RESEARCH ARTICLE

Standard based Deposit Guideline for Distribution of Human Biological Materials in Cancer Patients

Hwa Jeong Seo[1], Hye Hyeon Kim[2&], Jeong Soo Im[3], Ju Han Kim[2*]

Abstract

Background: Human biological materials from cancer patients are linked directly with public health issues in medical science research as foundational resources so securing "human biological material" is truly important in bio-industry. However, because South Korea's national R and D project lacks a proper managing system for establishing a national standard for the outputs of certain processes, high-value added human biological material produced by the national R and D project could be lost or neglected. As a result, it is necessary to develop a managing process, which can be started by establishing operating guidelines to handle the output of human biological materials. Materials and Methods: The current law and regulations related to submitting research outcome resources was reviewed, and the process of data 'acquisition' and data 'distribution' from the point of view of big data and health 2.0 was examined in order to arrive at a method for switching paradigms to better utilize human biological materials. Results: For the deposit of biological research resources, the original process was modified and a standard process with relative forms was developed. With deposit forms, research information, researchers, and deposit type are submitted. The checklist's 26 items are provided for publishing. This is a checklist of items that should be addressed in deposit reports. Lastly, XML-based deposit procedure forms were designed and developed to collect data in a structured form, to help researchers distribute their data in an electronic way. Conclusions: Through guidelines included with the plan for profit sharing between depositor and user it is possible to manage the material effectively and safely, so high-quality human biological material can be supplied and utilized by researchers from universities, industry and institutes. Furthermore, this will improve national competitiveness by leading to development in the national bio-science industry.

Keywords: Biobank - human biological materials - deposit - cancer - XML

Asian Pac J Cancer Prev, **15 (14)**, 5545-5550

Introduction

The term biobank has been defined in many ways but the definition adopted here will be an organized collection of human biological material and associated information stored for one or more research purposes. Human biological materials (human biospecimens and related cancer patients) are a vital component of a biobank and support many types of contemporary research like genomic personalized medicine, and translational bioinformatics studies. A biobank is organized to manage these resources throughout process that is conducted from deposit to termination (Watson, 2010; Denny, 2010; Wolf, 2012).

In South Korea, the National Biobank of Korea (NBK) (KBN, 2013) has performed this role of collecting, preserving, and distributing a large number of high quality biospecimens and their related information. NBK has collected and secured various high quality human biospecimens from population-based participants and disease-based participants through the Korea Biobank Project (KBP). Recently the first phase of KBP finished and accomplished its aims to collect human specimens from 500,000 participants with epidemiological and clinical information. Unlike successful securing biospecimens from the KBP, however, securing biospecimens and the outcomes of the national R&D projects of individual researchers has been difficult; as most researchers prefer to collect human biospecimens in their own data repositories. As some national R&D projects are related with certain disease or certain biospecimens, outcome biospecimens from this research cannot be ignored (Yoo, 2005a; 2005b; Lee, 2012).

The major cause of this issue among several other reasons is that current operating guidelines are old fashioned. The NKB's guidelines have several drawbacks. First, the procedure is too simplified for practical use, as there is only one deposit application form. It confuses

*[1]Medical Informatics and health Technology (MIT), Department of Healthcare Management, College of Social Science, Gachon University, Seongnam, [2]Seoul National University Biomedical Informatics (SNUBI), Seoul National University College of Medicine, Seoul, [3]Department of Preventive Medicine, Gachon University, Incheon, Korea, [&]Equal contributors *For correspondence: juhan@snu.ac.kr*

depositors and results in different responses. Second, the current guidelines do not mention the ownership of biospecimens or whether they resulted from a government funded project, though several laws that relate to this issue have been amended. Third, the entire deposit procedure is performed in a traditional paper-based process. Not only are paper-based processes difficult to efficiently secure, monitor, and measure, but they also are more time-consuming, error-prone, and costly than electronic processes (Ivan, 2009).

To solve those problems, this paper proposes a modified guideline that includes a specified procedure of deposit process with an XML standard based deposit form and a data schema for an electronic data process and communication.

