko-Varga	Respiratory Biological Sciences AstraZeneca R&D Lund
Nancy	Pedersen	Department of Medical Epidemiology and Biostatistics Karolinska Institutet
Leena	Peltonen	Human Genetics, The Wellcome Trust Sanger Institute
Ulf	Pettersson	Chair Medical Faculty ,Uppsala University
Camilla	Stoltenberg	Division of Epidemiology, The Norwegian Institute of Public Health
Lars	Terenius	Dept. of Clinical Neuroscience, Karolinska Insitutet
Elvar	Theodorsson	Dept. of Neurochemistry, Linköping University Hospital
Gunnel	Tybring	Department of Medical Epidemiology and Biostatistics Karolinska Institutet
Gert-Jan	Van Ommen	Centre for Human and Clinical Genetics, Leiden University Medical Centre
Lars	Wallentin	Department of Medical Sciences, Uppsala Universitet
Kurt	Zatloukal	Institut für Pathologie, Medizinische Universität Graz
Åsa	Ågren	Dept. of Public Health, Umeå University

Filled in the Questionnaire

Anders	Ahlbom	KI
Thor	Alvegård	Region Skåne
Tove	Andersson	CSL Behring AB
Sten-Magnus	Aquilonius	UU
Bo	Baldetorp	LU
Christina	Bark	KI
Lisbeth	Barkholt	Tobias Registret
Mikael	Benson	Sahlgrenska Akademin, GU
Lars	Benson	Luthagens Specialistmottagning AB
Göran	Berglund	LU
Peter	Bergsten	Landstinget i Uppsala län
Tomas	Bergström	GU
Sven	Bergström	UmU
Anna	Beskow	KI
Anna	Blom	LU
Lotta	Bodin	Landstinget Sörmland
Bertil	Borgencrantz	Capio Läkargruppen
Carl	Borrebaeck	LU
Gunilla	Bratthall	Malmö högskola
Tomas	Bremell	Västra Götalandsregionen
Anders	Brinne	Karolinska Institutet Biobank
Jonas	Broman	KI
Johan	Brun	Pfizer AB
Lars	Bäckman	Västra Götalandsregionen
Anette	Carlén	Sahlgrenska Akademin
Lena	Carlsson	GU
Sandra	Ceccatelli	KI
Björn	Cederin	Västra Götalandsregionen
Ing-Marie	Claesson	Landstinget i Östergötland
Gisela	Dahlqvist	UmU - Nationella biobanksrådet
Jan-Erik	Damber	Sahlgrenska Akademin, GU
Elisabeth	Darj	Landstinget i Uppsala län
Barbro	Diderholm	UU
Joakim	Dillner	LU
Sonja	Eaker	Chef RBC Uppsala/Örebro. Forskare Akademiska
Anders	Ekbom	KI
Marinne	Ellman	Active Biotech Research AB
Jan-Erik	Elverby	Västra Götalandsregionen
Stefan	Emdin	Västerbottens läns landsting
Kerstin	Engholm	Västra Götalandsregionen
Carl-Göran	Ericsson	Stockholms läns landsting/Stockholms Regionala
Henry	Eriksson	Biobankscentrum
Lennart	Eriksson	Västra Götalandsregionen
Sture	Eriksson	KI
Charli	Eriksson	UmU - Nationella biobanksrådet
Karin	Eriksson	Örebro Universitet
		Läkemedelsindustriföreningen - Nationella

		biobanksrådet
David	Erlinge	LU
Magnus	Fall	Västra Götalandsregionen
Maria	Feychting	KI
Pia	Forsberg	Landstinget i Östergötland
Jan	Forslid	Stockholms Regionala Biobankscentrum
Johan	Franck	KI
Hans	Fredlund	Örebro läns landsting
Barbro	Fridén	Fertilitetscentrum AB, Stockholmskliniken
John Eric	Frisell	Landstinget Dalarna
Gun	Frisk	UU
Gabriela	Godaly	Gambro Lundia AB
Leif	Groop	LU
Ludger	Grote	Consleep AB, Sömnlaboratoriet Carlanderska
Ludger	Grote	Västra Götalandsregionen
Anders	Gustafsson	KI
Ulf	Gyllensten	UU
Bertil	Hamberger	KI
Anne	Hammarström	UmU
Lennart	Hammarström	KI
Peter	Hansell	UU
Gunnar C	Hansson	GU
Per	Hansson	Västra Götalandsregionen
Börje	Haraldsson	Västra Götalandsregionen
Eva	Hardmeier	Täby Kommun
Mikael	Hasselgren	Landstinget i Värmland
Hans	Hedenström	Landstinget i Uppsala län
Karin	Hellström	InDex Pharmaceuticals AB
Ewa	Hellström	Landstinget i Östergötland
Nels	Henningsen	Cityvården Gustav Adolf Läkargrupp
Leif	Hernefalk	Landstinget Dalarna
Olle	Hernell	Västerbottens läns landsting
Lars	Holmgren	KI
Outi	Hovatta	KI
Vuokko	Höglin	SBL Vaccin AB
Peter	Höök	Medicare
Annica	Inerot	Västra Götalandsregionen
Bo	Jacobsson	Västra Götalandsregionen
Peter	James	LU
Per-Anders	Jansson	Västra Götalandsregionen
Ulf	Jansson	Landstinget Västernorrland
Anders	Johansson	Landstinget i Östergötland
Ingibjörg	Jonsdottir	Västra Götalandsregionen
Bertil	Kaijser	Västra Götalandsregionen
Anna	Karlsson	Medicin
Magnus	Karlsson	Region Skåne
Lars	Karlsson	Västra Götalandsregionen
Diana	Karpman	LU
Lena	Kjellen	UU
Elisabeth	Kjellén	Region Skåne
Torbjörn	Kjerstadius	Landstinget i Värmland
Sonja	Klingén	Västra Götalandsregionen
Anna	Krook	KI
Johan	Kärholm	Västra Götalandsregionen

