

Minnesanteckningar 2011 04 27

Jan-Eric Litton reports with input from NCI's Biorepository Informatics Systems on the international consensus concerning the necessity of a highly controlled vocabulary with a single controlled term for each given concept. This common vocabulary is a prerequisite for international collaboration and for e.g. Meta-analyses.

In this context a major issue arises, being a discrepancy between the Swedish national concepts of Sample collection (provsamling) – called Study in international contexts. The fact that Sweden must reach a consensus as to whether to adopt the international vocabulary or to adopt a standard translation list to be used in all international collaborative efforts was agreed upon (but no such agreement was reached during this meeting).

A thought-provoking slide shows that an ontology grasps a particular entity in a specific geographical area from a defined perspective at a particular time.

Examples include annotating databases such as SNOMED CT, MeSH, ICD-N, etc.

An important goal is to make sure that as understanding changes, original meaning is not forgotten -
To provide a bridge between what we record and how we reason

The infrastructure of information is also important – a goal of SPREC.

Thomas Bergman, a molecular biologist working within LifeGene, presented some basic aspects of the Life-Gene protocol and the application of SPREC within this context. More than 10 thousand participants have now been recruited, contributing biometrics, questionnaires and samples.

Biometrics include audiometri, height, bio-impedance, weight, hip, waist, PEF, BP, HR, Blood and urine samples

SPREC, a Standard Preanalytical Code has been developed by a working group within ISBER to aid harmonisation of Biospecimen research¹. It provides standard nomenclature for sample collection, processing and storage procedures.

The code is comprised of 7 elements indicating

Type of sample

Type of primary container
Precentrifugation (delay between collection and processing with temperature range)
Centrifugation (RCF and time, braking y/n)
Second centrifugation
Post centrifugation delay until storage
Storage (container and temperature)

An example for a serum sample, obtained in an SST tube may have the following code:

SER-SST-A-E-N-A-G

For detailed explanations see the code reference ¹.

There are 2 versions for biological samples with a potential 11 letters for blood and 13 for tissues.

The SPREC elements for solid samples are more comprehensive, building on Laboratory Data Management System (LDMS) standards².

Type of sample
Type of collection
Warm ischemia time
Cold ischemia time
Fixation type
Fixation time
Storage conditions.

Within LifeGene all samples are assigned a preliminary SPREC value, which may be corrected after processing.

Advantages of SPREC codes include

Facilitation of exchange of samples and data

Facilitation of multicenter studies

Ease of implementation

Good support, with feedback for improvements – a new version will be published 2012.

(Joyce would like to add an initial preanalytical category with the letter F for fasting patient, N for non-fasting and X for unknown status.)

SPREC is being used in numerous studies internationally, by NIH and others.

Disadvantages include

Insufficient alternatives for tissue samples

Does not eliminate the need for own coding systems

No space for qualitative comments (hemolysis, lipidemia – sample damage)

Anna Beskow asks if SPREC code is saved as a string of data. If saved as individual components it would be possible to use components as search terms.

Martin Fransson presented the current Bioinformatic status of BBMRI.EU

Febr 2008 – jan 2011 (ca 5 M€)

- 50 partners
- 7 WP
- WP5 – biobankinformatics
- Officially 11 partners in 8 countries

Format for collection and harmonisation of data

- 1) bbmri.eu biobank questionnaires
- 2) WP5 metadatamodel
- 3) WP5 minimal dataset

5 different questionnaires

1. major questionnaire of 13 pages and ca 300 answers
2. financial support 1 page
3. ELSI (4 pages)
4. IT (10 pages)
5. Sample collection (Study?) (4pages)

A presentation of results is available at the www.bbmriportal.eu

Martin suggests that perhaps the questions are too many or too specific.

Few Swedish Biobanks have been reported/included, perhaps because a single Swedish “biobank” often includes numerous “studies” or sample collections.

The questionnaires have developed within the entire bbmri.eu – not within a single WP, and have raised considerable interest.

Results will soon be published as an article in Nature.

Bbmri.eu structure is described as a meta-model consisting of a hub and spokes – both internationally and within each country.

It uses digital descriptions at 2 levels using values and or existence attributes (ex. BMI = 24 or BMI exists yes/no)

Martin speculates that this model may be too complicated.

OECD 2000 has collected information concerning biobank museum collections

GBIF Global biodiversity information facility – has created the Darwin Core with few obligatory attributes within natural historical groups.

This was the inspiration behind a minimal dataset

Data is categorised according to Biobank, Study, participant/case/sample

Questionnaires have been derived from p3g –

Sample type definition has been taken from SPREC so much harmonisation has already occurred.