Since 2003, the guidelines for biological resource management have been developed based on the federal law of South Korea. The initial guidelines detailed the required description as well as the receiving confirmation form and standard operating procedures (SOPs), which are detailed written instructions for five representative biospecimens. However, the initial guidelines only covered limited biospecimens, which is problematic.

For a simplified deposit procedure, especially, at least two interactions with four different types of forms such as a submitting deposit form and a receiving deposit permission reply (conforming or rejecting) to share deposit information between a depositor and an administrator in the KBN would be preferable, rather than having only one deposit form.

While revising the guidelines, the issue of collecting rich resources has also been raised. Basically the KBN acquires human resources periodically in cooperation with 13 regional biobanks (Lee, 2012), but rich human samples, the research results from government funded projects, have been overlooked and have not been gathered properly. In addition, the South Korean government established the law of 'regulation on management of national research and development project, etc.' (Ministry of Health and Welfare, ICT and Future Planning, 2011) to handle the above limitation.

Materials and Methods

We developed the deposit guideline for a systematic literature review by considering regulations, ownership of national R&D outcomes, biobanks in global status and big data (Hyun, 2013).

Relevant regulations

A biobank is a resource of various biospecimens and clinical data that are collected from national R&D projects, and provide research infrastructure to researchers for vitalizing the bio-medical industry. Several relevant regulations were reviewed in order to look into the sources of national R&D outcomes, and the deposit process that in the relevant regulations.

The purpose of health and medical service technology promotion ACT is to contribute to the development of the health and medical service industry and the promotion of national health by establishing a basic plan for the promotion of health and service technology as well as performing research and development activities (Ministry of Health and Welfare, 2011).

The purpose of ACT on development, management and utilization of biological resources is to ensure the efficient development and systematic management of biological resources that will lead to the sustainable utilization thereof as well as to build infrastructure for the development of biotechnology (Ministry of Health and Welfare, ICT and Future Planning, 2011).

Ownership of national R&D outcomes

The South Korean law and regulations pertaining to the ownership of national R&D outcomes, including the health and medical service technology promotion ACT (Ministry of Health and Welfare, 2011) as well as the regulation on administration of national research and development project, etc (Ministry of Health and Welfare, ICT and Future Planning, 2011) mandate that the ownership of outcomes of national research and development projects is determined for in accordance with the initial contract. However, for the purpose of public interest, it can be owned by the government.

To promulgate the above information, the guidelines must contain a statement that involves specific information of the related law and regulations. The deposit form should contain the type of deposit, whether it is a R&D outcome deposit or a general deposit, and whether the reason for the deposit was a request from a biobank or if it was a voluntary deposit from a donor.

Departmentalization of deposit procedure

Korean law for managing biological resources, known as Act on development, management and utilization of biological resources (Ministry of Health and Welfare, 2011) mandate that the essential procedure of deposit be identified as such. To practical use, we modified the deposit guideline as departmentalizing the deposit procedure and adding relative forms according to other international guidelines. As described below, we incorporated into our deposit procedure with appropriate forms required by the law, as well as those recommended by best practice guidelines for biobanking (National Institutes of Health, 2013).

Biobanks in global status

The EuroBioBank network is the first operating network of biobanks in Europe providing human DNA, cell and tissue samples as a service to the scientific community conducting research on rare diseases. It is the only network dedicated to rare disease research in Europe. About 100,000 samples are available across the network and can be requested via the online catalogue (EuroBioBank, 2014).

UK Biobank is a major national health resource, and a registered charity in its own right, with the aim of improving the prevention, diagnosis and treatment of a wide range of serious and life-threatening illnesses – including cancer, heart diseases, stroke, diabetes, arthritis, osteoporosis, eye disorders, depression and forms of dementia. UK Biobank recruited 500,000 people aged

between 40-69 years in 2006-2010 from across the country to take part in this project (UK Biobank, 2014).

