

201102004A

厚生労働科学研究費補助金
政策科学総合研究事業（統計情報総合研究）

I C D 1 1 におけるオミックス情報モデルの研究

平成23年度 総括・分担研究報告書

研究代表者 中谷 純
（東京医科歯科大学）

平成24（2012）年 5月

目 次

I. 総括研究報告	
I C D 1 1 におけるオミックス情報モデルの研究 中谷 純	-----1
II. 分担研究報告書	
I C D 1 1 におけるオミックス情報モデルの研究 今井 健	----- 14
III. 研究成果の刊行に関する一覧表	----- 27

平成 23 年度厚生労働科学研究費補助金（政策科学総合（統計情報総合）研究事業
総括研究報告書

「ICD11におけるオミックス情報モデルの研究」
研究代表者 中谷 純 東京医科歯科大学 特任准教授

研究要旨 2014 に改定が予定されている WHO-ICD11 では、医療情報の電子化の流れに対応することを目的として、情報モデル、オントロジック的整理を導入し、多軸構造、多言語対応、統一的用語解説などを実現しようとしている。この ICD11 では、オミックス時代に対応するためジェノミクス情報を取り込むことが計画されているが、ジェノミクス情報を組み込むためには、臨床コンテンツモデルと整合性のあるオミックス情報モデルを作成することが必要である。さまざまな臨床医学においてオミックス情報を利用することが今や常識的なアプローチとなってきた中、ICD11 に親和性を持つ臨床オミックス情報モデルを作成することは、ICD11 の可用性を高めオミックス領域への親和性を確保し、医科学全体の発展を促す上で最も重要でありかつ必須の課題である。本研究では、ISO において正式に国際標準(IS)と認められた日本発信の技術である ISO 25720 GSVML(Genomic Sequence Variation Markup Language) [1] [2]を元に、ICD11 コンテンツモデルに対応するオミックス情報モデルを作成する。

平成 23 年度は、平成 22 年度に行ったオミックス医学内の用語間の関連性、臨床オミックス情報モデルに必要な構成要素について検証作業に基づいて作成したアルファ版を土台として、標準臨床医学情報モデルとの連携性、概念対応の違いなど、臨床オミックス医学情報モデルを ICD11 の中で実効させるために必要な情報を明らかにしつつ、ベータモデルの作成を行い、その上で、ICD11 コンテンツモデルとの整合性検証を行う。

事実上の国際標準である ICD11 改定において追加されるオミックス情報カテゴリーをどのように記述するかは、今後の国際医学研究を左右する大きな課題である。日本が世界に先駆け国際標準として樹立した ISO 25720 GSVML 情報モデルを応用する形で、ICD11 に親和性のあるオミックス情報モデルを世界に先駆けて提示することは、日本発進の新技術として誇ることのできる大きな特色であり、独創的な点である。本研究はオミックス医学という分子生物医学の集大成ともいえる医学分野を現在の世界保健の主流である疾病分類に組み込むために必須のものであり、ICD11 に整合性の取れたオミックス情報モデルを世界に先駆け提示しようとする点で大きな意義をもつ時宜にかなった研究である。

研究分担者

田中 博 東京医科歯科大学 教授
今井 健 東京大学 助教

A. 研究目的

ICD11 コンテンツモデルに親和性を持つ臨床オミックス情報モデルを作成する。

B. 研究方法

1 臨床オミックス情報モデル（ベータモデル）の作成

先行して素描したアルファバージョンを元に、臨床オミックス情報モデル（ベータモデル）の電子化を行う。電子化は、主に XML 化することで達成する。その際、XML Editor などを用いて作成を行うが、必要に応じて Protege や LexWiki というオントロジーツールの利用も行う。特に、オミックスエンティティどうしの関連性表現について、その表現方法についての研究と開発が必要となる。α版でデザインされたコンテンツモデルサブ構造における、内部エンティティの相互関連性解析を行い、それに基づいて、内部相互関連を包含した XML 表現を行う必要がある。これは、GSVML で実現されている技術であるので、それを応用する形で研究開

発作成を行う。

2 ICD11 コンテンツモデルとの整合性検証

整合性検証は、ベータモデルで定義される項目ごとの意味的整合性とオミックス情報記述性、GSVML などの標準モデルとの整合性の観点から行う。また、臨床オミックス情報モデル作成に必要な構成要素、他の標準臨床医学情報モデルとの連携性、概念対応の違いなど、臨床オミックス医学情報モデルを ICD11 の中で実効させるために必要な情報を明らかにする。特に、オミックスエンティティどうしの関連性をインターフェース解析し、その表現方法をどのようにするかについての研究と開発が必要となる。α版で課題を指摘された内部的関連性をどのように表現するかがベータ版の整合性検証で重要となる。

3 WHO-TAG-HIM 会議への参加ならびに情報交換

随時ジュネーブで行われる会議に出席し、ICD のモデル化作業に関する情報収集と意見交換を行う。また、随時、WHO-FIC、内科 TAG などへも参加し、情報交換を行う。また、本プロジェクトはその重要性から、HIM-TAG を超え、別プロジェクトとして起動させるべきとの意見

もあり、そういった方向性にも対処する形で、会議への参加ならびに情報交換を行い研究へのフィードバックを行う。

C. 研究結果

1. ICD11 対応臨床オミックスサブ情報モデル (β版: iCOS β)

別添図1～図12にICD11オミックスサブ情報モデルのβ版 (ICD11 Omics Sub Information Model β版: iCOS β) を示す。添付資料に、XMLプログラムを示す。

図1にβの概略構造を示す。アルファ版と比べると、より機能が分割されたモジュール構成になっている。大きく分けると、OML (Omics Markup language)、ICD11 Entity、Transcription Moduleの三つのモジュールから構成されている。

OMLモジュールは、主にオミックスに関連した内容を記述するモジュールであり、ICD11entityモジュールはICD11コンテンツモデルに対応するモデリングモジュールであり、Transcription Moduleは、OMLとICD11 entityの間の転写関係を規定する双方向の転写モジュールである。

ICD11は、いったん、ICD11と鏡像関係を持つICD11 Entityに投射される。このICD11 Entityの中の、Causal Propertiesの中の、Genomic Linkagesから、Transcription Moduleを通じてOMLモジュールの中のオミックスデータカテゴリーにポイントされる。

Transcription Moduleは、ポイントの方向とその関係の信頼度を規定できるようになっている。つまり、Transcription Moduleで規定される関係がどちらからどちらへの方向性を持っているのか、その関係がどの程度信頼できるのかといったことを重みづけできるようになっている。これは、プログラム技術的には難解な命題であるが、矢印とその強さを、ベクトルエンティティとして定義することで実現した。

OMLの中のモジュールは、ISO国際標準 (IS) であるIS25720 GSVMLの構造を参考として、オミックス情報を許容できるように変更することで作成した。

図2は、iCOSの概略構成図である。iCOSは、ICD11 Entity、Transcription module、OMLの三つのサブ要素から構成される。Transcription moduleは、ICD11 OML TranscriptionとOML-ICD11 Transcriptionから構成される。OMLはOmics Markup Languageの略であり、Omics Date、Omics Annotation、ICD annotationから構成される。

図3は、ICD11 Entityである。これは、いわゆるICD11コンテンツモデルと等価なモデルであり、ICD11に対してON TO ON (1対1)で対応している。内部的にICD11へのミラーモデルを独立モジュールと

して持つことにより、発展途上モデルであるICD11に発生する微調整を緩衝することができると同時に、抜本的な改正に対しても、iCOS内部の変更を最小限に抑えることができ、iCOSのプログラム上のホメオスターシス (恒常性) を維持することができる。

