



SNOMED CT Patologian löydösluokituksen hyödyntämisestä tietojärjestelmissä

Esa-Matti Tolppanen

SNOMED CT Seminaari 20.10.2020

Esityksen käytetyt lyhenteet

- SNOMED II – Systematized Nomenclature of Medicine, versio 2
 - American College of Pathologists kehittämä luokitus, jonka alkuperäinen käyttötarkoitus oli patologian löydösten luokittelu palvelemaan näytearkiston indeksointia (SNOP – Systematized Nomenclature of Pathology)
- SNOMED CT – SCT - Systematized Nomenclature of Medicine Clinical Terminology
 - SNOMED:sta ja NHS:n Clinical Terminology version 3 (CTV3) termistöstä yhdistämällä luotu terminologinen järjestelmä (7/2020 versiossa voimassa olevia käsitteitä 477 950 kpl ja termejä 1 228 361 kpl)
- SNOMED CT PAT – SCT PAT - SNOMED CT Patologian löydösluokitus.
 - THL:n projektina kehittämä suomalainen osajoukko käytettäväksi patologian laboratoriolöydösten rakenteistamiseen.
- ICD-O-3.2 – International Classification of Diseases – Oncology versio 3.2
 - WHO:n julkaisema kasvaintautien luokitus.
- LIMS – Laboratory Information System
 - Patologian laboratoriossa käytössä olevan tietojärjestelmä
- SSR – Suomen Syöpärekisteri
 - Tärkein patologian laboratorioden tuottaman löydöstiedon toisiokäyttäjä

SNOMED CT PAT lähtökohdista

- Tietojärjestelmien kannalta tavoitteet:
 - Laboratorioissa käytössä olevan SNOMED II pohjaisen luokituksen päivittäminen nykyiseen SNOMED CT luokitukseen ja luokituksen harmonisointi, jotta tietojen yhteiskäyttö olisi käytännössä mahdollista
 - Sisältää ominaisuuksia, joilla helpotetaan uuteen luokitukseen siirtymistä – myös vähentää tarvittavia muutoksia laboratorioiden tietojärjestelmiin
 - Luoda sellainen ylläpitomekanismi ja rahoitus, joka vähentää tietojärjestelmien ylläpitotyötä ja varmistaa järjestelmien diagnoositaustarekisterien pysymisen yhtenäisenä
- Patologian löydöstietojen ”ydin”:
 - *Näytetieto*: mistä näyte on otettu?, millainen näyte on kyseessä (näytetyyppi)?
 - *Patologin tulkinta* 1-n:sta kudus-/solunäytteestä → diagnoosi (tai patologisanatominen diagnoosi = ”PAD”)
 - Mitä ihmisen anatomian osaa tulkinta koskee (*topografia*)?
 - Millainen muutos / ei-muutosta näytteessä näkyi (*morfologia*)?
 - PAD voidaan myös ilmaista yhtenä tautikäsitteenä (*disease* - SNOMED CT:ssa ”disorder”)
 - Esimerkiksi: 681631000119108|Primary adenocarcinoma of body of stomach (disorder)
 - Josta SNOMED CT järjestelmästä on tieto, että se koostuu osista: *topografia*: 450594004|Gastric corpus structure ja *morfologia*: 681631000119108|Primary malignant adenomatous neoplasm

SNOMED CT PAT tietosisällöstä 1/3

- SNOMED CT PAT standardoi tässä vaiheessa pääasiassa topografia, morfologia ja tauti –käsitteitä ja niistä käytettäviä tunnisteita
 - Esimerkkinä käsite: 56665009|Infiltrating basal cell carcinoma (morphologic abnormality)
- Kustakin käsitteestä on kansallinen suositus ensisijaisesti käytettävästä termistä
 - Esimerkkinä olevassa käsitteessä: 198680017 | Infiltrating basal cell carcinoma, sclerosing
- Kullekin käsitteelle ylläpidetään 0-n synonyymitermiä
 - Yllä olevassa käsitteessä: 198681018=Infiltrating basal cell carcinoma, non-sclerosing|198682013=Basal cell carcinoma, morpheic|198684014=Infiltrating basal cell carcinoma
 - Jakelussa on myös käsitteen ”synonyymitermeinä” ruotsinkieliset termit, yllä olevasta esimerkistä: 51291000052117|Invasiv basalcellscancer (*SNOMED CT Swedish edition* jakelusta)
- Lisäksi paikallisesti laboratorio voi ylläpitää kustakin käsitteestä omia synonyymitermejä – kunhan tietojärjestelmän taustarekisterissä on tieto siitä, mihin SNOMED CT PAT käsitteeseen (SNOMED CT käsite-ID:hen) kyseinen synonyymitermi viittaa.