The largest Swedish biobanks are those that are collected during routine medical care. The total number of samples in the biobanks of the Swedish Health Care system is estimated to about 50-100 million human samples, increasing with about 3-4 million samples per year. Another 985 000 new samples are collected each year during the cervical cytology screening (Swedish Biobank, 2014).

International Society of Biological and Environmental Resources (ISBER) is a global organization which creates opportunities for sharing ideas and innovations in biobanking and harmonizes approaches to evolving challenges for biological and environmental repositories. ISBER fosters collaboration, creates education and training opportunities, provides an international showcase for state-of-the-art policies, processes, and research findings, and innovative technologies, products, and services. ISBER will be the leading global forum for promoting harmonized high quality standards, ethical principles, and innovation in the science and management of biorepositories (ISBER, 2014).

Most of biobanks in developed countries such as the UK, Japan, Canada, Sweden, etc. have been managed by the government. A common requirement for those biobanks is to have an association between u-health and big data. This shows they consider sharing biobank data, beyond collecting and storing.

Developed countries such as the United States and the United Kingdom have recognized that the use of public information is directly connected to nationwide service, and have competitively promoted various open policies. For the systematic utilization of public data, data based "mash-up" is needed. In order to achieve this, the distribution of data based on a structured model, 'structured data' is very important.

Big data and government data

Big data' is a new term that highlights the fact that this increase in the amount of data is not a problem, but a new challenge and opportunity. If we can handle the complexity of data, the competitive power of the organization will be increased (Lee, 2011). Extracted meaningful and valuable information from massive data can be used for active management and change prediction. This big data is increasingly being introduced in the enterprise, as well as public sector. The next computing paradigm has been formed based on this data based environment. Not only big data from global companies, but also worldwide public information are in the high interest. USA PCAST (President's Council of Advisors on Science and Technology) reported 'Designing a Digital Future', that emphasized the need for big data strategy to all U.S. federal government agencies (PCAST, 2010).

Developed countries such as the United States and United Kingdom have been recognized that the use of public information is directly connected to nationwide service, and have been competitively promoted various open policies. For systematic utilization of public data, data based 'Mash-up' is needed. In order to achieve this, distribution of data based on a structured model, 'Structured data' is very important (Data.gov, 2014; Data.go.kr, 2014).

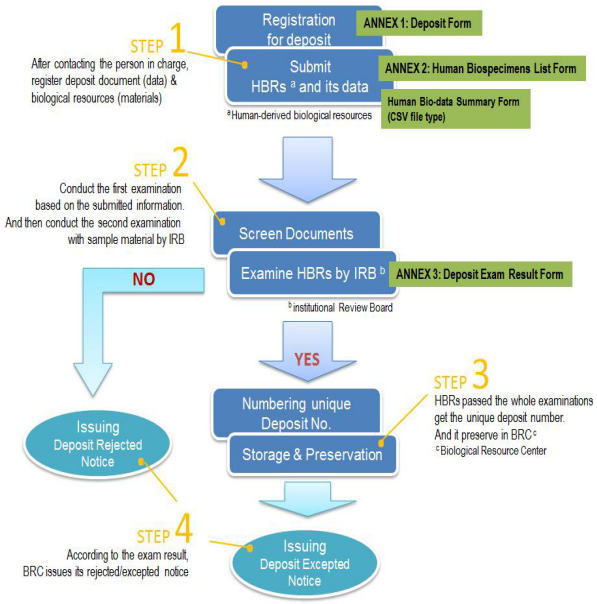

Figure 1. Human Biological Materials Deposit Process

Results

Deposit procedure

For the deposit of biological research resources, it is necessary to have a standard process and forms based on certain regulations. When a project is finished, the project director should be able to donate his or her project outcome. For the registration of a deposit, first of all, he or she fills out a deposit form which should be written about the deposit type and R&D information such as project-project name, serial number, and project performing period, and information pertaining to the project director and participating researchers. Then, the biological research resources and their data should be submitted with a human biospecimens list form and a human bio-data summary form. Secondly, based on the submitted documents and resources, there would be two examinations as to whether it is proper to be deposited. According to whether or not this is granted, the result form is transmitting to a donor. If the result is success, each biological research resource is then granted a unique deposit number and preserved in the standard format.