図4は、Transcription Moduleである。Transcription moduleは、ICD11_OML_Transcription、OML_ICD11_Transcriptionという二つのモジュールから構成されている。それぞれのモジュールは、その投射方向により機能分割されている。ICD11_OML_Transcriptionは、ICD11からOMLへの投射を行う場合に使用するモジュールであり、OML_ICD11_Transcriptionは、OMLからICD11への投射を行う場合に使用するモジュールである。ICD11_OML_Transcriptionモジュールは、ICD11GenomicLinkageというICD11 Entityの項目とOmics Dataとの間の関係を一方向に規定する。関係の重みづけは、確信度という形で用意してあるが、自分なりの特殊な定義を用意して、参照する形で使用することもできる。OML_ICD11_Transcriptionモジュールは、OML ICD11 Annotation GroupというOMLの項目とICD11 Entityとの間の関係を一方向に規定する。こちらも、関係の重みづけは、確信度という形で用意してあるが、自分なりの特殊な定義を用意して、参照する形で使用することもできる。

図5は、OMLモジュールである。OMLは、Omics Data、Omics Annotation、ICD Annotationという三つのサブ構造から構成される。Omics Dataは、Variation Data、Transcription Data、Proteomics Data、Signalomics Data、Metabolomics Data、Organomics Dataといったプリセットの代表的オミックスデータ毎のモジュールとその他のオミックスデータをユーザーが定義して使用することができるようになっている。

図6は、OMLのオミックスアノテーションの構造を示している。オミックスアノテーションは、オミックスの種類ごとに選択できる構造となっている。簡便性のため、代表的なオミックス構造は羅列してあるが、当然網羅的ではないので、other_omics_annotationでユーザーがオミックスの種類ごとにアノテーション定義を行うことができるようになっている。このother omics annotationは、ネスト (階層再帰呼び出し; 入れ子) できるので、オミックス階層内だけでなくオミックス階層間にわたる情報も同時に記述することができる。

図7は、Genomic Sequence Variationの場合のAnnotation情報が書かれている。Whole Genome

Sequence, mendelian segregate, homozygote detect, somatic mutation, experiment analysis, epidemiology analysis, epidemiology_gsv, variation annotation といった、Variation データに特化したアノテーション情報を格納できるようになっている。

図8は、ICD Annotationの構造を示している。Personal Info には、オミックスデータを取得した個体(個人)に関する情報が入る。Phenotype には、オミックスデータからの連結結果としてのフェノタイプが入る。inter relation of omics annotation には、オミックスアノテーション同士の連関関係情報が格納される。environmental condition には、環境情報が格納され、ICD classification annotation には、ICD11 entity との関連性を示すマップ情報が格納される。Clinical annotation には、臨床アノテーション情報が格納される

図9は、Variation データを例とした場合の、ICD annotation 構造の説明である。

図10は、ICD annotation の詳細構造を示したものである。

図11は、clinical annotation の構造を示している。clinical annotation は、disease と clinical observation、clinical annotation database reference とからなる。disease は、単一疾病ごとの臨床アノテーションを記述することができ、clinical observation は現実の臨床例(複数の疾病が合併している状態)におけるその観察項目を記述することができる。いくつかの細項目が重なっているため、それらはプログラム上、entity として共通化した構造となっている。

図12は、Transcription module の詳細構造である。Omics ICD11 annotation group は、ICD Classification annotation と disease NCK から構成され、ICD classification annotation は、ICD concept title と ICD Classification properties とから構成される。Disease NCK は、Clinical annotation の中で、対象とするオミックスに関する臨床情報を疾病ごとに記述するための枠組みとして利用する。一方、ICD classification annotation は、疾病ごとに決まっている ICD コードおよびそのプロパティに対応させているものである。従って、これらに対応させることにより、ICD コードと対象とする臨床アノテーション情報を疾患ごとに対応させることができるようになる。

D. 考察

β モデルにおける課題は、オミックスアノテーションの表現妥当性を維持しつつ単純化することである。現在、オミックスアノテーションは、オミックスの種類ごとに選択できる構造

となっているが、この方式は苦肉の策であり、本来は単一の構造で、オミックス種別だけを選択する構造にするのが、情報学的には望ましい。しかし、オミックスは、分子に関するもの、ネットワークに関するもの、化学反応特性に関するもの、環境に関するものと一緒にオミックスといっても多数あるので、これらを単一の構造で表現することは難しい。 β モデルでは、簡便性のため、代表的なオミックス構造は羅列してあるが、当然網羅的ではないので、other_omics_annotation でユーザーがオミックスの種類ごとにアノテーション定義を行うことができるようになっている。この other_omics_annotation は、ネスト(階層再帰呼び出し; 入れ子)できるので、オミックス階層内だけでなくオミックス階層間にわたる情報も同時に記述することができる。このオミックス表現は、オミックス研究自体の発展に合わせて、その構造を変化させていく必要がある。

E. 結論

今後は、本研究で作成した β モデルを、実データを用いて検証しつつ、現実化に向けた改善を行う必要がある。我々の持つ国内唯一の臨床オミックスデータベースである、iCOD (integrated Clinical Omics Database)のデータを用いて、iCOS β を真の国際標準とすることができるよう、今後とも努力したい。

参考文献

- [1] Nakaya, J., Kimura, M., Hiroi, K., Ido, K., Yang, W., Tanaka, H. 2010. Genomic Sequence Variation Markup Language (GSVML). "International Journal of Medical Informatics(Int J Med Inform.)". Volume 79, Issue 2, 130-142. PMID: 19969503.
- [2] Nakaya, J., Hiroi, K., Yang, W., Ido, K., Kimura, M. 2006. "Genomic Sequence Variation Markup Language (GSVML) for Global Interoperability of Clinical Genomics Data(Best Paper Award)". Asia Pacific Association for Medical Informatics 2006 Proceedings. A01. 1-8.
- [3] Iwei Yeh, Peter D. Karp, Natalya F. Noy, Russ B. Altman. Knowledge Acquisition, Consistency Checking and Concurrency Control in Gene Ontology. Bioinformatics 2003. 19: 241-248.
- [4] Jun Nakaya, Tetsuo Shimizu: Knowledge Architecture based on Evidence Based Logical Atomism for Translational Research. 2006. International Journal of Computer Science and Network Security, vol. 6-2, ISSN:1738-7906, 175-179
- [5] SNOMED-CT, SNOMED International, URL = [http://www.snomed.org/index.html] Accessed at Feb. 22, 2011

F. 健康危険情報 なし

G. 研究発表

1. 論文発表

A) 中谷 純、田中 博. (2012). 地域医療・福祉ネットワーク化白書. シードプランニング. in Press

B) Jun Nakaya, Michio Kimura, Riichiro Mizoguchi, Kouji Kozaki, Kaei Hiroi, Keisuke Ido, and Hiroshi Tanaka. 2012. Semantic Mechanisms in Clinical Omics Database. InterOntology2012. pp 81-91.

2. 学会発表

招待講演

A) ICD11におけるオミックス情報モデルの研究. 中谷 純、田中 博. 母子愛育会 日本子ども家庭総合研究所、2011. “平成22年

度厚生労働科学研究 統計情報総合研究講演会:保健医療統計における国際基準への対応と活用”