SNOMED CT PAT tietosisällöstä 2/3

- SNOMED CT PAT:ssa jokaisella termillä on nk. input-id:n, jonka tarkoitus on helpottaa siirtymistä SNOMED CT PAT käyttöön
- Input-id koostuu kahdesta osasta:
 - <perinneosa>”-”<SNOMED CT TERM ID>
 - Esimerkiksi termi *Infiltrating basal cell carcinoma, sclerosing* input-ID on muotoa: M80923-198680017
 - Perinneosa on muodoltaan sama kuin tämän hetkinen SNOMED II koodi (helpottaa siirtymistä).
 - Kasvaintautien morfologian osalta perinneosa myös sama kuin ICD-O-3.2 tunnistekoodi (helpottaa siirtymistä)
 - Mutta rakenne mahdollistaa huomattavan paljon tarkemman löydöksen kuvauksen koska se yhdessä SNOMED CT osan kanssa tunnistaa kaikki SNOMED CT käsitteet

SNOMED CT PAT tietosisällöstä 3/3

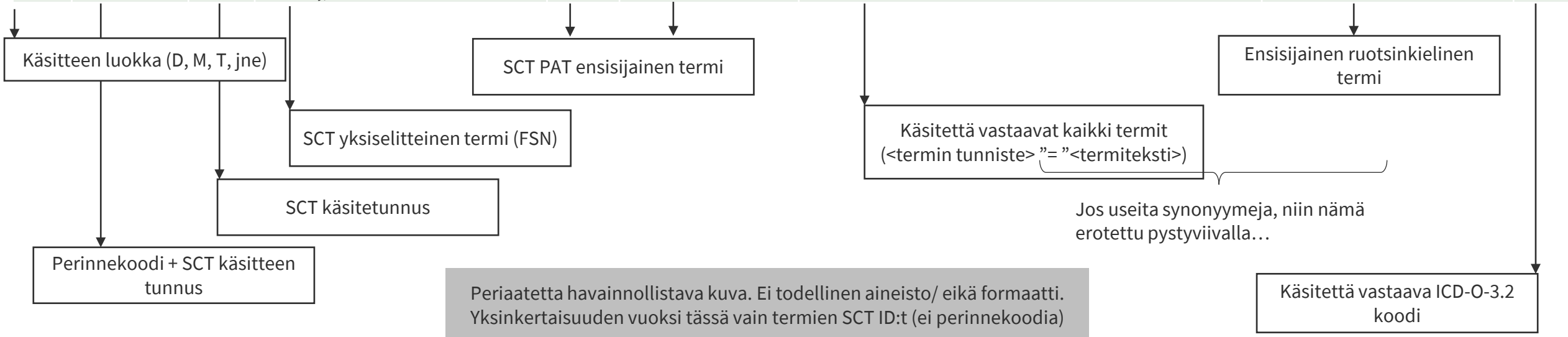
- Kukin termi viittaa yhteen käsitteeseen. Jokaisella käsitteellä on yksiselitteinen tunnus on (SCT concept ID ja ”perinnekäsite-ID”
 - ”Perinnekäsite-ID” on input-id:n tapaan muotoa <perinneosa>”-”<SNOMED CT CONCEPT ID>, jossa perinneosa on muodoltaan morfologian osalta SNOMED II / ICD-O-3 muotoinen
- Kustakin käsitteestä on lisäksi muuta tietoa:
 - SNOMED CT Ruotsin laitoksesta käsitettä vastaava ensisijainen ruotsinkielinen termi
 - ICD-O-3.2 vastaavuus (tunniste ja termi)
 - Erilaista ylläpitoa helpottavaa hallinnollista tietoa (lisäys, muutos, poisto pvm jne)
- Näin ollen SNOMED CT PAT luokituksen ydin on jotakuinkin seuraavan näköinen (kuva seuraavalla sivulla)

SNOMED CT Patologian löydösluokitus

- Käsitteiden mukaan jaoteltuna (kukin rivi on uniikki käsite)

