PROJECT TITLE: Anatomy alignment of SNOMED-CT and the Foundational Model of Anatomy Ontology (FMA)

OBJECTIVE:
1. Evaluate and verify the accuracy and correctness of existing mappings between SNOMED CT anatomical terms and the FMA concepts.
2. Deliver the first set of artifacts that includes the previously mapped (Map_asserted) and unmappable (Asserted_no_map) anatomical SNOMED CT concepts.

METHODOLOGY:

From the original SNOMED CT Access database, we derived data from four column fields, namely: *E-sepid*, *base name*, *E-name* and *FMAID* and transferred them to a new workbook where we added 4 new column headings for evaluation purposes: *FMA preferred term* (human-readable terms to assist FMA editors in their review), *Updated FMAID*, and *Updated FMA term* (fields for correction of erroneous or inappropriate mappings) and *Notes* (Report of errors/inconsistencies and suggestions for resolution). This new workbook was pushed to a web repository supporting concurrent edits.

In the early stages of the review process we observed significant discrepancies between the entries in the *base name* and *E-name* columns. We therefore created the *Term mismatch* column to record and report these discrepancies to the SNOMED CT editor(s) for further review, evaluation and resolution. We classified the discrepancies into 5 categories, namely:

  a. count disparity: singular vs plural concepts (**C**)
       e.g. *base name* "paraganglion", *E-name* "Entire paraganglia (body structure)"
  b. different entities represented (**D**)
       e.g. *base name* "skeletal system", *E-name* "Entire skeleton (body structure)"
  c. 'E-name' missing (**M**)
       e.g. *base name* " bone tissue of maxilla", *E-name* (0)
  d. term "Entire" missing (**ME**)
       e.g. *base name* "lumbosacral joint", *E-name* 'Lumbosacral joint"
  e. synonyms (**S**)
       e.g. *base name* "skull", *E-name* "Entire cranium (body structure)"

The evaluation process for both "Map_asserted" and "Asserted_No_Map" work lists entailed at least two cycles of review, with a third review focused on all concepts that were previously identified and determined to be problematic or erroneous. Where the semantics of a SNOMED CT term was in question, we based evaluation and mapping on one or more of the following:

1. consulted with the anatomy domain SNOMED CT expert for the definition or clarification on the intended meaning of a term
2. consulted the SNOMED CT anatomy OWL file to determine the intended semantics as declared by the properties associated with the concept/term
3. consulted the IHTSDO SNOMED CT browser (http://browser.ihtsdotools.org/#) for the properties associated with the term (e.g. References to clinical entities).

RESULTS:

A. Evaluation of SNOMED CT "Map_Asserted" anatomical concepts.

The "Map_asserted" file includes SNOMED CT anatomical concepts previously mapped to the FMA by IHSTDO editor(s).

1. Incorrect mappings

Of the **8724** concepts in the "Map_Asserted" list, we identified **314** were incorrectly mapped to the FMA. (e.g. SNOMED CT term "posterior muscle of thigh" (E00335) mapped to FMA term "Set of hamstring muscles" [FMAID 81022]. The correct map should be to FMA term "Muscle of posterior compartment of thigh" [FMAID 22427]). All corrections were entered in the *Updated FMAID* and *Updated FMA term* columns.

2. Updated FMA preferred terms

There were **28** entries where the preferred FMA terms were erroneous. Updated preferred terms were entered in the *Updated FMA term* column only, while retaining the same FMAID. Example: Entry E11141 has *FMA preferred term* "Supraspinous ligament of T1-T2 vertebrae" updated to "T1-T2 segment of supraspinous ligament".

3. Ambiguous mappings

3.1 Unspecified structure

There were **13** entries that represent non-specific anatomical entities that were ambiguously mapped to specific FMA concepts. Example: SNOMED term "modiolus" (E00117) was mapped to FMA term "modiolus of cochlea" [FMAID 61278]. The term

"modiolus" is applied to different entities, such as "modiolus of cochlea" and "modiolus of labial part of mouth" [modiolus of face]. The concept should be explicitly specified to avoid confusion. These entries were flagged in the *Note* column.

3.2 Inappropriate SNOMED CT terms

Terms pertaining to the parts of the lung, such as lobes and bronchopulmonary segments, were inaccurate and misleading. For example: SNOMED CT term "right upper lobe of lung" (E01320) implies that a lung also has a left upper lobe. A lung does not have a right and left upper lobe. There is a pair of lungs, right and left, and each has an upper lobe, hence, "upper lobe of right lung" and "upper lobe of left lung". The inappropriate terms are from entries E01320 to E0133 and E01335 to E01345. We recommend the use of the FMA preferred terms for these entries, as indicated in the *Note* column.

3.3 Semantic conflict

This section describes a semantic conflict in term usage between the SNOMED CT and the FMA.