B) ICD11 改訂とオミックスコンテンツモデル. 中谷 純、今井健、田中博. 幕張メッセ、2011. 第15回日本医療情報学会春季学術大会

C) オミックス医療情報その光と影. 中谷 純. 鹿児島、2011. 11. 21 第31回医療情報学連合大会 学会企画「オミックス医療情報 その光と影」

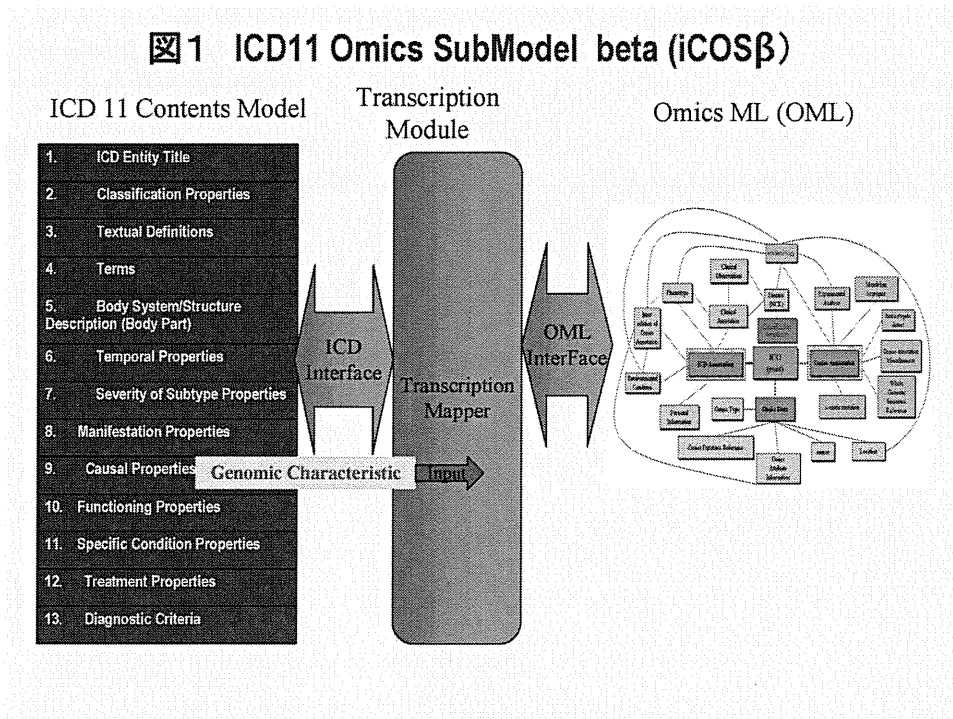
D) ICD11におけるオミックス情報モデルの基礎検討 . 今井健、中谷純. 鹿児島、2011. 11. 21. 第31回医療情報学連合大会 学会企画「オミックス医療情報 その光と影」

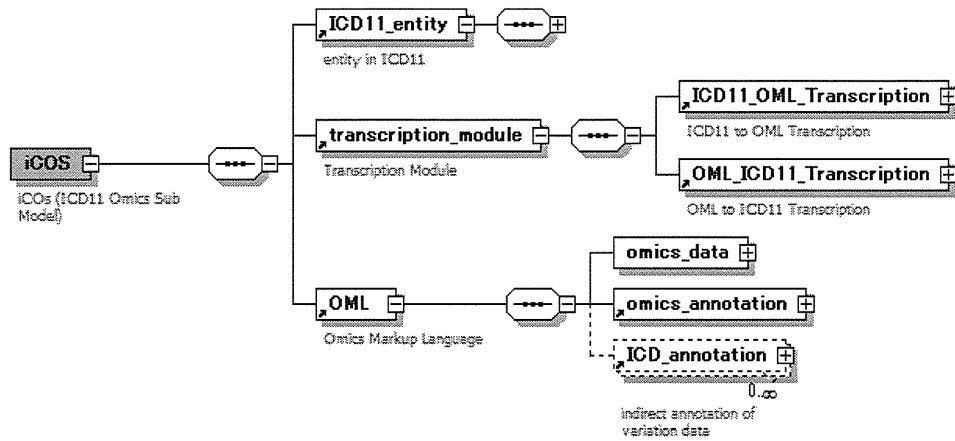
H. 知的財産権の出願・登録状況

1. 特許取得 なし

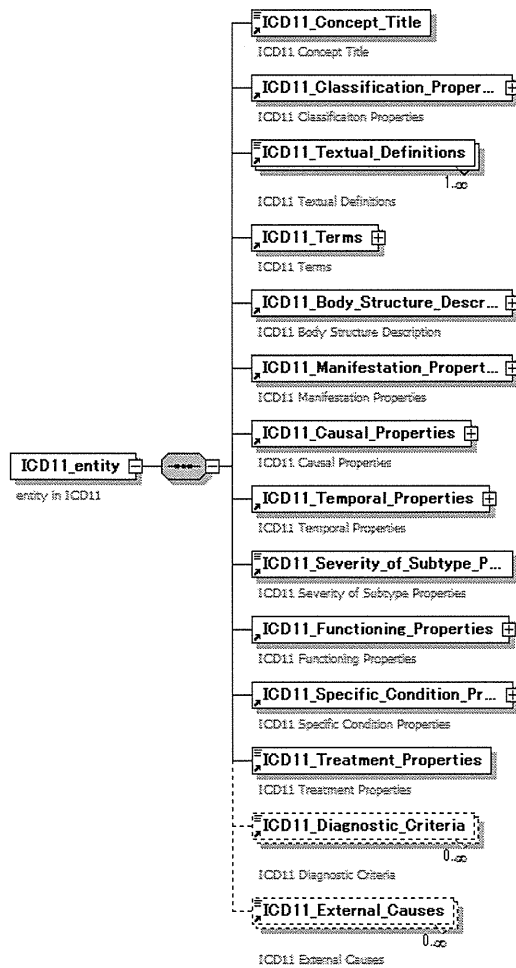
以 上

(別添図表)