TMD	Legacy_conceptID	Sct_conceptID	SCT yksiselitteinen nimi (FSN)	SCT PAT preferredID	SCT PAT_preferred_term	SCT PAT_synonyms	SCT Sweden preferred	ICDO-3
M	M80213-58248003	58248003	Carcinoma, anaplastic (morphologic abnormality)	96799014	Carcinoma, anaplastic	96799014=Carcinoma, anaplastic	Anaplastiskt karcinom	8021/3
M	M80223-16741004	16741004	Pleomorphic carcinoma (morphologic abnormality)	28346015	Pleomorphic carcinoma	28346015=Pleomorphic carcinoma	Pleomorft karcinom	8022/3
M	M80313-42596004	42596004	Giant cell carcinoma (morphologic abnormality)	71062012	Giant cell carcinoma	71062012=Giant cell carcinoma	Jättecelskarcinom	8031/3
M	M80400-128876004	128876004	Tumorlet, benign (morphologic abnormality)	207527012	Tumorlet, benign	207527012=Tumorlet, benign	Liten benign tumor	8040/0
M	M80401-72938002	72938002	Tumorlet (morphologic abnormality)	121155015	Tumorlet	121155015=Tumorlet		8040/1
M	M80413-74364000	74364000	Small cell carcinoma (morphologic abnormality)	123497012	Small cell carcinoma	123497012=Small cell carcinoma	Småcelligt karcinom	8041/3
M	M80416-709194001	709194001	Metastatic small cell carcinoma (morphologic abnormality)	3039041011	Metastatic small cell carcinoma	3039041011=Metastatic small cell carcinoma	Metastaterat småcelligt karcinom	8041/6
M	M80500-23730008	23730008	Papilloma, no International Classification of Diseases for Oncology subtype (except papilloma of bladder M-81201) (morphologic abnormality)	39844010	Papilloma	39844010=Papilloma	Papillom (förutom papillom i blåsa M-81201)	8050/0
M	M80503-25910003	25910003	Papillary carcinoma (morphologic abnormality)	43414013	Papillary carcinoma	43414013=Papillary carcinoma	Papillärt karcinom	8050/3
M	M80513-89906000	89906000	Verrucous carcinoma (morphologic abnormality)	149032010	Verrucous carcinoma	149032010=Verrucous carcinoma 149034011=Verrucous squamous cell carcinoma	Verruköst karcinom	8051/3
M	M80520-63451008	63451008	Squamous cell papilloma (morphologic abnormality)	105472015	Squamous cell papilloma	105472015=Squamous cell papilloma	Skivepitelpapillom	8052/0

Jokaisella termillä on myös input-id (puuttuu kuvasta)



SNOMED CT PAT työversio Koodistopalvelussa 1/3

- Luokituksen työversio 0.7 on ladattavissa tutustuttavaksi Koodistopalvelusta osoitteesta:
 - <https://koodistopalvelu.kanta.fi/codeserver/pages/classification-view-page.xhtml?classificationKey=3564&versionKey=4045>