    3.3.1   Tree structure vs. trunk

        3.3.1.1 Artery
        In the FMA, an artery or a vein represents an arterial tree or venous tree, respectively. However, many of the SNOMED CT entries for "artery" are mapped to the FMA arterial trunks, e.g. "coronary artery" (E00761) mapped to the FMA "Trunk of coronary artery" (FMAID). In some cases, SNOMED CT arteries were directly mapped to FMA arteries because the corresponding arterial trunks were not present in the FMA. But there were also instances where the arterial trunks exist in the FMA and yet some SNOMED CT arteries were still directly mapped to the FMA arteries, e.g. "superior anterior alveolar artery" (E04543) mapped to FMA "Anterior superior alveolar artery" (FMAID 49771). We checked the context by which the SNOMED CT artery was linked to clinical entities and we found in some cases that the term was used to imply both the trunk and the arterial tree. Example: "ulnar artery" (E04854) was linked to "Injury of ulnar artery at wrist" (trunk context) as well as to "Poor ulnar perfusion of hand" (tree context).

        3.3.1.2 Vein
        There was only direct mapping between the SNOMED CT and the FMA veins. We did not find any vein to trunk mapping in the list.

        3.3.1.3 Nerve
        We observed 4 different mapping patterns for the nerves:

- SNOMED CT nerve to FMA nerve (organ to organ)
- SNOMED CT nerve trunk to FMA nerve trunk (organ part to organ part)
- SNOMED CT nerve to FMA nerve trunk (organ to organ part)
- SNOMED CT nerve to FMA peripheral segment of nerve (organ to organ part)

The inconsistent mapping patterns will lead to confusion and, most likely, to errors, especially when such information is to be used for intelligent query or reasoning. We have identified a total of **1145** cases in this category and we flagged these entries in the *Note* column. For resolution, consistent and precise semantics must be assigned to "artery", "vein" and "nerve".

3.3.2   3-D surface vs. 2-D surface

"Surface" in the FMA is a 2-D immaterial anatomical entity and serves as a boundary for a 3-D material anatomical entity. In SNOMED CT, "surface" is a 3-D entity, based on its *is_a* attribute where for instance SNOMED CT "mesial surface of tooth" (E04985) *is_a* "Tooth part". "Surface" in this case refers to the surface layer (outer 3-D layer) of the tooth. The mappings therefore to the FMA "surface" concepts are incorrect. There were **153** cases of this type of conflict that are flagged in the *Note* column. Should the IHTSDO editor decide to keep the SNOMED CT "surface" concepts as a 3-D entity, then the appropriate mappings should be to the FMA "surface layer" concepts.

4. Term mismatch

There were **3417** cases of term mismatches between corresponding entries in *base name* and *E-name* columns, with the following breakdown:

Category	number of terms
C	68
D	66
M	2323
ME	663
S	297
Total	**3417**

The above discrepancies can contribute to problematic mappings. In some cases base names were used for mapping while in other cases, E-names were used. As examples, in one case, *base name* "paraganglion", which has correspondence to *E-name* "Entire paraganglia (body structure), was mapped to FMA "Paraganglion" but

in another case, *E-name* "Entire renal ganglia (body structure)", which has correspondence to *base name* "renal ganglion", was mapped to FMA "Set of renal ganglia". In each case, they share the same E-sepid but the base name is a different concept or entity from the E-name. In the case of missing E-names, mappings were based only on base names.

In an accompanying Excel file called "Extract of Base name - E-name mismatch_01-30-2015.xlsx", we clustered the entries to facilitate analysis of entries in each category.

In category **C**, there were **68** mismatches, where base names are in singular form while the corresponding E-names are in plural form. Some are mapped to singular FMA concepts while others are to the plural concepts. As a solution we suggest the creation of additional concepts to represent the two concepts separately. Retain the existing entries with singular base names and assign them their singular E-names Then create their corresponding plural concepts (both *base name* and *E-name*) as new entries in the table with new internal IDs (E-sepid). Re-submit these entries for proper FMA mapping.

In Category **D**, there are **66** mismatches. Retain the *base name* as the primary concept in each entry and assign the same term for its corresponding *E-name*, except for a couple of base names marked in RED. In these cases we suggest using the *E-name* as the primary concept and changing the *base name* to match the "structure concept" of *E-name*. For example, the *base name* "maxillary alveolus" should be changed to "maxillary tooth socket", which was derived from the E-name "Entire maxillary tooth socket (body structure)". Subcategory called **D ME** includes E-names that are missing the term "Entire".

Category **M** has the main bulk of mismatches, totaling around **2323** entries. Create E-names using the base names. Entries E13926 and E13927 should be deleted because they are not canonical structures.

In Category **ME** add the term "Entire" missing from the **663** E- name entries.

Category **S** includes about **297** term mismatches that are in fact synonyms. We suggest that the *base name* be treated as the primary or preferred name except for all E-names highlighted in RED in which case we suggest that the E-name be the preferred name. Subcategory **S ME** lists entries missing the term "Entire" from the E-names.