Loreana discussed standard formats and variables

Standard categories and questions for epidemiologic research are now available on the p3g home page³.

Standard formats should be used for new questionnaires and for reconstructed/converted questionnaires, using max. 30 alphanumeric (English) symbols, avoiding symbols that are international data commands. Do not mix small and capital letters.

Questions are presented according to categories in the hierarchy

Theme (tema)

Domain (domän)

Section (sektion)

Individual questions (enskilda frågor)

See details in p3g, DataShaper:

Example Theme: socioeconomy; Domain Life style, section diet, individual question alcohol

The p3g secretariat also provides tools for harmonisation in conjunction with collaborations.

See P3G secretariat, observatory, DataShaPER

[HTTP://WWW.DATASHAPER.ORG/DATASHAPER.HTML;JSESSIONID=941EC57644F48BA9976395DE949DC7EF#HARMONIZATIONTAB](http://www.datashaper.org/datashaper.html;jsessionid=941ec57644f48ba9976395de949dc7ef#harmonizationtab)

HARMONIZATION PLATFORM

The process follows a rigorous approach including:

1. The development of rules providing a formal assessment of the potential for each individual study to generate each of the variables in a given DataSchema.
2. The application of these rules to determine and tabulate the ability of each study to generate each variable, thereby identifying the information that can be shared.

	Study A	Study B	Study C	Study D
Variable 1	IMPOSSIBLE	PARTIAL	COMPLETE	COMPLETE
Variable 2	COMPLETE	COMPLETE	COMPLETE	COMPLETE
Variable 3	PARTIAL	IMPOSSIBLE	IMPOSSIBLE	IMPOSSIBLE

3. The development and application of a processing algorithm enabling that study to generate the required variable in an appropriate form.
4. Access to the Harmonization Platform is limited to collaborative context. Please contact us to see how we can work together.

Sonja Eaker Fält (Uppsala landsting) discusses SBR now and in the future

The Swedish Landstingen (counties) each own their own part of SBR

The system owner representative is Göran Elinder

The system is maintained by Inera AB

Software is delivered by Softwarepoint

Information structure

General structure (Organisation, Huvudman, biobank, ansvarig)

Information at the Individual level (this information is not available today)

Personal id (personnr)

Consent (date, type)

Sample (code, anatomical localisation, type of sample taken, etc. (see SPREC above)

SBR-IS SBR- Information System

ILis-Register, SBR Register and TBF Register (Tillägg för BiobanksForskning)

Implementation

Pathology/Cytology - 60-70% of all clinics are registered

Microbiology 30-40% are registered

Clinical Chemistry, Clin. Genetics, etc. – only a few are incorporated

Practical aspects

The Landstingens LIS must have access to and use software such as Sympathy and SafirLis. A major risk is that data will not be entered into the system.

According to the new suggested Biobank Law to take effect 2012 – all Biobanks will be required to submit their information.

The new law will apply to numerous additional areas with regulation of traceability and detailed information that must be available.

According to PUL (Personuppgiftslagen) a person can refuse registration within SBR.

A small working group with the purpose of establishing conditions necessary for collaboration between healthcare and research within a single or multiple landsting (counties), universities and/or the pharmaceutical industry will meet for the first time on May 11 with goals to create a uniform information infrastructure for sampling, registration, sample handling, storage, selection, retrieval and distribution by secure and integrity protective methods for future healthcare and research.

Pär-Ola Bendahl (Lund) and Loes Linsen (Belgium) Sapphire Biobank LIMS

Pär-Ola, a biostatistician working with Oncology Research and Loes Linsen, working with hematology have been involved in configuration and implementation of the Sapphire (LabVantage) Biobank LIMS version 5.0 purchased from Softwarepoint for use by members of the EU 6th framework Cancer Control and Prevention using Registers and Biobanks (CCPRB), Pi Joakim Dillner. The current version has been developed considerably since the initial purchase by KIs biobank. Developments have accelerated and communication improved after the recruitment of a consultant/programmer by Softwarepoint who is familiar with the world of biological samples (Chris Wilcoxon).

Pär Ola and Loes have worked to limit the visible menu options according to assigned roles and with validation of information import. For Pär-Ola a major challenge is the import of sample storage information from a highly heterogeneous collection. Both Loes and Pär-Ola have been working with the standard codes of information (t.ex. SNOMED) for classification of data. For details see their presentations.

Joyce Carlson reported on the current situation for Region Skånes Biobank (RSB).

Region Skåne made a political commitment September 6, 2010 to invest in and support a regional biobank (RSB).

Sample management will be handled by Labmedicine, which is now a separate organisation from that of the regional hospitals.