CodeId	Abbreviat	ShortNam	LongName	BeginningDat	ExpiringDate	LastModifiedDate	LastModif	Status	Descriptio	OID	CreatedDate	A:Dimens	A:SNOME	A2:sct_sy	A:short_na
10057001	Chronic lic	Chronic lic	Chronic lic	01.01.2020	31.12.2099	06.07.2020 13:12:45	Lehtonen,	1			06.07.2020 13:12:45	Morpholo	10057001	17552016	17552016
10087007	Schistosor	Schistosor	Infection k	01.01.2020	31.12.2099	06.07.2020 13:12:45	Lehtonen,	1			06.07.2020 13:12:45	Disease	10087007	17604017	17604017
10138007	Bone marr	Bone marr	Bone marr	01.01.2020	31.12.2099	06.07.2020 13:12:45	Lehtonen,	1			06.07.2020 13:12:45	Morpholo	10138007	17681011	17681011
1015003	Peritonsill	Peritonsill	Structure c	01.01.2020	31.12.2099	06.07.2020 13:12:45	Lehtonen,	1			06.07.2020 13:12:45	Topograpl	1015003	2781013	2781013
10200004	Liver	Liver	Liver struc	01.01.2020	31.12.2099	06.07.2020 13:12:45	Lehtonen,	1			06.07.2020 13:12:45	Topograpl	10200004	17776014	17776014
10209003	Parotid lyr	Parotid lyr	Structure c	01.01.2020	31.12.2099	06.07.2020 13:12:45	Lehtonen,	1			06.07.2020 13:12:45	Topograpl	10209003	17788011	17788011
10229700	Main bror	Main bror	Main bror	01.01.2020	31.12.2099	06.07.2020 13:12:47	Lehtonen,	1			06.07.2020 13:12:47	Topograpl	10229700	16546101	16546101
10230000	Gingival m	Gingival m	Structure c	01.01.2020	31.12.2099	06.07.2020 13:12:47	Lehtonen,	1			06.07.2020 13:12:47	Topograpl	10230000	16546601	16546601
10288008	Squamous	Squamous	Squamous	01.01.2020	31.12.2099	06.07.2020 13:12:45	Lehtonen,	1			06.07.2020 13:12:45	Morpholo	10288008	17919018	17919018
10293006	Iliac artery	Iliac artery	Structure c	01.01.2020	31.12.2099	06.07.2020 13:12:45	Lehtonen,	1			06.07.2020 13:12:45	Topograpl	10293006	17927010	17927010
10295004	Chronic vi	Chronic vi	Chronic vi	01.01.2020	31.12.2099	06.07.2020 13:12:45	Lehtonen,	1			06.07.2020 13:12:45	Disease	10295004	17930015	17930015
10361400	Follicular	Follicular	Follicular	01.01.2020	31.12.2099	06.07.2020 13:12:47	Lehtonen,	1			06.07.2020 13:12:47	Morpholo	10361400	16731201	16731201
10362008	Accessory	Accessory	Accessory	01.01.2020	31.12.2099	06.07.2020 13:12:45	Lehtonen,	1			06.07.2020 13:12:45	Morpholo	10362008	18039019	18039019
10364300	Atypical e	Atypical e	Atypical e	01.01.2020	31.12.2099	06.07.2020 13:12:47	Lehtonen,	1			06.07.2020 13:12:47	Morpholo	10364300	20189201	20189201

SNOMED CT PAT työversio Koodistopalvelussa 2/3

- Koodistopalvelusta ladattavan tiedoston sarakkeiden lukuohje:
 - A | Codeid = yksiselitteinen tunniste (nyt = SCT käsitteen tunniste)
 - B | Abbreviation = SCT PAT FI käsitteestä käytettävä ensisijainen termi
 - C | Shortname = SCT PAT FI ensisijainen termi (nyt sama kuin B)
 - D | LongName = SCT FSN käsitteen yksiselitteinen nimi
 - E-L | erilaista hallintaan liittyvää tietoa
 - M | A:Dimension = Topography/Morphology/Disease
 - N | A:SNOMEDCT = SCT käsitteen tunniste
 - O | A2:sct_synonyms = Synonyymitermien tunnukset ja termien tekstit
 - P | A:short_name_id = SCT PAT FI ensisijaisen termin (B, C) tunniste

SNOMED CT PAT työversio Koodistopalvelussa 3/3

- Teknisesti SNOMED CT PAT määrittäminen on relaatiotietokanta, josta tehdään tarvittavia aineistoja eri työvaiheita ja julkaisuja varten
- Koodistopalvelussa oleva työversio *THL/SNOMED CT - Patologian löydösluokitus* on käsiteluettelo. Se ei sisällä ruotsinkielisiä ensisijaisia termejä, ei ICD-O-3.2 vastaavuuksia eikä siirtymistä helpottavaa input-id koodia/perinneID:tä.
- Lopullinen jakelumuoto (termit vs. käsitteet, tietosisältö) päätetään myöhemmin saadun palautteen perusteella
- Versio tulee päivittymään nyt kun projekti on taas käynnistynyt puolen vuoden tauon jälkeen. Suunnitelmissa on, että seuraava työversio sisältää ICD-O-3.2 yhteensopivuudet ja se on tarkistettu ainakin kahden yo-sairaalan patologian laboratoriossa (12/2020)