Modified deposit form

There are many reports outlining the complexity of biobanking that provide strong recommendations and the identification of best practice for all aspects of the process (ISBER, 2012). Figure 2 outlines the steps involved in biobanks. As the requirements for complex multi-institutional and international collections to study disease processes have been established, this article focuses on some of the important practical and ethical issues related to the integration of biobanks.

The steps involved in biobanks are a continuum (Watson, 2010), starting with participant consent, leading to the collection, processing, storage and distribution of samples. The process finishes with the publishing of data

and informing participants of how their samples have been used, which then supports the continuation of recruitment to biobanks.

We now explain the 26 items in the checklist (Table 1), and give published examples for each item. Some examples have been edited by removing citations or spelling out abbreviations. We advise researchers to address all items somewhere in their paper or project, but we do not prescribe a precise location or order. For instance, we discuss the reporting of results under a number of separate items, while recognizing that researchers might address several items within a single section of text or in a table.

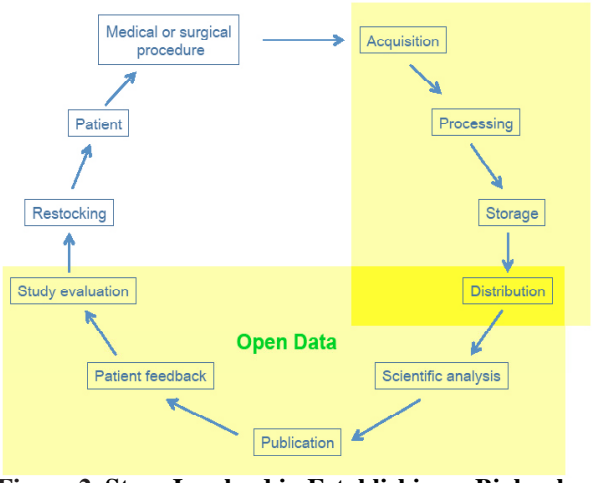

Figure 2. Steps Involved in Establishing a Biobank

XML standard based deposit form

XML data model is one of the reasonable and well represented data model among various data models for electronic representation (University of Essex, 2009). Among entire deposit procedure, we distinguished some phases that donor handle data such as submission phase and result notice phase that including contents whether deposit has been permitted. And we extracted data items from selected phases to construct XML data model. XML based deposit procedure data model consists of three components including deposit basic information, biospecimens information in submission phase, and permission information in result notice phase. Deposit basic information part in submission phase have information of researcher, organization, project, deposit type, and deposit date. For information of researcher and organization allow to be plural as research can be a collaborative project and researcher represents all participant researchers.

Biospecimen information part in submission phase has 12 data items including name, volume, volume unit, biospecimens class, and etc. Particularly, biospecimens class is classified with 5 classes; 1'B' for blood or blood derivation, 2'C' for human cell line, 3'F' for body fluid or substance, 4'G' for genomic, and 5'T' for tissue according to NIH classification criteria (NCI, 2011).

Permission information in result notice phase includes permission result, detail reason for approval, notice date, and etc. Permission result can be selected among four results; 1'appro' for approval, 2'condi' for conditional approval, 3'disappro' for disapproval, and 4're' for

Table 1. Order Form of Human Biological Material - Checklist of Items that Should be Addressed in Reports of Deposit

