

201002007A

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）

ICD11におけるオミックス情報モデルの研究

平成22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23（2011）年 5月

目 次

I. 総括研究報告	
ICD11におけるオミックス情報モデルの研究	-----1
中谷 純	
II. 分担研究報告書	
1. 国際情報収集に基づく ICD11 におけるオミックス情報モデルの基礎検討	
今井 健	----- 21
III. 研究成果の刊行に関する一覧表	----- 40

平成 22 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、オントロジー的整理を導入し、多軸構造、多言語対応、統一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しうるオミックス情報モデルを作成する。

平成 22 年度は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を進め、WHO-TAG-HIM での情報交換収集などを行いながら、ICD11 コンテンツモデル、ISO 25720 GSVML, GO (Gene Ontology) [3], NCK (Normalized Clinical Knowledge) [4]などを参考にして、アルファ版を作成した。また、SNOMED-CT [5]との対応解析も行った。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、日本発進の新技术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
広井 嘉栄 東京医科歯科大学 特任助教

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

平成 22 年度は、以下のような調査研究を行った。

1 オミックス医学における用語間の関連性についての検討

用いるオミックス医学分類はまだ存在しないが、ISO 25720 GSVML, GO (Gene Ontology), NCK (Normalized Clinical Knowledge)などを参考にして、これらのモデルで用いられている用語を基に、用語同士の関連性を解析する。関連性検討に際して、それぞれの用語の意味を多方面から分解し多軸的な意味情報（メタデータ）として記述し、共通の軸を抽出する。抽出した共通軸を中心として、非共通軸を重複の無い形で付加し、すべてのメタデータを記述しうる構造を素描し、アルファバージョンとする。

2 アルファ版の素描

用語の関連性解析検討において抽出した共通軸を中心として、非共通軸を重複の無い形で付加し、関連性解析において作成したすべてのメタデータを記述しうる構造を素描し、アルファバージョンとする。

3 WHO-TAG-HIM 会議への参加ならびに情報交換収集

WHO-TAG-HIM (Topic Advisory Group for Health Informatics and Modeling)では、ネットワークミーティングが随時行われるので、その会に出席し、ICD のモデル化作業に関する情報収集と意見交換を行う。必要に応じ、ジュネーブで行われる Face To Face Meeting に参加する。また、随時、WHO-FIC, 内科 TAG などへも参加し、情報交換収集を行う。

（倫理面への配慮）今回使用する情報は、臨床オミックス医学の情報構造ならびに用語であり、氏名・年齢等個人を特定できるもの、個人情報に関する内容を含まない。

C. 研究結果

・ オミックス医学における用語間の関連性についての検討を行った。ISO 25720 GSVML, GO (Gene Ontology), NCK (Normalized Clinical Knowledge)などを参考にして、これらのモデル

で用いられている用語を基に、用語同士の関連性を解析した。まず、関連性検討に際して、それぞれの用語の意味を多方面から分解し多軸的な意味情報（メタデータ）として記述し、共通の軸を抽出した。抽出した共通軸を中心として、非共通軸を重複の無い形で付加し、すべてのメタデータを記述しうる構造を素描した。

・ 解析に基づいて、アルファ版の素描を行った。用語の関連性解析検討において抽出した共通軸を中心として、非共通軸を重複の無い形で付加し、やはり関連性解析において作成したすべてのメタデータを記述しうる構造を構築し、アルファバージョンとした。

・ WHO-TAG-HIM 会議への参加ならびに情報交換収集を行った。WHO-TAG-HIM (Topic Advisory Group for Health Informatics and Modeling) の随時行われるネットワークミーティングに出席し、ICD のモデル化作業に関する情報収集と意見交換を行った。今回は、ジュネーブで行われる全体会議 (Face To Face Meeting) は行われなかった。なお、随時、WHO-FIC、内科 TAG などへも参加し、情報交換収集を行った。

1. 現在の ICD11 Content Model におけるジェノミクス情報の取り扱い

本年度はこれまで行われてきた WHO-TAG HIM の対面会議は実施されなかった。そこで数回行われた電話会議による議論に参加し、Content Model 並びに記述フレームワークの調査を行った。