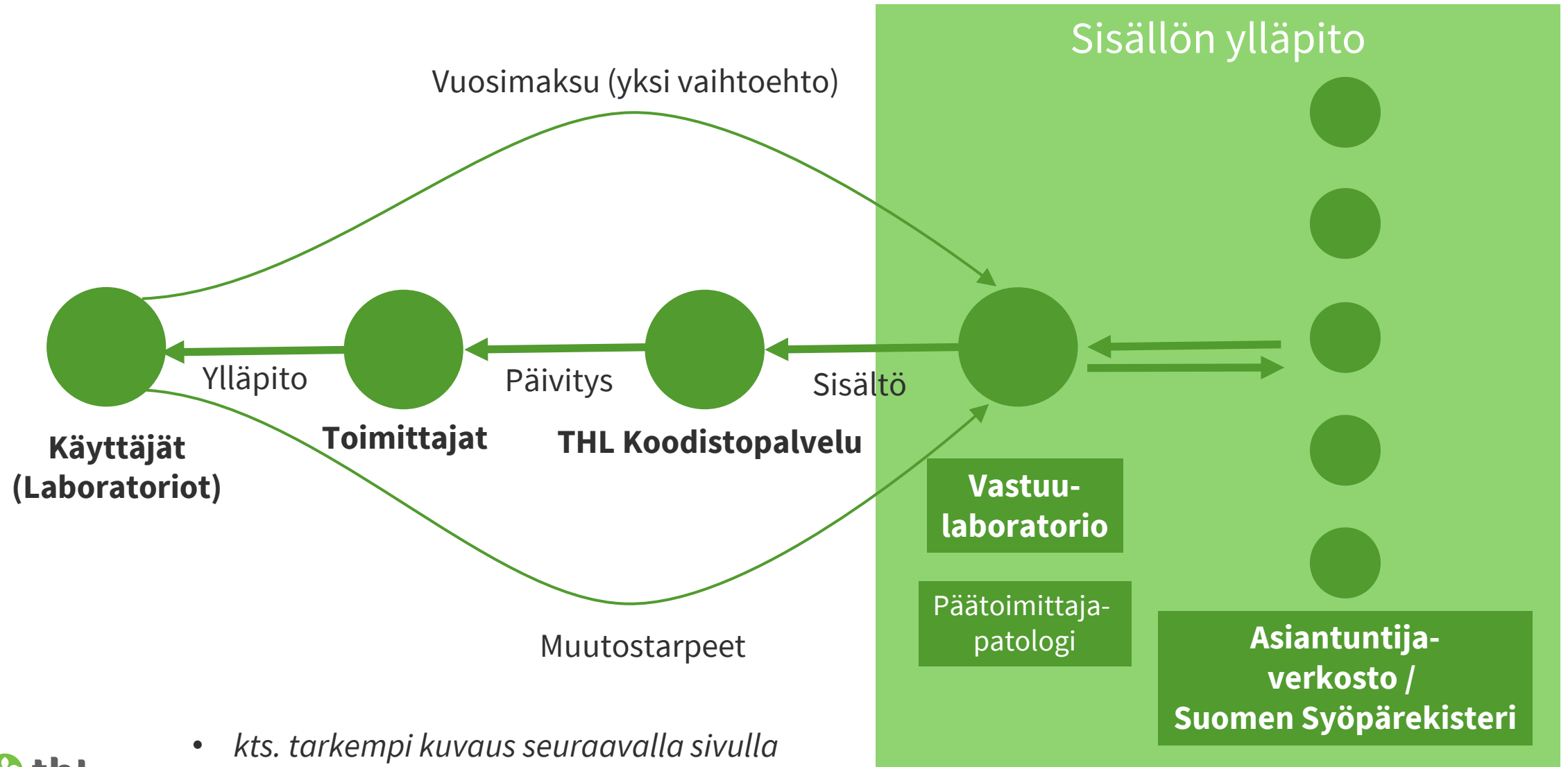
SNOMED CT Patologian löydösluokitus

- Muut tuotettavat tietoaineistot

- Lisäksi SNOMED CT patologian löydösluokitukseen kuuluu joukko lisätietoaaineistoja, joiden avulla helpotetaan SNOMED CT järjestelmän hyödyntämistä laboratorion tietojärjestelmässä.
- Alla on luettelo (tässä vaiheessa) suunnitelluista lisätietoaaineistoista:

Aineisto	Kuvaus
Käsitteiden väliset hierarkiat	Kaikkien Body structure ja Morphologic abnormality käsitteet ja niiden alakäsitteet. Voidaan käyttää esim. hakutoiminnallisuudessa kun haetaan jonkin käsitteen / termin kaikkia alakäsitteitä ilman että tarvitsee luetella niitä kaikkia.
Parilliset anatomiset sijainnit	Voidaan käyttää tarkistukseen yms. käyttöliittymässä – tähän tauluun kuuluville 'body structure' tiedoille pitää (pitäisi) antaa myös puolisuutta kuvaava tieto.
Disorder -> Topography, Morphology	Taulun avulla esim. maligniteeteissa ja tulehduksissa Disorder käsitteet voidaan purkaa tietokantaan myös Body Structure ja Morphologic abnormality kenttiin.
SNOMED CT → ICD-O-3	Mappaustaulu morfologian osalta SNOMED CT SCT_ID → ICD-O-3 (Suomen Syöpärekisterin ylläpitämä versio). Tämä taulu sisältää myös tiedon siitä, pitääkö tietty löydöstieto siirtää Syöpärekisteriin. (tieto on myös SCT PAT aineistossa)

SNOMED CT PAT luokituksen sisällön ylläpito ja jakelu



SNOMED CT Patologian löydösluokituksen ylläpito

- Osapuolten suunnitellut roolit 1/2

- Suunnitelman mukaan sisällön ylläpitoa koordinoivana ”vastuulaboratoriona” toimii ensimmäiset 3-4 vuotta FIMLAB.
 - Vastuulaboratoriossa on päätoimittajapatologi ja tarvittava IT-tuki, jotka vastaavat sisällöstä ja luokituksen sisällön ylläpito prosessista
 - Samalla pyritään huolehtimaan osaamisen jatkuvuudesta ja osaamis pohjan laajentamisesta
- Asiantuntijaverkostoon kuuluu erikoisalueittain yo-sairaaloiden nimeämät henkilöt. Kasvaintautien osalta Suomen syöpärekisteri vastaa asiantuntijaverkostosta ja sisällön ylläpidosta.
- International Academy of Pathology (IAP) Suomen osasto hyväksyy jakeluun menevän luokitusversion

SNOMED CT Patologian löydösluokituksen ylläpito

- Osapuolten suunnitellut roolit 2/2

- THL huolehtii Koodistopalvelun kautta tapahtuvasta jakelusta, yhteydenpidosta SNOMED CT ylläpitävään IHTSDO organisaatioon
- Tietojärjestelmätoimittajat vastaavat aineistojen toimittamisesta Kansallisesta koodistopalvelusta patologian laboratorioiden käyttöön taustarekistereihin ylläpidon kannalta sopivassa muodossa
- Kussakin laboratoriossa diagnoosirekisteristä vastaava patologi ja käytännön ylläpidosta vastaava tietojärjestelmän pääkäyttäjä toimivat yhdyshenkilöinä vastuulaboratorioon. Lisäksi he vastaavat oman laboratorion luokituksen ajankohtaisuudesta ja mm. muutostarpeisiin vastaamisesta.
- Kansallisesta Patologian löydösluokituksesta julkaistaan uusi versio kaksi kertaa vuodessa. Kukin laboratorio vastaa omien taustarekistereittensä päivittämisestä.

SNOMED CT PAT soveltamisesta patologian tietojärjestelmissä 1/2

- Yksityiskohtaisia ohjeita (puhumattakaan määräyksistä) siitä miten SCT PAT luokittelua tulisi soveltaa tietojärjestelmässä, ei voida antaa (eikä ole syytäkään)
- Käyttäjien ja tietojärjestelmien kannalta minimaalinen tapa toteuttaa SCT PAT luokituksen käyttöönotto on rakentaa sen avulla taustarekisteriin mappaustableu (vastaavuustableu), jossa on järjestelmän sisäisten koodien → SCT PAT käsitekoodien väliset vastaavuudet
 - Mappaustableun avulla löydöstietoja voidaan käsitellä myös kansallisen standardin mukaisesti: esim. laboratorion tilastoinnissa, siirrettäessä tiedot tietovarastoon ja kun tietoja toimitetaan Suomen Syöpärekisteriin ja/ tai Kanta-palveluun.
 - Mappaustableun avulla voidaan myös osaltaan ohjelmallisesti päätellä löydöstiedon kuuluvan Suomen syöpärekisterille ilmoitettavien joukkoon