Item	Item number	Recommendation
ORDER FORM of HUMAN-DERIVED MATERIAL	1	Indicate the ANNEX's type with a commonly used term in the title
RESEARCH INFORMATION		
R&D Project ID	2	Describe the unique number of Research and Development project
Government Department ID	3	Describe the unique number of government agency including department
Government Department Name	4	Describe the name of government agency
R&D Project Name	5	Describe the name of Research and Development project
Research Title	6	Describe the research title with a commonly used term
Primary Investigator(PI)	7	Describe the name of primary investigator
Institutions	8	Describe the name of institutions
Research Period	9	Describe the research period of time
Research Stage	10	Indicate the stage of research divided by foundation, practical application, development, and et cetera
Science&Technology Standard Classification (Code)	11	Describe the proper code by science & technology standard classification
APPLICANT INFORMATION		
Primary Investigator Name	12	Describe the name of primary investigator
Primary Investigator Position	13	Describe the position of primary investigator
Primary Investigator TEL.	14	Describe the Telephone number of primary investigator
Primary Investigator FAX	15	Describe the FAX number of primary investigator
Primary Investigator email	16	Describe the email address of primary investigator
Researchers Name	17	Describe the name of participate researchers
CONTENTS OF REGISTRATION		
Original Report	18	Indicate the number of original report
Paper	19	Indicate the number of paper
Patent	20	Indicate the number of patent
Biological Material Info.	21	Indicate the number of biological material Information
Technical report	22	Indicate the number of technical report
Software	23	Indicate the number of software
Research Hardware	24	Indicate the number of research hardware
CONTENTS OF DEPOSIT		
Biological Material	25	Indicate the number of biological material
Components	26	Indicate the number of components

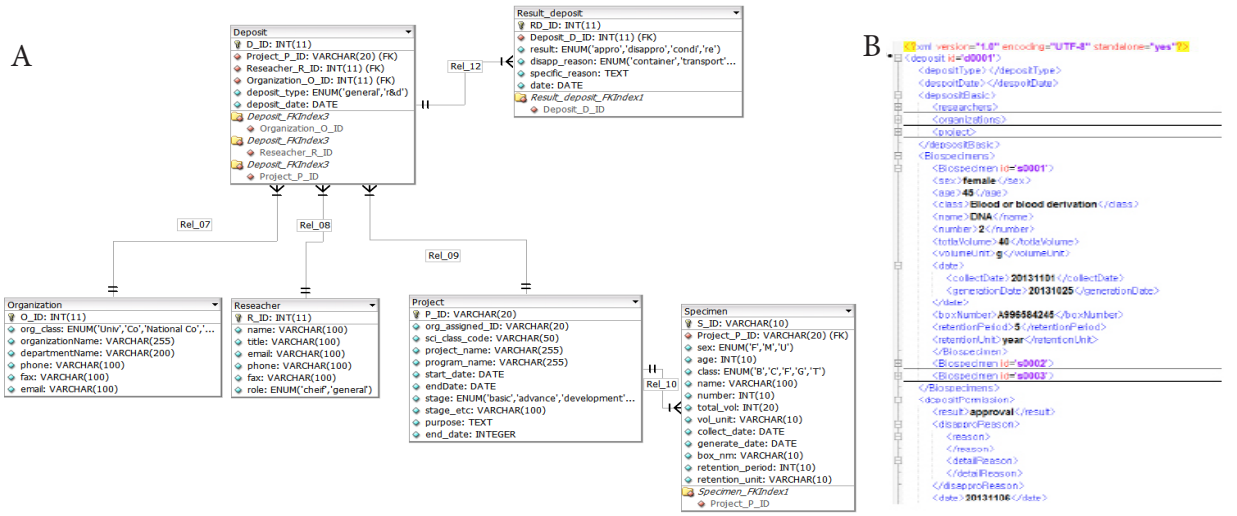

Figure 3. A) Database Table Schema of Deposit Procedure: 1) Tables with Red outline are Related with Deposit form, 2) Tables with Blue outline are Associated with Biospecimens information Collection form, and 3) Tables with Green outline are Deposit Permission Result form, B) Sample XML with Deposit Data

supplement and re-review. For the result with disapproval, disapproval reason can be selected among six reasons; 1'container' for damaging a container of specimen, 2'transport' for poor transmission condition, 3'discordance' for discordance between deposit specimen and recorded specimen, 4'short volume' for short amount of specimen volume, 5'abnormal' for abnormal of specimen quality control (QC) and 6'etc.' for other reasons.