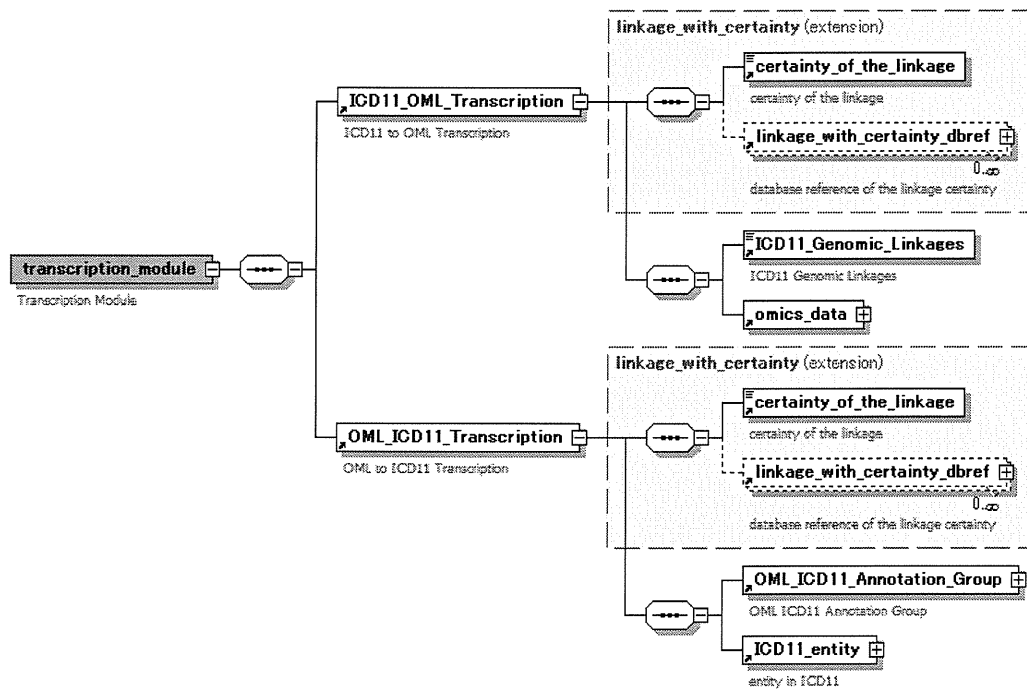




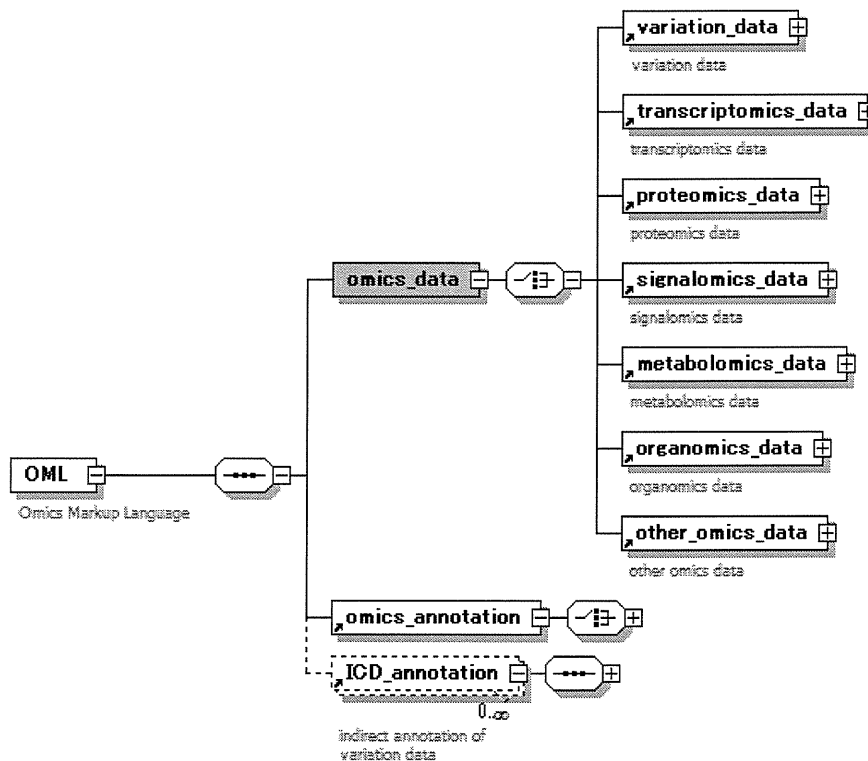
(图 2: iCOS 概略構成図)



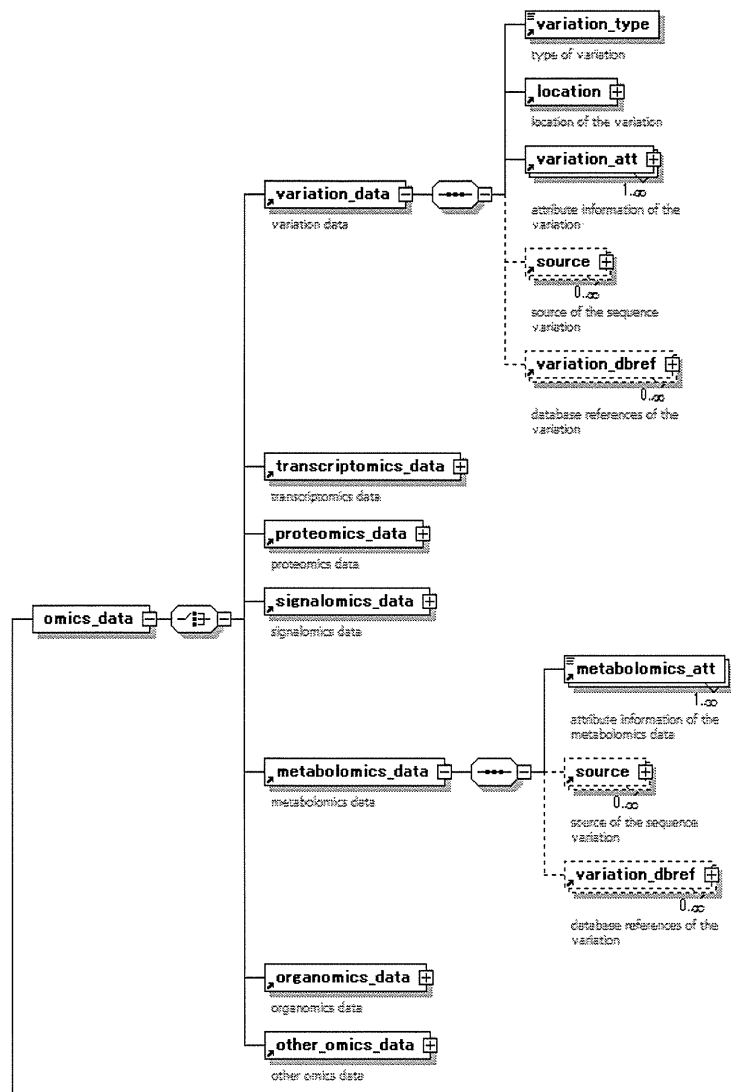
(图 3: ICD11 Entity)



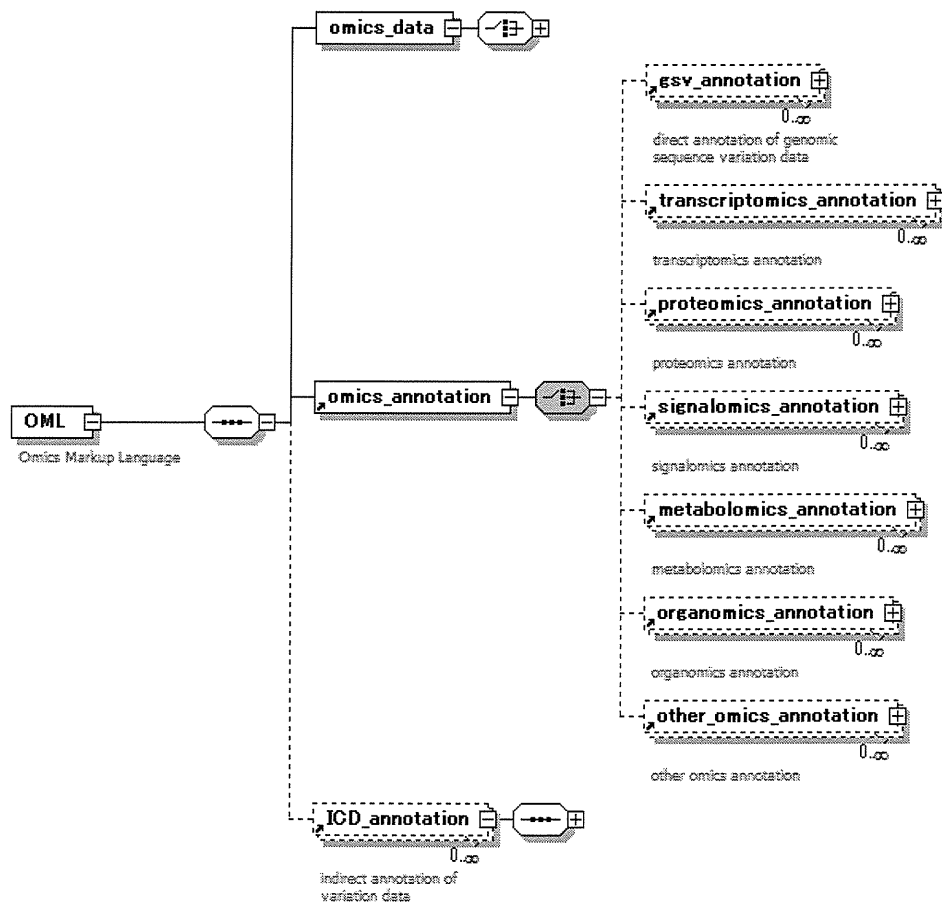
(図 4: Transcription Module)



