New concept proposal
Tobacco Exposure

Author	Sylvain Pradervand, Amanda Herbrand	Date last updated	14.11.2023
Project	SPO	Contact	DCC
Dataset release	2024.1	Consulted expert	-

1   Rationale

Long term exposure to tobacco has proven to be a major factor for cancerous and non-cancerous diseases, especially those related to the respiratory system. The concept "Tobacco Exposure" carries information of importance for both diagnosis and prognosis, as well as helping to administer the best possible treatment.

2   Comparison to other standards/data models

2.1   SNOMED CT

In SNOMED CT, there is the observable entity 229819007 |Tobacco use and exposure (observable entity)| with the corresponding value set from the findings hierarchy, 365980008 |Finding of tobacco use and exposure (finding)|. The descendants within the findings hierarchy include different kinds of information, such as the relation of the subject to a possible attempt to stop, to the modality of smoking (cigarette, cigar, pipe), or intensity like heavy or low.
The widely used quantity "pack years" is also covered (401201003 |Cigarette pack-years (observable entity)|).

2.2   NCI Thesaurus

In the NCI Thesaurus, there is a concept "Tobacco Use" (C0543414) with the children Electronic cigarette user, ex-tobacco user, passive smoke exposure, tobacco smoking and smokeless tobacco use. There are only two descendants related to the smoking modality: cigarette and cigar.
https://ncimeta.nci.nih.gov/ncimbrowser/pages/concept_details.jsf?dictionary=NCI%20Metathesaurus&code=C0543414&type=relationship

A project of

SAMW ASSM
Schweizerische Akademie der Medizinischen Wissenschaften
Académie Suisse des Sciences Médicales
Accademia Svizzera delle Scienze Mediche
Swiss Academy of Medical Sciences

SIB
Swiss Institute of Bioinformatics

SIB | Swiss Institute of Bioinformatics
PHI | Personalized Health Informatics Group
www.sphn.ch | dcc@sib.swiss

2.3 openEHR

In the openEHR architecture standard, there are two concepts for the use of tobacco containing products, "Tobacco smoking summary" and "Smokeless tobacco summary". "Vaping summary" is a third concept, with a very similar structure as the other two concepts. The concepts are very comprehensive, including both dates for start and cessation of smoking as well as overall years, amount and pack years.

https://ckm.openehr.org/ckm/archetypes/1013.1.2466/mindmap
https://ckm.openehr.org/ckm/archetypes/1013.1.2817/mindmap
https://ckm.openehr.org/ckm/archetypes/1013.1.4823/mindmap

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	Tobacco Exposure	episode of exposure of an individual to tobacco in various forms	Tobacco Exposure	episode of exposure of an individual to tobacco in various forms				SNOMED CT:229819 007 |Tobacco use and exposure (observable entity)|	
composedOf	type code	coded information specifying the type of the concept	exposure type code	code describing the exposure of the individual to tobacco and related substances	Code	SNOMED CT	702979003 |Never used tobacco (finding)|; 43381005 |Passive smoker (finding)|; 77176002 |Smoker (finding)|; 8517006 |Ex-smoker (finding)|; 228494002 |Snuff user (finding)|; 228493008 |Ex-snuff user (finding)|; 81703003 |Chews tobacco (finding)|; 228513009 |Ex-tobacco chewer (finding)|;		0:1

							722499006 |Electronic cigarette user (finding)|; 35361000087100 |Ex-electronic cigarette user (finding)|		
inherited	agent code	coded information specifying an active power or cause (as principle, substance, physical or biological factor, etc.) that produces a specific effect	consumed substance code	code for the tobacco derived substance consumed by the individual	Code	SNOMED CT	descendant of: 39953003 |Tobacco (substance)|		0:1
inherited	degree code	coded information specifying the degree of the concept	exposure degree code	code for the degree of exposure to the substance. For cigarettes: 1-9 cigs/day = low, 10-20 cigs/ day = moderate, 20+ cigs/day = high	Code	SNOMED CT	62482003 |Low (qualifier value)|; 1255665007 |Moderate (qualifier value)|; 75540009 |High (qualifier value)|		0:1
inherited	duration	elapsed time from start to end of the concept	exposure duration	duration of exposure to tobacco in years	Quantity	UCUM	unit -> code restricted to: a		0:1
inherited	quantity	an amount or a number of the concept	pack years	calculated number of years that the individual consumed one pack of cigarettes per day for	Quantity	UCUM	unit -> code restricted to: a		0:1
inherited	route code	coded information specifying the route of the concept	route code	coded information specifying the route of the exposure to the agent	Code	SNOMED CT			0:n

inherited	start datetime	datetime at which the concept started	exposure start	datetime at which the exposure started	temporal				0:1
inherited	end datetime	datetime at which the concept ended	cessation datetime	datetime the individual stopped being exposed to tobacco related products	temporal				0:1
inherited	record datetime	datetime the concept was recorded	record datetime	datetime the tobacco exposure was recorded	temporal				0:1

General concept	Cardinality for concept to Administrative Case	Cardinality for concept to Data Provider	Cardinality for concept to Pseudo Identifier	Cardinality for concept to Source System
Tobacco Exposure	0:1	1:1	1:1	1:1

SIB
Swiss Institute of Bioinformatics

4 Impact on the SPHN Dataset

No impact is expected on the other concepts of SPHN Dataset. Tobacco Exposure inherits from the new concept [Exposure](#) (release 2024.1).