To complete XML based data, we first stored the above designed and structured data into database tables. We extracted data from relative tables to form deposit information based on XML model (Figure 3).

Discussion

Securing and utilizing biomedical research resources ensures the future growth of the life science and healthcare industries. The South Korean government is interested in maximizing the return on public investment in biomedical research, especially including human biological materials in cancer patients. The lack of proper infrastructures and guidelines to capture and manage research data for re-use and re-analysis has become an issue. This paper presents guidelines for managing e.g. capturing, storing, and distributing human biological material resources, the outcome of the national R&D project in deposit and registration.

Under the South Korean regulations "Governmental Regulations on Capturing and Managing of Biological Research Resources" established recently with South Korean government-led efforts and some original regulations and rules in the KCDC (Korea Center for Disease Control & Prevention). Each resource center and research director of a national R&D project can utilize the guidelines to manage human biological material resources effectively and safely in a well-organized structure with consecutive order.

Structured guidelines for the deposit and registration of human biological materials in cancer patients were developed which are specifically national R&D project outcomes. These aim at helping donors, research directors, to comprehend deposit policy easily. They include comprehensive explanations about the policy of depositing of human biological materials including the R&D outcome, the object of deposit, the deposit standard, the deposit process, and the relevant regulations.

Despite this policy, depositing is in further need of revitalization. It is important to attract voluntary deposits from research directors by changing their perception of sharing their research outcomes. It is important that effective and safe preservation of the high value-added HBRs be developed by national R&D projects with systematic management in the long-term, so that useful and applicable resources can be distributed to many researchers. This will, in turn, induce them to develop other high quality resources which are the core of biological industry.

Also, this paper attempts to solve the issue of paper-based distribution through constructing a structured deposit form and XML based deposit procedure forms. The biggest problem of a paper based system is high maintenance costs, redundant writing, data error, and so on. With this XML based data model based on the revised and structured deposit procedure forms, researchers can share their data freely.

Those data that have been collected and will be deployed in the future will be big data through communication among researchers, and will be able to perform as public data. With high growth these public data, the government wishes to distribute and utilize those data. According to the government plan to switch paradigms from data collection to data utilization, the introduction of the concept of Health 2.0 is proposed here.

Health 2.0 is a new market based on government data. The discussion of health 2.0 has been held in health camp. The keywords in health 2.0 are consumer centered health care and web 2.0 communities. The basic idea of health 2.0 originated from combining some philosophies and technologies from web 2.0 with healthcare products (Hughes, 2008). Health 2.0 is participatory healthcare.

Enabled by information, software, and community that we collect or create, we the patients can be effective partners in our own healthcare, and we the people can participate in reshaping the health system itself.

The main concept is that health 2.0 is participatory healthcare. User participating productions like Search, Tools and social network systems (SNS) are playing important roles in health 2.0. Unlikely search engine like 'Naver' or 'Google' that have been used in the past, advanced search engines like 'Pubmed' or 'Healthline' are more likely to be used. With vertical results, these engines provide in-depth possible causes of symptoms by simply typing a symptom into the engines.

In this research, a structured deposit form was developed to implement government data which could be collected during the management of resource allocations, registration, and distribution according to National R&D outcomes. These guidelines result in following effects: existing contents about human biological materials easily accessed via 'Search', new types of information expressed and used via 'Tools', and new contents produced via 'SNS'. This will be the first step for the utilization of human biological materials.

Acknowledgement

This research was supported by a fund (2011-E63008-00)by Research of Korea Centers for Disease Control and Prevention, a grant of the Korean Health Technology R&D Project, Ministry of Health and Welfare (HI13C2164) and a grant of the National Research Foundation of Korea (NRF), Ministry of Science, ICT and Future Planing (20 10-0028631).