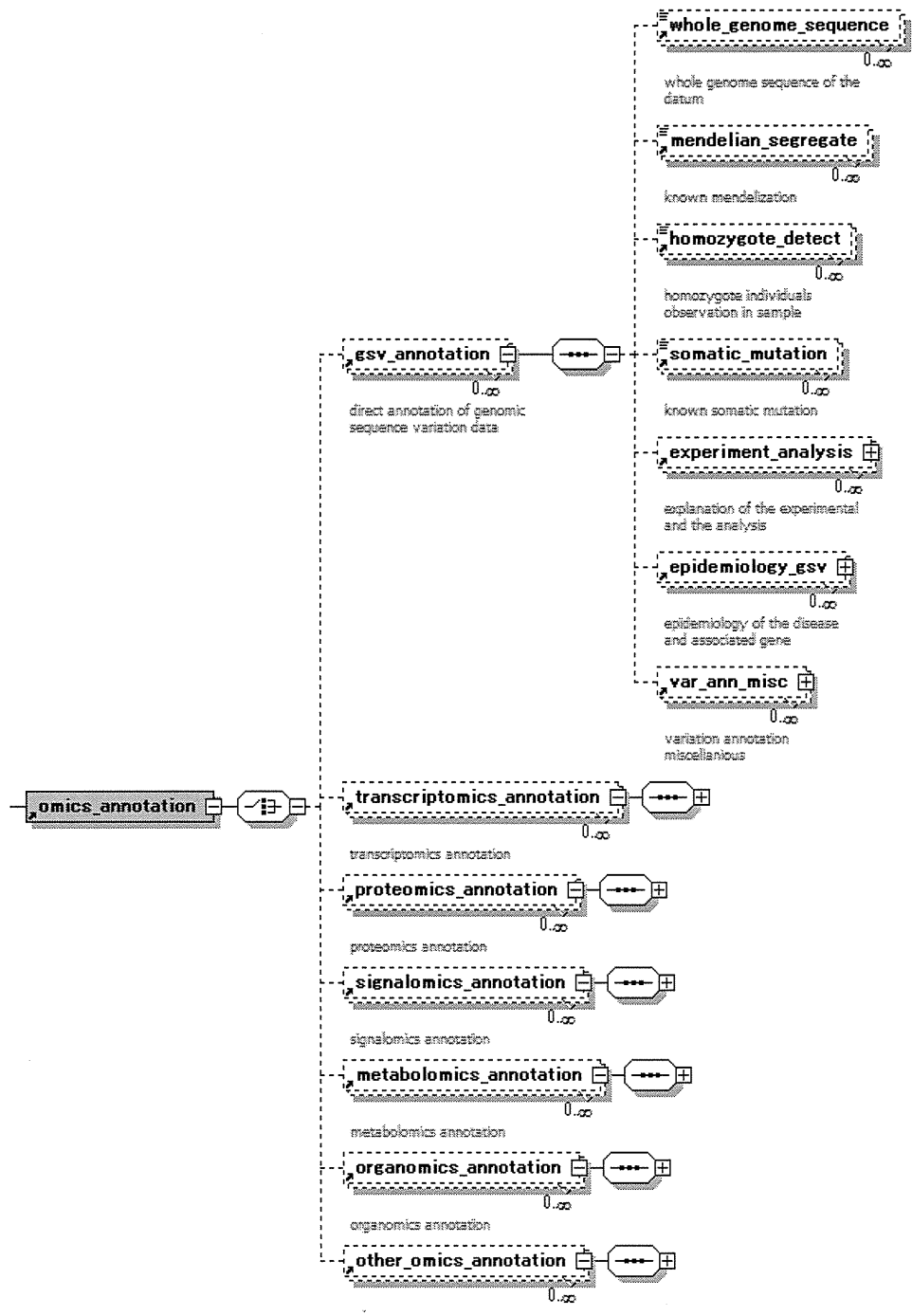
(図 5: OML モジュール)



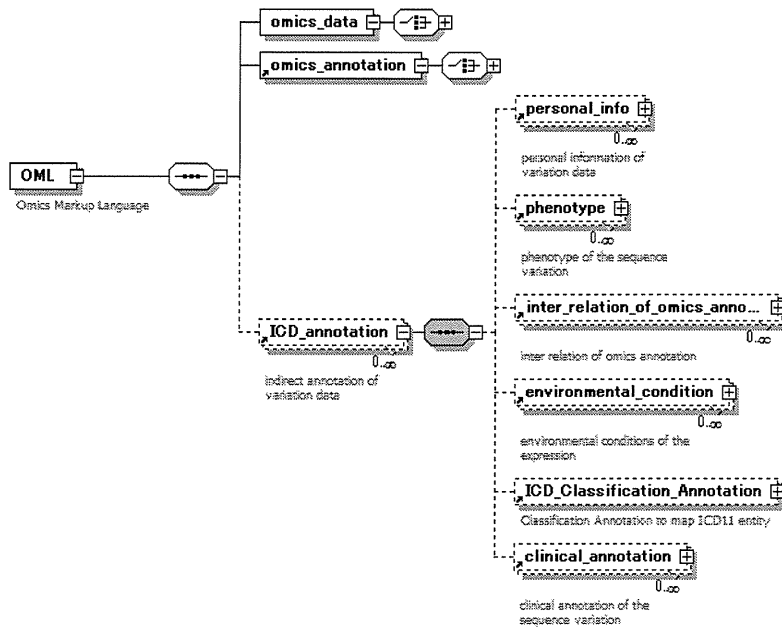
(図 6: オミックスデータの中の Variation Data とメタボロミクスデータ)



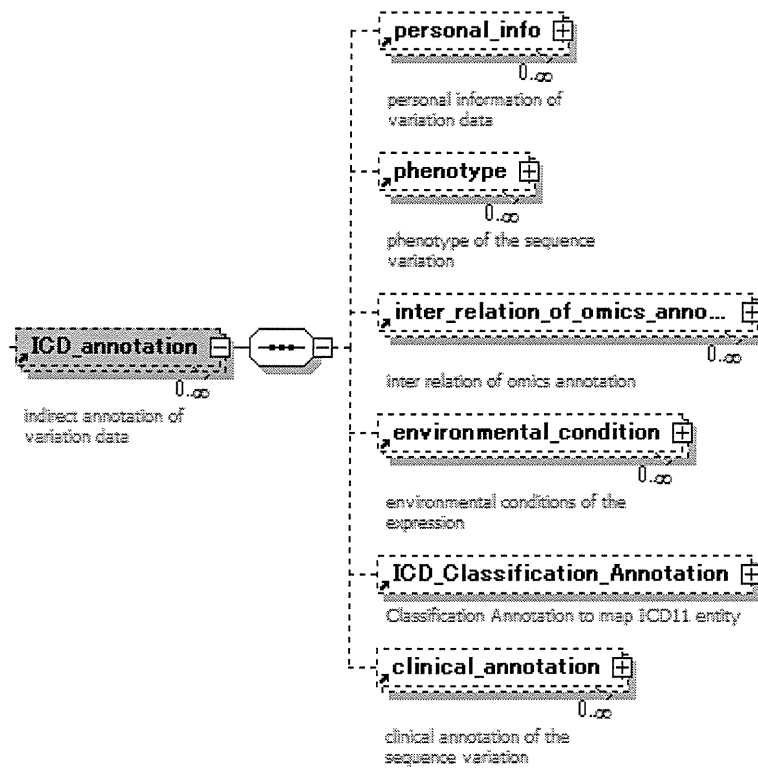
(図 7: OML のオミックスアノテーション構造)