5 Discussion

The concept of Tobacco Exposure can be instantiated to express different episodes of consuming or being exposed to different modalities or degrees of exposure.

5.1 Naming of the concept

The name "Tobacco exposure" was chosen in contrast to other suggestions such as "Smoking Behaviour" (making it independent of Tobacco) or the previous title "Tobacco Smoking Behaviour" (focusing more on the active part and also the modality of smoking). "Smoking Behaviour" would possibly allow the inclusion of other substances or media such as vaping. However, this concept is meant for assessing risks regarding certain diseases. This risk is best established for exposure to tobacco, associations to diseases for other media such as non-tobacco-containing, nicotine-based e-cigarettes is not clear as of today. Other consumptions of tobacco, such as chewing, are, however, clearly associated with a risk for cancerous diseases and should hence be included in this concept.

5.2 Data elements

type code (contextualized name: **exposure type code):**
Value Set for the composedOf "exposure type code" was chosen with distinctive values because these are the most general and also the most common options. The SNOMED-value set for descendants of 365980008 |Finding of tobacco use and exposure (finding)| is very broad and also mixes in substance consumed and duration. That's why we suggest staying with the most high-level values that also depict the information available in the hospitals best.
Nevertheless, use of e-cigarettes (vaping) is included in the value set to depict the whole spectrum of possible risky behavior without including it in the concept name.

quantity (contextualized name: **pack years):**
Pack years is one of the most widely used variables when it comes to assessment of exposure to tobacco (compare, for example, SNOMED CT: 401201003 |Cigarette pack-years (observable entity)|). Although it is calculated by using the amount and duration of cigarette consumption, it is often the only variable documented in the EHR. Because it only applies to cigarettes, the concept was extended by the compsedOfs "degree code" (exposure degree code) and "duration" (exposure duration) to also be able to capture the amount and duration.
Note: The number of "pack years" shall be represented by the property "quantity". This is obvious from the contextualized name of the composedOf ("pack years"), however, not necessarily from the general name "quantity" which is inherited from the parental concept '[Exposure](#)'.

end datetime (contextualized name: **cessation datetime):**

A project of

SAMW ASSM
Schweizerische Akademie der Medizinischen Wissenschaften
Académie Suisse des Sciences Médicales
Accademia Svizzera delle Scienze Mediche
Swiss Academy of Medical Sciences

SIB
Swiss Institute of Bioinformatics

SIB | Swiss Institute of Bioinformatics
PHI | Personalized Health Informatics Group
www.sphn.ch | dcc@sib.swiss

Currently it is not possible to only capture the year of cessation. DateTimes are of type xsd:dateTime which captures year, month, date and timestamp in a single string. To express "2018" one would have to use "2018-01-01T00:00:00+00:00" as a proxy.
Source: https://www.w3schools.com/xml/schema_dtypes_date.asp

5.3 Possible extensions

We discussed also including a composedOf for "reason stopped" because it might be relevant for some research questions. However, at the moment the data in EHR will most likely not yield this information and there is no appropriate value set within SNOMED CT as of today. If necessary, the concept could be extended by this composedOf if there's a suitable proposition for a value set.

6   Examples

6.1   Example 1: Ex-smoker (moderate smoking for 40 years, quit 2018)

Tobacco Exposure

    type code:
        identifier: 8517006
        name: Ex-smoker (finding)
        coding system and version: SNOMED-CT-2023-07-31
    agent code:
        identifier: 66562002
        name: Cigarette smoking tobacco (substance)
        coding system and version: SNOMED-CT-2023-07-31
    degree code:
        identifier: 1255665007
        name: Moderate (qualifier value)
        coding system and version: SNOMED-CT-2023-07-31
    duration:
        quantity:
            value: 40
            unit:
                code:
                    identifier: a
                    name: a
                    coding system and version: UCUM
                comparator: -
    quantity:
        value: 20
            unit:
                code:
                    identifier: a
                    name: a
                    coding system and version: UCUM
                comparator: -
    route code:

identifier 284009009
name Respiratory tract route (qualifier value)
coding system and version SNOMED-CT-2023-07-31
start datetime: -
end datetime: 2018-01-01 00:00
record datetime: 2023-08-04 10:22:15

6.2   Example 2: Current e-cigarette smoker (low consumption since 5 years)

Tobacco Exposure

type code:
      identifier: 722499006
      name: Electronic cigarette user (finding)
      coding system and version: SNOMED-CT-2023-07-31
agent code: -
degree code:
      identifier: 62482003
      name: Low (qualifier value)
      coding system and version: SNOMED-CT-2023-07-31
duration:
      quantity:
            value: 5
            unit:
                  code:
                        identifier: a
                        name: a
                        coding system and version: UCUM
                comparator: -
quantity: -
route code: -
start datetime: -
end datetime: -
record datetime: 2023-09-01 09:42:18