Göran	Landberg	LU
Ulf	Landegren	UU
		GU dekan Sahlgrenska Akademin - Nationella biobanksrådet
Olle	Larkö	Sahlgrenska Akademin
Sven	Larsson	Västra Götalandsregionen
Anders	Lasson	Västerbottens läns landsting
Per	Lenner	UU
Anders	Lewén	KI
Tülay	Lindberg	Regionala etikprövningsnämnden i Lund
Göran	Lingman	KI
Jan-Eric	Litton	Västerbottens läns landsting
Börje	Ljungberg	Landstinget i Uppsala län
Östen	Ljunggren	AstraZeneca R&D Headquarters
Sverker	Ljunghall	LU
Stefan	Lohmander	Kirurgkliniken vid Sophiahemmet
Göran	Lundegårdh	Landstinget i Uppsala län
Gudmar	Lönnnerholm	Västra Götalandsregionen
Peter	Lönnroth	Kungsgårdets VC, Luleå
Gulliksson	Mats	SLU
Sofia	Mikko	Region Skåne
Lennart	Minthon	Nycomed AB
Mats	Mogard	Dynamic Code AB
Catrin	Molander	Landstinget Dalarna
Frank	Niklasson	GU
Mikael	Nilsson	KI
Monica	Nistér	Västra Götalandsregionen
Lars	Nistor	Landstinget i Östergötland
Conny	Nordin	Västra Götalandsregionen
Urban	Norén	Smittskyddsinstitutet
Ragnar	Norrby	Landstinget i Uppsala län
Matts	Olovsson	Region Skåne
Bertil	Olsson	KI
Nancy	Pedersen	Sahlgrenska Universitetssjukhuset
Katarina	Peltz	Västra Götalandsregionen
Lennart	Persson	HPR Teknik AB
Per	Pettersson	LU
Eeva	Piitulainen	Lundby Sjukhus
Erik	Pileblad	Novo Nordisk Scandinavia AB
Marianne	Pilgaard	GU
Marianne	Quiding-Järbrink	Västerbottens läns landsting
Solbritt	Rantapää-Dahlqvist	Landstinget i Östergötland
Karin	Roberg	Biovitrum AB
Elin	Rosendahl	Landstinget i Östergötland
Ulf	Rosenqvist	KI
Tove	Rylander Rudqvist	UU
Lars	Rönblom	KI
Karin	Schmekel	TikoMed AB
Peter	Schmidt	UU
Tobias	Sjöblom	GU
Henrik	Sjövall	Landstinget i Uppsala län
Alkistis	Skalkidou	UU
Ingiridur	Skírnisdóttir	Landstinget i Östergötland
Thomas	Skogh	Landstinget i Jönköpings län
Christer	Slotte	

Torgny	Stigbrand	UmU
Magnus	Ström	Landstinget i Östergötland
Mikael	Sundfeldt	Västra Götalandsregionen
Christer	Sundström	Landstinget i Uppsala län
Marie	Sverud	Regionalt biobanksregister, Södra sjukvårdsregionen
Ann-Christine	Syvänen	UU
Torbjörn	Söderström	Landstinget i Uppsala län
Anders	Tegnell	Socialstyrelsen
Ulf	Théen	Landstinget i Uppsala län
Ulf	Tidefelt	Örebro läns landsting
		Stockholms läns landsting/Stockholms Regionala
Eva	Tillman	Biobankscentrum
Tibor	Tot	Landstinget Dalarna
Gunnar	Tufveson	Landstinget i Uppsala län
Torsten	Tuvemo	Landstinget i Uppsala län
Gunnel	Tybring	KI
Hans	Törmä	Landstinget i Uppsala län
Agneta	Törning	CRO filial till PharmaNet Services GmbH
Fredrik	Uhlin	Landstinget i Östergötland
Göran	Wadell	UmU
Torkel	Wahlin	Västra Götalandsregionen
Kurt	Wahlstedt	Wahlstedt Medicinsk Konsult AB
Lars	Wallentin	Landstinget i Uppsala län
Bengt	Wallner	UU
Lars	Wesslen	Landstinget Gävleborg
Per	Westermarck	Landstinget i Uppsala län
Anna	Winkvist	Sahlgrenska Akademin, GU
Ola	Winqvist	Sentoclone AB, Sophiahemmet
Neus	Visa	SU
Ulrika	von Döbeln	Karolinska Univeristetssjukhuset
Christian	Åkermark	Sport Medical Centre - Sport Med AB
Ann-Kristin	Öhlin	Universitetssjukhuset MAS
Agneta	Öjehagen	LU