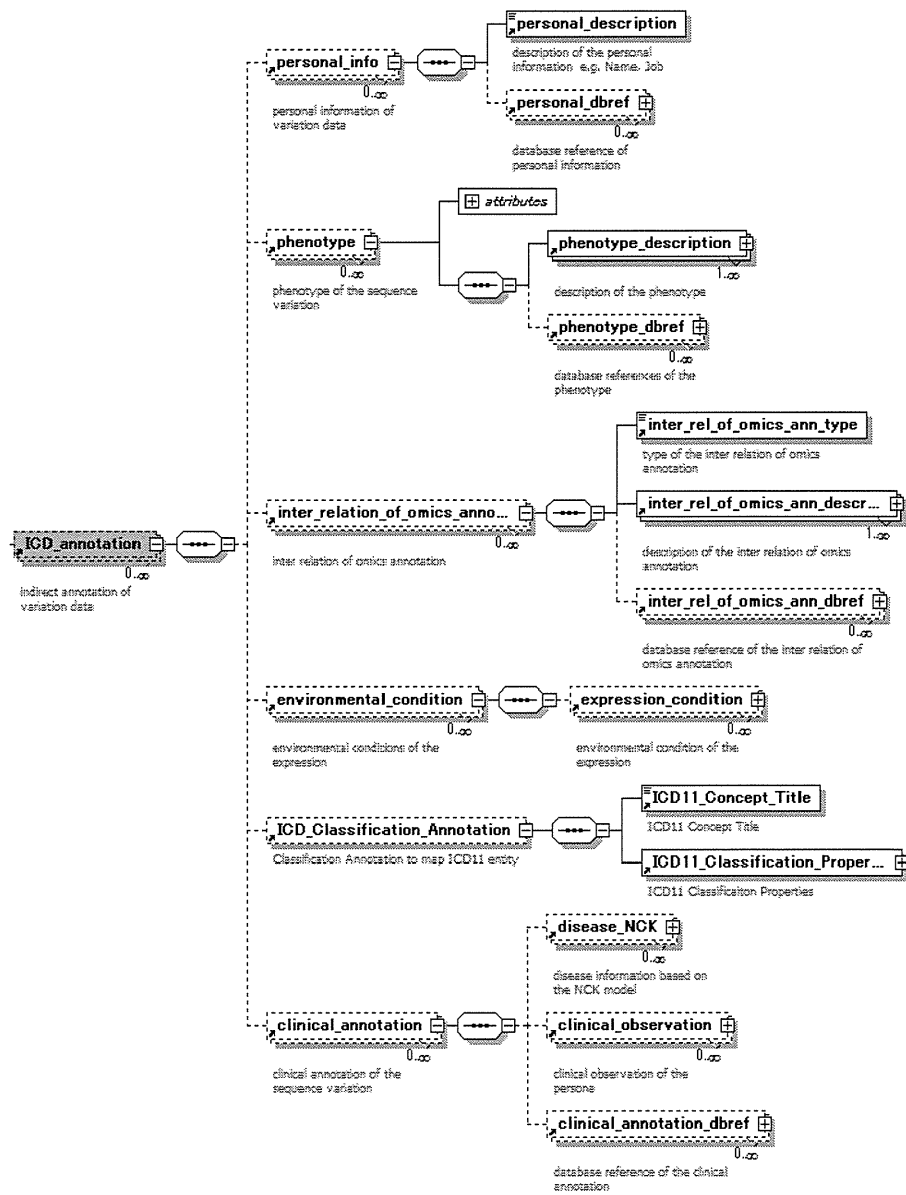
(図 8: Omics Annotation 情報 (Genomic Sequence Variation))



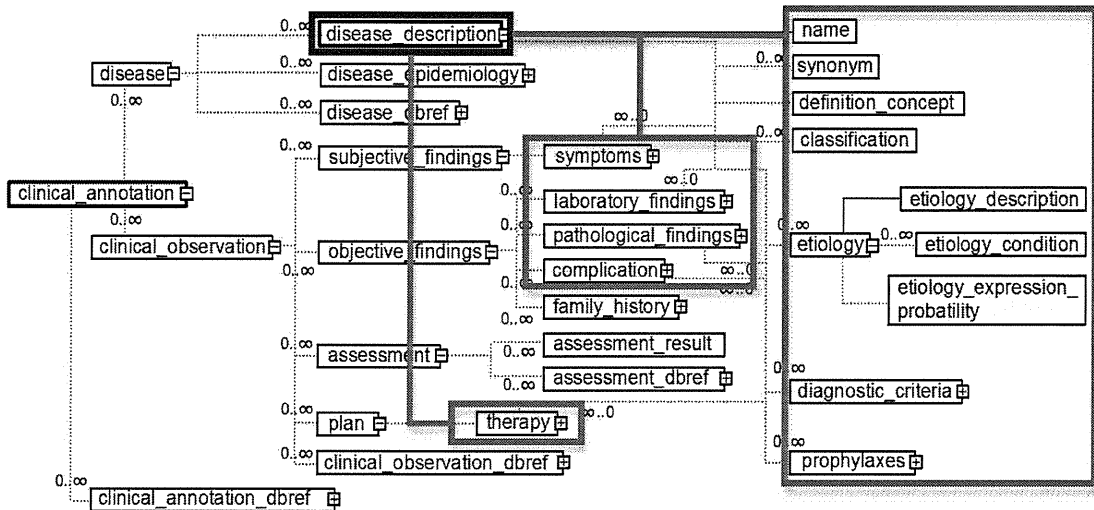
(図 9: ICD Annotation 構造)



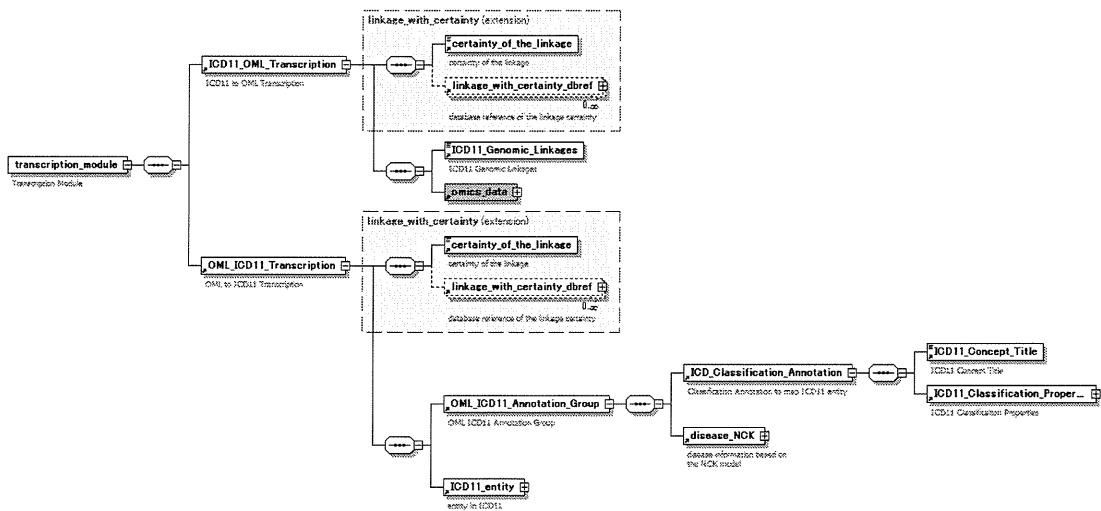
(図 10: ICD Annotation 構造 (Variation データの場合))



(図 11: ICD Annotation 詳細構造)



(図 12: Clinical Annotation 構造)



(図 13: Transcription module 詳細構造)

H23 年度 厚生労働科学研究費補助金
(政策科学総合研究事業 統計情報総合研究事業)
分担研究報告書

ICD11 におけるオミックス情報モデルの研究

研究分担者 今井 健 (東京大学 大学院医学系研究科 疾患生命工学センター 助教)

研究要旨

本研究全体の目的は、現在改訂作業が進められている WHO-ICD11 に親和性を持つ臨床オミックス情報モデル iCOS (ICD11 Omics Sub model) の構築である。本分担研究では、前年度に構築された iCOS- α 版を発展させ、ICD11 における疾患カテゴリーの特性記述モデル (Content Model) とオミックス情報を記述するための OML (Omics Markup Language) モデルを相互に接続する部分構造 (転写モジュール) を中心に、モデルの詳細化と基礎検討を行った。

A. 研究目的

2015 年に改訂が予定されている WHO-ICD11 では、従来と同様の分類階層構造改訂のみならず、“Content Model” と呼ばれる一種の情報モデルに基づいた各疾患カテゴリーの特性記述 (原因や症状といった各種の臨床的特性など) が大きな特徴となっている。これは昨今の医療情報の電子化の流れの中、単なる分類体系の提供にとどまらず、既存のデファクトスタンダードとなっている各種情報モデルや SNOMED-CT 等のターミノロジーとの連携を視野に入れてのことである。Content Model ではジェノミクス情報の記述を取り込むことが計画されているが、既存のリソースとの連携の観点から ICD11 の可用性を高め、オミックス領域への親和性を確保するためには、ICD11 に親和性を持つ臨床オミックス情報モデルの構築が必須の課題であり、これが本研究全体の目的である。本分担研究では、前年度に構築された当該臨床オミックス情報モデルのアルファ版である iCOS- α (ICD11 Omics Sub model - α version) を元に、モデルの詳細化と検証を行った。