B. Evaluation of SNOMED CT "Asserted_no_map" anatomical concepts.

The "Asserted_no_map" file includes SNOMED CT anatomical concepts previously deemed unmappable to FMA concepts.

1. Completed mappings

Of the **1506** entries, we were able to map **549** SNOMED CT terms to the FMA concepts. The process required the creation of **311** new concepts (terms) in the FMA while the remaining concepts already existed in the ontology. The mappings were entered in the *Updated FMAID* and *Updated FMA term* columns.

2. No equivalent FMA concepts

We have verified and confirmed that **957** SNOMED CT concepts were unmappable to the FMA based on a number of reasons:

2.1. Ambiguous terms
2.1.1 Non-specific terms
There were **12** cases where the concepts represented by the terms were not precisely specified. Example: "axillary vessel" (E13206) may refer to an artery, a vein or a lymphatic vessel.

2.1.2 Unconventional partition
The terms in this category represents concepts defined by unconventional methods. The term "upper limb distal to the distal humeral metaphysis" (E20496) implies that the fiat boundary is the plane that bisects the metaphysis. This may be a special partition for a specific application but we have not been able to identify what application uses it. We can incorporate the concept in the FMA as a regional part of the upper limb but we will await clarification from the IHTSDO editor(s). There are **7** cases of this kind.

2.2. Inappropriate terms
We have identified **6** concepts that improperly associated immaterial anatomical entities (spaces) with material anatomical entities (anatomical structures). Spaces cannot have material components, as in the case of "bone tissue of cranial cavity" (E09588).

2.3. Ontological issue
The FMA adheres to a set of principles that defines the taxonomic framework of the ontology, which is a single inheritance hierarchy. Classes in the FMA are assigned to defined organizational units, such as organ, tissue, cells and anatomical spaces. The SNOMED CT concept "musculoskeletal entity" (E19953) refers to any structure that belongs in the musculoskeletal system and that can either be a muscle (organ), a bone (organ), a joint, a cartilage (organ) or a skeletal ligament (organ). In the FMA, each of those structures is assigned to a specific class that is defined in terms of specific attributes, e.g. bone consists of regular connective tissue while muscle consist of muscle fibers. Hence, separate parent organ classes are created to group together structures of the same types. However should there be a need for "broad classes" such as "musculoskeletal entity" for a

specific application, the classes can be generated and logically defined using existing content: A 'musculoskeletal entity" is a subclass of a class that is an organ or an organ part which is a constitutional part of the musculoskeletal system.

For terms representing 2 or more structures as combined sites (e.g. "bronchus and skin of thorax" [E17277]), they can be defined as subclasses of classes that has members organs and/or organ parts.

A total of **320** concepts falls into this category.

2.4. Non-structural context

2.4.1. Functional context

There were **84** terms that represented concepts based on function, namely, muscles acting on joints (n=**56**) and functional spinal units (n=**28**). Depending on the hierarchical structure of SNOMED CT anatomy, these concepts may be associated with structural entities and defined based on structural classes and entities. Example: "Muscle acting on hand joint" (E15326) may be logically defined as a muscle organ whose part has insertion to part of a bone that is part of a joint in the hand.

2.5.2 Clinical context

Some concepts are beyond the domain of canonical anatomy. "Fetal vertex" (E06286) and "fetal parietal boss" (E06285) are obstetrical terms while "acanthocyte" (E18551) and "tumor cells" (E19195) are pathological structures. We have identified **10** concepts in this category. However canonical anatomy can serve as an organizing framework for these clinical entities.

2.5. Non-canonical anatomical entities

We have identified **7** non-canonical concepts, e.g. "intervertebral disc space of sixth lumbar vertebra" (E01099). Canonical anatomy can be extended to accommodate normal variations.

2.6. Lack of information

There are **497** terms that represent anatomical landmarks for procedures, such as acupuncture (Huatuochiachi points [E18253]) and ECG (Lead I [E17873]). We don't have enough reliable information to specify the anatomical location of the different acupuncture points. Should we obtain proper information for these points, we can accommodate them in the FMA.

All comments and suggestions for the unmappable SNOMED CT terms are entered in the *Note* column.

3. Term mismatch

There were **1351** cases of term mismatches between corresponding entries in *base name* and *E-name* columns, with the following breakdown:

Category	number of terms
C	4
D	1
M	449
ME	867
S	30
Total	**1351**

As in section A4 above, use the base names for creating the E-names in category **M** and add the term "Entire" to E-names in **ME**. In category **C**, retain the current mappings and create, as new entries, the corresponding singular or plural form for those existing terms. We recommend using the base names as the preferred names in the **S** category.

C. For this quarter we are submitting as deliverables this report, two EXCEL files "SNOMED-FMA_mappings_01-30-2015.xlsx" and "Extract of Base name - E-name mismatch_01-30-2015.xlsx".