Appendix 5 – National Healthcare Quality Registries in Sweden

National Healthcare Quality Registries in Sweden – 2007

Respiratory Diseases

Swedevox – Respiratory Failure Registry
Swedish Quality Register of Otorhinolaryngology

Childhood and Adolescence

BORIS – Childhood Obesity Registry in Sweden
SWEDIABKIDS – The Swedish Childhood Diabetes Registry
PNQn – Perinatal Quality Registry / Neonatology
SÖK – National Registry of Suspected/Confirmed Sexual Abuse in Children and Adolescents

Circulatory Diseases

RiksSvikt – Heart Failure Registry
SCAAR – Swedish Coronary Angiography and Angioplasty Registry
Swedish Heart Surgery Registry
RIKS-HIA – Registry on Cardiac Intensive Care
National Registry on Out-of-Hospital Cardiac Arrest
Riks-Stroke – National Quality Register for Stroke
Swedvasc – Vascular Registry in Sweden
GUCH – Grown-Up Congenital Heart disease Registry
SEPHIA – Registry on Secondary Prevention in Cardiac Intensive Care
National Catheter Ablation Registry
Auricula – National Registry of Atrial Fibrillation and Anticoagulation

Endocrine Diseases

NDR – National Diabetes Registry
Scandinavian Quality Register for Thyroid and Parathyroid Surgery
SOREg – Swedish Obesity Surgery Registry

Gastrointestinal Disorders

Swedish Hernia Registry
GallRiks – Swedish Quality Registry on Gallstone Surgery
Swedish Quality Registry for Ventral Hernia

Musculoskeletal Disorders

RIKSHÖFT – National Hip Fracture Registry
Swedish National Hip Arthroplasty Register
Swedish Knee Arthroplasty Register
National Pain Rehabilitation Registry
Swedish Rheumatoid Arthritis Registry
Followup in Back Surgery
Swedish Shoulder Arthroplasty Registry
Swedish Cruciate Ligament Registry – X-base
Swedish National Elbow Arthroplasty Register (SAAR)

Diseases of the Nervous System

SMS – Swedish Multiple Sclerosis Registry
CPUP – Quality Registry for Children with Cerebral Palsy
WebRehab Sweden – Quality Registry in Rehabilitation Medicine
SveDem – Swedish Dementia Registry

Genitourinary Disorders

GYNOP – National Quality Registry for Gynecological Surgery
SRR – Swedish Renal Registry

Cancer

National Prostate Cancer Registry
National Breast Cancer Registry
National Quality Registry for Esophageal and Stomach Cancer (NREV)
Swedish Rectal Cancer Registry
Swedish Gyn-Oncology Registry
Swedish Colon Cancer Registry

Eye Disorders

Swedish National Cataract Register
Swedish Corneal Transplant Register
Macula Register

Other

RIKSÄT – National Quality Registry for Specialized Treatment for Eating Disorders
SIR – Swedish Intensive Care Registry
PsoReg – Swedish Psoriasis Registry
InfCare HIV
Swedish Therapeutic Apheresis Registry
SKaPa – Swedish Quality Register in Caries and Periodontitis
Swedish National Registry of Palliative Care
Senior Alert – National Registry on Nutrition, Fall Prevention, and Pressure Sores
Quality Registry for Emergent Care

Appendix 6 – List of Abbreviations

BBMRI	Biobanking and Biomolecular Resources
BIMS	Biobank Information Management System
BISC	Biobank Infrastructure Committee
DISC	Database InfraStructure Committee
ELSI	Ethical Legal Societal Issues
GU	Göteborg University
KI	Karolinska Institute
KTH	Royal Institute of Technology
LiU	Linköping University
LU	Lund University
NBP	National Biobank Program
SALAR	Swedish Association of Local Authorities and Regions
SCB	Statistic Sweden
SLU	Swedish University for Agricultural Sciences
SU	Stockholm University
UmU	Umeå University
UU	Uppsala University
VR	Swedish Research Council
WHO	World Health Organization