B. 研究方法

WHO-ICD11 の特性記述モデル (Content Model) を含む ICD11 の最新動向について WHO-ICD11 HIMTAG における議論や、厚生労働省 ICD11 国内内科 TAG 検討会に参加し、情報収集を行った。得られた結果と前年度までに

作成された iCOS α モデルを基に、ICD11 とオミックス情報との相互接続の役割を果たす部分構造 (転写モジュール) を中心に情報モデルの構築を行った。構築した iCOS モデルは Altova XML Spy を用いて XML Schema として記述した。本研究では、患者情報を扱わないため特に倫理面での配慮は必要としない。

C. 研究結果

以下いくつかの主要な部分について iCOS モデルの概要とその特徴を述べる。iCOS 全体の詳細については、別添の XML Schema を参照されたい。

C-1 最上位構造

図 1 に iCOS モデルの最上位構造を示す。iCOS は ICD11 における情報単位 (疾患カテゴリー) を表すエンティティ ICD11_entity とオミックス情報を記述するためのエンティティ OML (Omics Markup Language)、そしてそれらとの間の相互接続を行うための transcription_module というエンティティの 3 つからなっている。これにより、ICD11 における疾患カテゴリーの特性情報の記述と OML によるオミックス情報の記述を連携させる構造となっている。

C-2 ICD11_entity

図 2 に ICD11_entity の構造を示す。これは以下の 3 つの内容から構成されている。

- (1) 各カテゴリに付随する基本情報
各疾患カテゴリのタイトルである `Concept_Title`、その文章定義である `Textual_Definition`、実際にそのカテゴリを指し示す用語（各言語、あるいは種々の同義語等）である `Terms` が含まれる。
- (2) 分類体系としての階層構造情報
`Classification_Properties` は、上位階層へのポインタや、どのような目的（`Primary Care`, `Morbidity Coding`, `Mortality Coding`）において使用されるカテゴリか、といった情報を記述する項目である。
- (3) ICD11 で導入される疾患カテゴリ特性
従来の ICD で記述されていた上記(1),(2)の情報に加え、ICD11 において初めて追加される特性である。現在の時点で WHO-ICD11 HIMTAG (Health Information Modeling Topic Advisory Group) において議論されている最新の Content Model (原因・症状・発生部位・治療など、各疾患カテゴリの特性を記述するための項目セット) を元に設計されている。

C-3 Transcription_module (転写モジュール)

図3にICD11で記述された情報とOML(オミックス情報モデル)で記述された情報を連結するための `Transcription_module` (転写モジュール) を示す。

- (1) ICD11→OMLでは、ICD11のContent Model内の `Genomic_Linkages` にて記述された内容から関連するオミックス情報(OML内の `omics_data`) にアクセスすることを想定しており、その2つの項目が確信度と情報源のデータベースへの参照が付いたリンク(`linkage_with_certainty`) にて関連づけられている。
- (2) 一方のOML→ICD11については、OML内のオミックス情報から対応するICD11_entityへのマッピングとなっている。OML側の情報は、特定の `omics_data` が持つ臨床情報(`ICD_annotation`・後述)のうち、ICD11へのマッピングに必要な(A)ICD上の分類情報(`ICD_Classification_Annotation`)と、(B)ICD11のcontent modelに対応するOML内の疾患特性(`disease_NCK`)を取り出し、`OML_ICD11_Annotation_Group`としてグルーピングされている。(図4参照)これらはA)と同様に、ICD11内の特定のICD11_entityと、`linkage_with_certainty` リンクで関連づけられている。

ンクで関連づけられている。

C-4 OML (Omics Markup Language)

図5～図11にOMLの構造を示す。OMLはオミックスデータ(`omics_data`)、それに対するアノテーション情報(`omics_annotation`)、またICDに關係するような臨床関連情報(`ICD_annotation`)からなっている。

- (1) `omics_data` は図5,6に示すように `genomic sequence variation`, `transcriptomics`, `proteomics`, `signalomics`, `metabolomics`, `organomics`, `other_omics` など各種の `omics` データ(択一)を記述する項目である。例えば `genomic sequence variation` については日本発の国際標準規格である ISO25720:2009 Health informatics - Genomic Sequence Variation Markup Language (GSVML) の構造を利用して記述される。
- (2) `omics_annotation` では図7,8のように各種 `omics_data` に対するアノテーション情報を記述する。例えば `genomic sequence variation` に対するアノテーションとして図8のような情報が記述可能である。
- (3) `ICD_annotation` (図9)は各 `omics_data` と関連する高位の臨床情報を記述する項目であり、図10に示すように、`personal_info`, `phenotype`, `environmental_condition`, `clinical_annotation` 等の項目を含む。これらの情報の中で、ICD11に關係する項目は各 `omics_data` が関連する疾患へのポインタである `ICD_Classification_Annotation` と `clinical_annotation` 中の `disease_NCK` であり、これらをまとめたものが C-3(2)で述べた `OML_ICD11_Annotation_Group` である。`disease_NCK` は疾患とその臨床特性(症状・所見, 病因, 治療, 診断クライテリア, etc.) を記述した、我が国で構築されたリソースである `NCK` (Normalized Clinical Knowledge) の構造を元にしており、ISO25720 GSVMLでも採用されているモデルである。図11に各疾患の臨床特性の詳細を示す。`disease_NCK` の下位の `disease_description` を起点に辿られる太字の部分がICD11のcontent modelと対応しており、これらは内容的にほぼ互換であること、並びにICD11のバックボーンオントロジーとして予定されているSNOMED-CTにおける疾患記述モデルからもほぼ対応付

け可能(参照モデルである iCOS の方が、記述粒度が細かいため)であることが昨年度の研究にて検証されている。

D. 考察

iCOS モデルでは ICD11 における疾患情報から関連するオミックス情報を、またそのオミックス情報から関連する疾患情報を、と双方向に情報を辿ることができるように設計されており、既存のモデルでは実現されていない大きな特徴の 1 つである。それぞれの方向はいずれも 1 対多の対応となることが想定されるが、本モデルではそれを転写モジュールを介して再帰的に辿ることが可能である。

例えば、ICD11 においてある疾患カテゴリーの特性とし、ICD11_genomic_linkages において関連する omics_data が確信度と DB ソースを伴って (1 つ以上) 記述されているとする。その情報は ICD11_OML_Transcription によって OML 内部の omics_data と関連づけられる。各 omics_data は genomic_sequence_variation のようなデータから transcriptomics, proteomics, metabolomics など様々な種類が想定されるが、これらには関連するオミックスレベルでの付加情報(omics_annotation)と、もしあれば高位の臨床的な付加情報(ICD_annotation)が (1 つ以上) 存在する。これには phenotype や environmental_condition といった情報の他に関連する疾患の ICD 情報 (ICD_classification_annotation) や臨床情報(clinical_annotation) が含まれ、さらにそこから disease_description を介して複数の疾患特性 (symptoms, findings, therapy, etiology, etc.) と関連を辿ることができる。そこで記述される OML 内の疾患特性と ICD 情報の組 OML_ICD11_Annotation_Group から、また対応する ICD11_entity を辿ることができる。

また疾患情報とオミックス情報の記述において基となっているのは、それぞれ疾患分類の国際デファクトスタンダードである ICD、また日本発の国際標準規格である ISO25720 GSVML である。既存の国際標準を最大限再利用した形でモデルを作成することは国際化推進の観点から極めて重要であり、これが本研究の 2 つめの大きな特徴である。ISO25720 は元々主に SNIP を対象としたモデルであったが、本研究の iCOS では他の種類のオミックス情報も包含するため GSVML を拡張しており、これが従来のモデルとは異なる点である。