SNOMED CT PAT soveltamisesta patologian tietojärjestelmissä

- SCT PAT avulla voidaan toteuttaa monipuoliset hakupalvelut – erilaisille termeillä löytyy yksi ja sama käsite.
- SCT PAT avulla voidaan parantaa tiedonsyötön tarkistuksia – parilliset elimet käsitejoukon avulla voidaan tietyissä topografioissa pakottaa käyttäjä antamaan puolisuustieto
- SCT PAT mahdollistaa käytön kannalta tehokkaan kantarakenteen, jossa syötetyn tiedon perusteella voidaan populoida aina T ja M kentät laadukkaita hakuja varten
 - kun löydös ilmoitetaan valitsemalla vastaava disease/disorder luokka, niin ohjelmallisesti voidaan ”populoida” topografia ja morfologia –kentät
 - Siten esim. haut saadaan yksinkertaisimmaksi ja kattavammaksi.
 - Kts. esimerkki sivulla 3.
 - Toinen esimerkki – syötetty disease käsite: *erosive gastritis* → käyttöliittymä populoi topografiakenttään: *stomach structure* ja morfologiakenttään: *ulcerative inflammation*.

(jatkuu)

SNOMED CT PAT soveltamisesta patologian tietojärjestelmissä

- Poiminnat Suomen syöpärekisteriin 1/2

- Suomen syöpärekisterin (SSR) ylläpitää SCT PAT:ssa kasvaintietojen osalta tietokokonaisuutta joka palvelee standardointia ja myös rekisteritiedon keruun laatua
 - Suomen syöpärekisteri koordinoi kasvaintautien asiantuntijaryhmän toimintaa (täältä osin yhteistyökumppani vastuulaboratoriolle)
 - Ylläpidettävä tietoaineisto pyrkii sisältämään tiedot, joiden avulla laboratoriotietojärjestelmässä voidaan mahdollisimman pitkälle algoritmisesti päätellä mitkä näyte/vastaustiedot kuuluu tiedonsiirron piiriin
- (jatkuu)

SNOMED CT PAT soveltamisesta patologian tietojärjestelmissä - Poiminnat Suomen syöpärekisteriin 2/2

(jatkoa edelliseltä sivulta)

- Suunnitelman mukaan SSR siirtyy SNOMED CT PAT mukaiseen määritelmään laboratorioden tiedonsiirrossa. Pilotointi on mahdollista käynnistää vuoden 2021 kuluessa.
 - Laboratorioilmoituksissa siirtyvä tieto on SCT PAT mukaan koodattua + alkuperäiset termi/tekstitiedot
 - Sisäisesti SSR koodaa Syöpärekisteriin löydökset ICD-O-3.2 mukaisesti
-
- Vastaavalla tavalla voidaan toteuttaa tiedonsiirrot LIMS – Kanta palvelu (2024-?), LIMS – Biopankit, LIMS – Shp tietovarasto

Mistä SNOMED aineistot saa käyttöönsä?

- Perehtymiseen ja suunnitteluun yms. käyttöön browser.ihtsdotools.org kotisivulla oleva SNOMED CT selain riittää pitkälle
- Vaativampaan kehittämistyöhön voi olla tarpeen ottaa käyttöön paikallisesti SNOMED CT määrittelyaineisto – helpottaa merkittävästi kehittelytyötä SNOMED CT Patologian löydösluokitusprojektin kokemusten mukaan
- Miten edetä:
 - Ensin liitännäissopimus/-jäsenyys, haetaan mlds-palvelimen kautta – olisi toivottavaa joka tapauksessa koska siten paremmin THL kertyy tieto käyttäjistä yhteydenpitoa varten
 - Aineistot saa käyttöönsä mlds-palvelimelta kun käyttäjätunnus on kunnossa. Määrittelyaineisto on käytännössä joukko SNOMED CT lataustiedostoja (tsv-muodossa)
 - Valmiit perl ja SQL –skriptit on saatavissa Githubista, esim. MySQL kantaa varten skriptit ja dokumentaatio: <https://github.com/rorydavidson/SNOMED-CT-Database/tree/master/mysql-loader-with-optimized-views>
 - Eri tietokantaversioille (ainakin Oracle, MS SQL Server, PostgreSQL) on olemassa vastaavat sivut.



Kiitos!