References

Data.go.kr. [cited at Jan. 6, 2014]. Available from: https://www.data.go.kr/

Data.gov. [cited at Jan. 6, 2014]. Available from: http://www.data.gov/

Denny JC, Ritchie MD, Basford MA, at el (2010). PheWAS: demonstrating the feasibility of a phenome-wide scan to discover gene-disease associations. *Bioinformatics*, **26**, 1205-10.

EuroBioBank. [cited at Jan. 6, 2014]. Available from: http://www.eurobiobank.org/

ISBER (2012). Best practices for repositories: collection, storage, retrieval and distribution of biological materials for research. BIOPRESERVATION AND BIOBANKING. US: International Society of Biological and Environmental Repositories.

KBN, Korea BioBank Network [Internet]. Osong, Korea; 2013 [cited at Jan. 6, 2014]. Available from: http://kbn.cdc.go.kr/

Yoo KY, Shin HR, Chang SH, et al (2005). Genomic epidemiology cohorts in Korea: present and the future. *Asian Pac J Cancer Prev*, **6**, 238-43.

Yoo KY, Shin HR, Chang SH, et al (2002). Korean multi-center cancer cohort study including a biological materials bank (KMCC-I). *Asian Pac J Cancer Prev*, **3**, 85-92.

Lee JE, Kim JH, Hong EJ, et al (2012). National biobank of Korea: quality control programs of collected-human biospecimens. *Osong Public Health Res Perspect*, **3**, 185-9.

Lee MJ (2011). Big data and the utilization of public data. *Int Inf Secur*, **2**, 47-64.

Hyun MK, Hwang JS, Kim JH, et al (2013). Survival outcomes after whole brain radiation therapy and/or stereotactic radiosurgery for cancer patients with metastatic brain tumors in Korea: a systematic review. *Asian Pac J Cancer Prev*, **14**, 7401-7

Ministry of Science, ICT and Future Planning [partial amendment, Mar. 9, 2011] (2011). Act on development, management and utilization of biological resources.

Ministry of Science, ICT and Future Planning [partial amendment, Mar. 9, 2011] (2011). Regulation on management of national research and development project, etc.

NCI (2011). NCI Best practices for biospecimen resources [internet]. u.s.: national cancer institute. [cited at Jan. 6, 2014]. Available from: http://biospecimens.cancer.gov/bestpractices/2011-NCIBestPractices.pdf

NIH (2013). Guidelines for human biospecimen storage and tracking within the nih intramural research program [internet]. u.s.: national institutes of health. [cited at Jan. 6, 2014]. Available from: http://sourcebook.od.nih.gov/oversight/BiospecimenGuidelines.pdf

Ivan P, Kern T, Damijan M (2009). Comparison of paper-based and electronic data collection process in clinical trials: costs simulation study. *Contemporary Clinical Trials*, **30**, 300-16.

PCAST (2010). Designing a digial future: federally funded research and development in networking and information technology. US: President's Council of Advisors on Science and Technology.

University of Essex (2009). Managing and sharing data, a best practice guide for researchers [Internet]. U.K. [cited at Jan. 6, 2014]. Available from: http://www.admin.ox.ac.uk/media/global/wwwadminoxacuk/localsites/researchdatamanagement/documents/managingsharing.pdf

Watson RW, Kay EW, Smith D (2010). Integrating biobanks: addressing the practical and ethical issues to deliver a valuable tool for cancer research. *Nat Rev Cancer*, **10**, 646-51.

Wolf SM, Brittney NC, Brian VN, et al (2012). Managing incidental findings and research results in genomic research involving biobanks and archived data sets. *Genet Med*, **14**, 361-84.

Ministry of Health and Welfare [partial amendment, Aug. 4, 2011] (2011). Health and medical service technology promotion act.

Swedish Biobank (2014). [cited at Jan. 6, 2014]. Available from: http://www.biobanks.se/

UK BioBank (2014). [cited at Jan. 6, 2014]. Available from: http://www.ukbiobank.ac.uk/

Hughes B, Joshi I, Wareham J (2008). Health 2.0 and medicine 2.0: tensions and controversies in the Field. *J Med Int Res*, **6**, 23.