一方、今後モデルの記述能力と妥当性の検証を進めるためには、実際のゲノムデータ、階層オミックス情報、臨床情報を用いた実証研究が必要と考えられる。2015 年のリリースに向けて記述が進められる ICD11 は現時点では疾患特性記述作業が余り進んでおらず、記述粒度など未だ不明な点も多い。そのため、今後 ICD11 構築作業の動向を踏まえながら、iCOS モデルの調整並びにユースケースごとの検証を行う必要があると考えられる。

E. 結論

本分担研究では、昨年度の研究成果を発展させ、ICD11 に親和性の高い臨床オミックス情報モデル(iCOS)について、特に ICD11 の疾患特性情報とオミックス情報の相互連結機構を中心に構築した。これにより、ICD11 で記述される疾患の特性から、関連するオミックス情報を辿り、またそのオミックス情報から関連する疾患を辿る、という双方向の記述が可能となった。今後さらなる検証のためには、実データを用いた記述能力と妥当性に関する実証研究が必要と考えられる。

F. 健康危険情報

なし

G. 研究発表

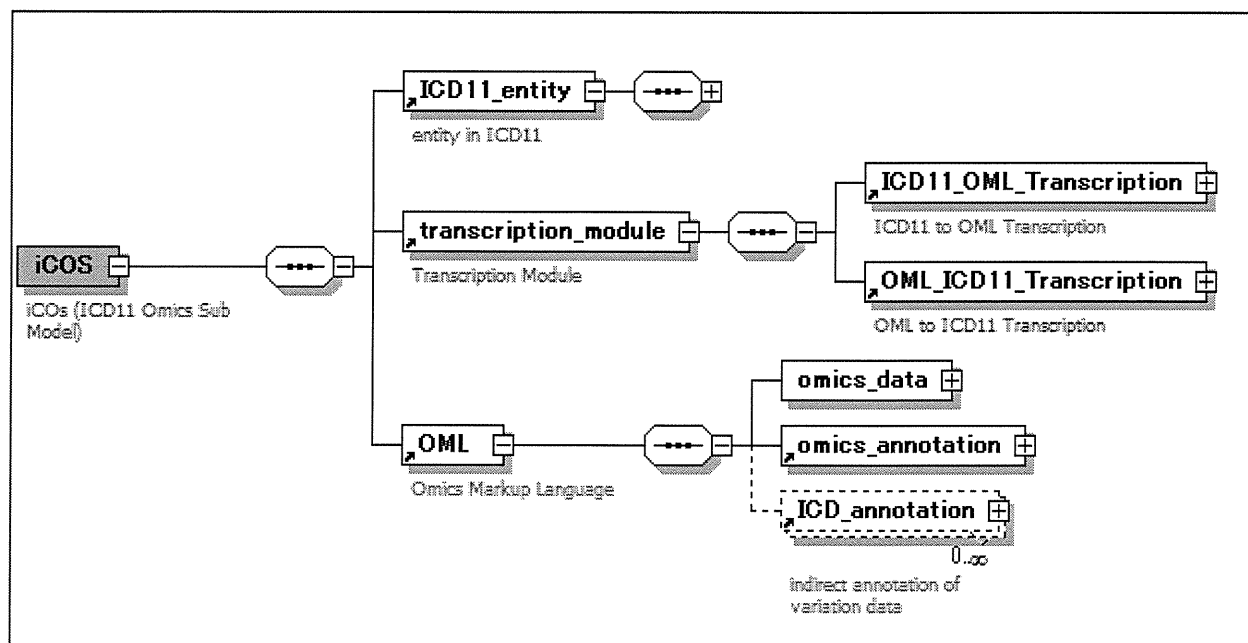
1. 論文発表

今井健, 中谷純. ICD11 におけるオミックス情報モデルの基礎検討. 医療情報学 31 (Suppl.), pp.81-86, 2011.

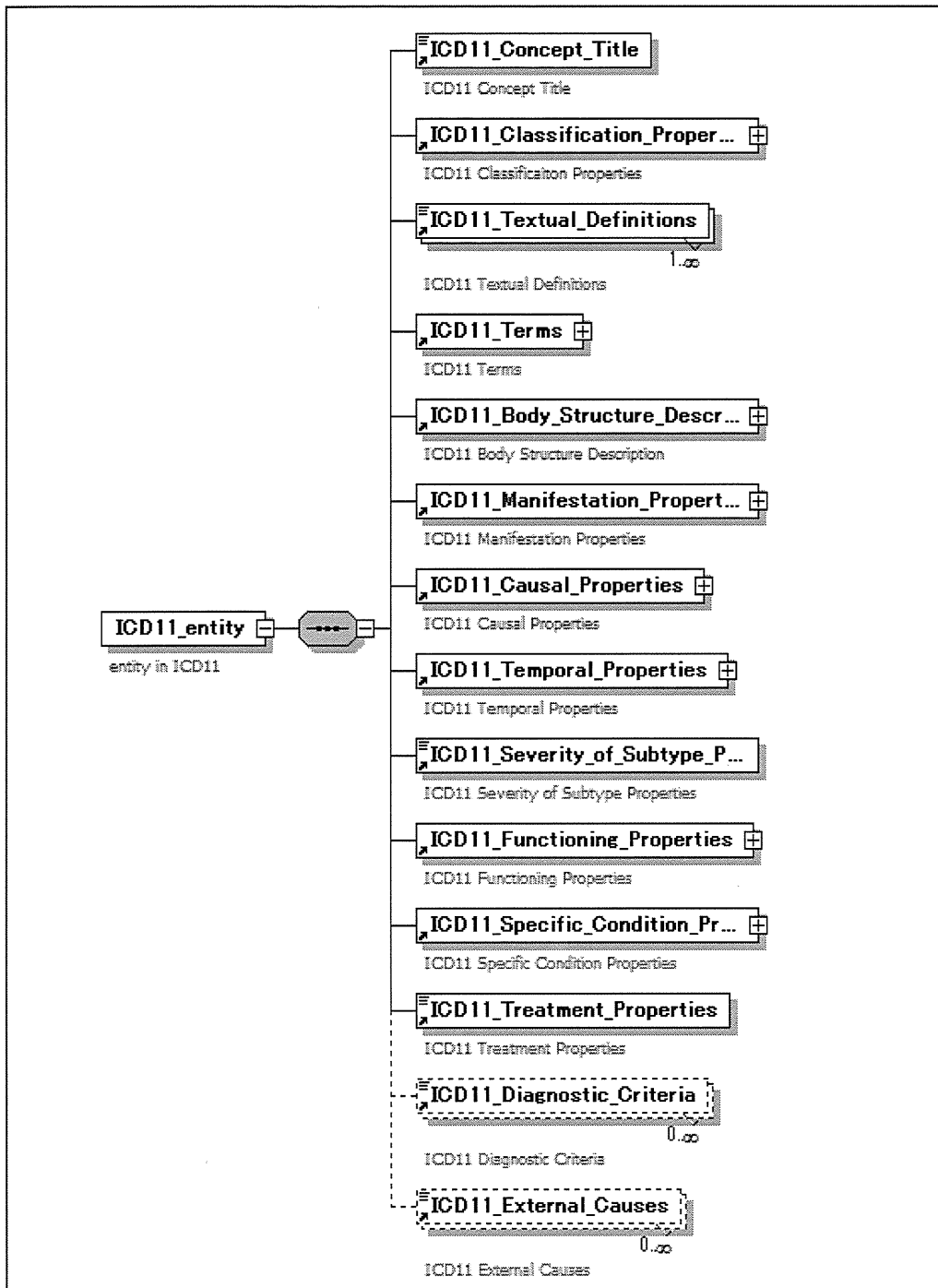
2. 学会発表

H. 知的財産権の出願・登録状況

なし



(図1. iCOS モデルの最上位構造)



(図2. ICD11_entity の概要 (第1階層まで展開))