Three working groups are currently focussing on administration, sample receipt and handling and a biobank LIMS. A new director of the Biobank (Verksamhetschef) has been recruited and will start on August 1.

The Sample handling working group has recommended decentralisation with Sampling and initial sample handling including freezing at -80oC within 2 hours at four hospital labs, followed by intermittent transport of sample batches in frozen condition to a central repository.

The central repository will offer DNA/RNA extractions, robotic storage and retrieval and many other services. Tissue samples will mainly remain within the decentralised pathology departments with centralisation of specific sample types or collections. Thus a challenge for the LIMS group is to incorporate activities at all of these sites.

The LIMS working group includes members from Clinical Pathology, Microbiology, Chemistry, from Labmedicine's IT/MT and from the regional participant in NBR, as well as a previous AstraZeneca employee now representing the planned Big-3 biobank study. The initial purpose of this group was to evaluate the possibility to use existing contracts/licenses of LabVantage Sapphire. The Purchasing department has decided that existing contracts cannot be used. IT/MT representatives, currently involved in an active purchase of a new LIMS for all of laboratory medicine are eager to ensure compatibility of the Biobank LIMS with the LIMS to be used within the rest of laboratory medicine. Region Skåne has directed the center for clinical pharmacological research within Skånes University Hospital to develop support for extraction of information from electronic health records and population based health registers. The LIMS group limits its focus to sample registration and handling within the Labmedicine facilities. Results of specific analytical tests will not be stored in the LIMS, but possibly "meta information, e.g. the type of tests that have been performed may be stored. Awareness of SPREC codes and other conventions (e.g. from p3g) at this timepoint before the infrastructure is finalized is very helpful.

Loreana

SBR och bbmri.5 WP 5

Collaboration for the development of SBR-IS is essential to adapt the system so that it will be useful for both healthcare and research.

SKL and VR have together delegated the task to ensure a common minimal core dataset for both healthcare and the research groups that choose to participate.

Interviews with a number of researchers have identified a few basic user cases:

1. Search for samples for which consent has been changed
2. Search for Biobanks following certain criteria
3. Search for individual sample donors according to specific criteria
4. Search for biobank samples from specific sample donors
5. Create/maintain a No Consent Register

Pilot studies are ongoing

Pilot försök pågår

General Discussion

Jan-Erik Litton defines data categories:

Metadata	not sensitive or riskfylld
Aggregerade data	usually not sensitive
Objekt data	frequently sensitive with integrity risks

Disclosure filters can be used to protect sensitive data

The next meeting is planned for Uppsala in November. A general format of initial international information followed by national and then local from the hosting site is agreed upon.

Roxanna är sammankallande för en liten arbetsgrupp som snart skall börja definiera ett gemensamt datasätt som kan samlas med ett nationellt templat.

There is a brief discussion on the need for a minimal data set for cancer, Diskussion om minsta dataset för cancer.

MF suggests that sample data (analytical data) can be an information domain to allow harmonisation.

JEL mentions that in international discussions, some instances assume the use of commercial software solutions while others develop their own systems and completely forbid commercial solutions.

JEL suggests that an agreement and recommendation from within the biobank sphere, i.e. make decisions and show the way forward toward development of an excellent biobank LIMS, could likely influence adoption by NCI and other international actors.

SKL uttrycker Landstingens intresse för samarbete för att skapa en biobanks katalog på metanivå
Samordning av hur många prover som finns och inventering börja med enstaka områden – jättestort jobb.
Sonja kallar för att börja inventera biobanker (möte den 11/5)..

Open source programvara diskuteras.

Det anses viktigt att vi skapar gemensamma kriterier, kategorier, och format. Sedan kan vi sannolikt importera lagrade data till ett nytt system om företaget kraschar, etc.

ERIC processen diskuteras – om det blir en legal entitet kommer Huvudkontoret att placeras i Graz.

Det uttrycktes önskemål att leverantörer av kommersiella lösningar skall helst inte få vara med på våra möten för att undvika känsliga situationer.

Ovanstående Engelska -> Svenska kommentarer har registrerats under samtalets gång efter bästa förmåga

References

1. Fotini Betsou, Sylvain Lehmann, Garry Ashton, et al. Standard Preanalytical coding for biospecimens: Defining the Sample PREanalytical Code. *Cancer Epidemiol Biomarkers Prev* 2010;19:1004-1011.
2. <https://www.fstrf.org/apps/cfmx/apps/ldms/manual/manual.html>
https://www.fstrf.org/apps/cfmx/apps/manual/5.8/AppendixA_codes.pdf.