New concept proposal

Sample Processing

Author	Eelke van der Horst, Femke Kopmels	**Date last updated**	31/10/2023
Project	General interest	**Contact**	DCC
Dataset release	2024.1	**Consulted expert**	-

1 Rationale

Sample processing is an essential part of the (omics) experimental workflow. It comprises all processes that manipulate a sample before it can be analysed, such as dissociating tumor cells or culturing. Some sample processing steps are characteristic for a particular omics type, such as library preparation, others are more general, such as culturing.

2 Comparison to other standards/data models

2.1 SNOMED CT

SNOMED CT has some sample preparation concepts under its 'Procedure (procedure)' concept (SNOMED:71388002).

2.2 OBI and EFO

OBI has 'material processing' (OBI:0000094), defined as "A planned process which results in physical changes in a specified input material". However, this includes processing of any input material entity, not just samples but also complete animals or synthesis, although most subclasses of OBI 'material processing' would qualify as Sample Processing, EFO does not include this class, and also does not have a class with similar meaning to the proposed concept. However, many of its 'experimental process' (EFO:0002694) subclasses are sample processing steps, for instance 'sample splitting' (EFO:0030016).

3 Concept information

Concept or concept compositions or inherited	General concept name	General description	Contextualized concept name	Contextualized description	Type	Standard	Value set or subset	Meaning binding	Cardinality for composedOf
concept	Sample Processing	a process that prepares a sample for a subsequent process	Sample processing	a process that prepares a sample for a subsequent process					
composedOf	code	coded information specifying the concept	code	code information specifying the type of sample processing	Code				0:1
composedOf	input	input associated to the concept	input	the input sample	Sample				0:n
composedOf	output	output associated to the concept	output	the output sample	Sample				0:n
composedOf	start datetime	datetime at which the concept started	start datetime	datetime at which the sample processing started	temporal				0:1
composedOf	quality control metric	quality control metric associated to the concept	quality control metric	quality control metric related to the output of the sample processing	Quality Control Metric				0:n
composedOf	predecessor	a preceding process associated to the concept	predecessor	process preceding this sample processing	Sample Processing			RO:0002087 |immediately preceded by|	0:n
composedOf	standard operating procedure	standard operating procedure associated to the concept	standard operating procedure	standard operating procedure that was followed for this sample processing	Standard Operating Procedure				0:1

General concept name	Cardinality for concept to Administrative Case	Cardinality for concept to Data Provider	Cardinality for concept to Subject Pseudo Identifier	Cardinality for concept to Source System
Sample Processing	0:n	1:1	0:n	1:n

4 Impact on the SPHN Dataset

Optional (if existing concepts need to be adapted because of this new concept, state here the currently released version of the existing concept and the proposed adapted version)

5 Discussion

As with all experimental processes, Sample processing steps may be chained in sequence, and may consist of individual steps that provide additional metadata. For instance, it can be part of an assay concept to provide essential metadata on the sample processing that is required for the particular assay type.

To indicate the type of sample processing, descendant terms of EFO's 'experimental process' (EFO:0002694), OBI's 'planned process' (OBI:0000011), or SNOMED CT 'Procedure (procedure)' (SNOMED:71388002) can be used.

Since *Sample Processing* has input and output of type *Sample*, having a composedOf named 'sample' is ambiguous. Hence the names 'input' and 'output'. Since input and output may not always be relevant or known, for instance in case of intermediate samples between two processing steps, the minimum cardinality of these properties is 0. Note that a sample processing step may have multiple input samples, for instance in the case of a tumor sample and antibody sample, or when multiple input samples are pooled into the same library. Consequently, there could also be multiple output samples for a sample processing step. For some scenarios, it is practical to have multiple samples as output of a sample processing step (For example, in the case of creating multiple isolates from a given sample). While in other scenarios, particularly in the context of genomics, it is much more prudent to have a sample processing step for each corresponding output sample.

Processing datetime is an optional attribute to the *Sample Processing* concept. As collection datetime is mandatory for Samples, the collection datetime for the output sample is equal to the processing datetime of *Sample Processing*.

6 Example

Isolating DNA from blood sample

code: 104166004 |Nucleic acid molecular isolation or extraction method (procedure)|
start datetime: 2023-06-26
input:
    identifier: BloodSample1
    collection datetime: 2023-06-26
    material type code: 119297000 |Blood specimen|
output:
    identifier: DNASample1
    collection datetime: 2023-06-26
    material type code: 258566005 |Deoxyribonucleic acid specimen (specimen)|

standard operating procedure:
    name: DNA isolation from blood sample
    description: To describe how DNA from blood must be isolated and stored.
    version: V.1.1
    data file:
        format code: EDAM:format_3508 |PDF|
        code: EXAMPLE01

Imaging Mass Cytometry (IMC) staining sample preparation for IMC.

code: OBI:0302887 |staining|
start datetime: 2023-06-26
input:
    identifier: FFPE_block_123
    collection datetime: 2023-06-24
    material type code: OBI:1200000 |FFPE specimen|
output:
    identifier: FFPE_block_123_stained
    collection datetime: 2023-06-26
    material type code: OBI:1200000 |FFPE specimen|
standard operating procedure:
    name: Tissue Staining for Imaging Mass Cytometry
    description: This protocol describes the preparation and staining of human FFPE material for multiplex visualization using Imaging Mass Cytometry (IMC). This protocol uses basic standard immunohistochemical staining techniques. The tissue is incubated with antibodies that have specific affinity for different cells and tissues in the context of a formalin fixed thin section of human tumor material. Antibodies are labeled with heavy metals instead of the typical fluorochromes, and are visualized using the Hyperion Imaging System manufactured by the Fluidigm corporation.
    version: FEB 25, 2021
    data file:
        uniform resource identifier: dx.doi.org/10.17504/protocols.io.bspyndpw
        format code: EDAM:format_3508 |PDF|
        code: EXAMPLE01