別添表 1 に WHO-TAG HIM で議論が行われた最新の ICD11 Content Model と原則案を示す。各疾患カテゴリーの特性は 1~14 の主項目により記述され、いくつかの項目は副項目を持つ階層モデルとなっている。1~4 は分類体系において基本的情報である項目名や上下階層関係、定義等の記述であるが、5~14 は各疾患カテゴリーの臨床的特性を記述するための項目となっている。これらの記述項目は全ての疾患においてその十分性が網羅的にチェックされたものではない。そこで ICD10 カテゴリーの疾患特性を、「他のカテゴリーと区別するのに必要最小限な特性セット」として網羅的に記述した試みである既存研究（科学研究費補助金 若手研究 (B) 19700128 「自由入力病名から国際疾病分類体系へのオントロジーに基づく自動分類手法に関する研究」研究代表者：今井 健）の成果を元に、両者の特性記述項目の比較を行った。

結果、既存の ICD10 カテゴリーを記述し直すことを考えた場合、現在の Content Model は「ある疾患の続発症/後遺症/合併症となっている」という意味関係については不十分ではあるものの、その他についてはほぼ項目間の対応が存在することを確認した（論文発表 (A) 参照）。若干の微修正は必要なものの Content Model はほぼ十分な項目カテゴリーを備えていると言える。一方で項目設定の範囲としてはほぼ十分だとしても、記述粒度によっては他のリソースとの

連携という目的に合致しないことも予想される。ジェノミクス情報は 7.4 の ” Genomic Linkages ” という項目で記述されることとなっているが、現時点では表にある以上の詳細な記述フレームワークは定められておらず、GO (Gene Ontology) など既存のリソースとの連携という観点からは不十分と考えられる。

2. ISO 25720 GSVML における疾患特性記述モデルと ICD11 content model

ISO 25720 GSVML (Genomic Sequence Variation Markup Language, ISO25720) は主に SNP に焦点を当て、Genomic Sequence Variation のデータを交換するための XML ベースのマークアップ言語であり、2009 年に ISO 標準となっている。

各 Variation のデータは別添図 1 に示すように、直接的な付加情報である Direct_annotation (allele, type, position, length, region 等) とより高位の説明情報である Indirect_annotation (omics data, clinical information, environmental data) から成っている (下位に大量の情報項目が存在するが一部のみ抜粋した)。ICD11 Content model の “Genomic Linkages” は前述の通り未だ記述フレームワークが存在していないが、これらの構造を利用可能であると考えられる。

一方、1 つのジェノミクス情報は複数の疾患に関連している場合もあるため、「ジェノミクス情報→関連する疾患」という方向の記述も整合性を持って行われなければ成らない。この関連疾患特性の記述は、ISO 25720 GSVML では Indirect_annotation 中の Clinical_annotation に含まれる Disease_description を中心に行われる。別添図 2, 図 3 にそれぞれ Clinical_annotation, Disease_description の構造を示す。Disease_description の構成要素は他の Clinical_description の要素を参照する形で記述されている。ICD11 content model には存在していない臨床特性記述項目も数多く存在している。また Symptoms や Laboratory/Pathological findings など多くの項目はまたさらに下位の項目を含み、ICD11 content model の臨床的疾患特性 (項目 5~14) よりも記述粒度が細かいが、対応づけに必要な程度の粒度で比較すると別添表 2 のようになる。ISO 25720 GSVML の特徴上、本来対象としていない外傷の外因 (External Causes) を除けば、基本的には ICD11 content model の内容は網羅されている。ただし項目によって 1:1 対応になっていない箇所も存在し、今後臨床オミクス情報モデルとして発展させていくためには Etiology_description, Diagnostic_differential, Symtoms_description などの項目を詳細化し ICD11 content model とマッピングするための Reference Model を作成する必要があると考え

られた。一方で、研究結果 1 で述べたように ICD10 を書き直す際に ICD11 content model では足りていないと思われる特性として、後遺症・続発症・合併症などの関係があるが、ISO 25720 GSVML では Disease_description > Complication として項目がすでに存在しており、今後追加されたとしても対応が可能であることが判明した。

3. ISO 25720 GSVML と SNOMED-CT における疾患特性記述モデル

SNOMED-CT は 1999 年米国病理学会が定めた病理疾病分類 SNOMED をベースとし、英国病院用の医療用語集 United Kingdom's Clinical Terms Version3 (Read Code)、MeSH (米国国立医学図書館科学用語集) などを統合して作成され、2007 年からは The IHTSDO (International Health Terminology Standards Developing Organization) により管理されているターミノロジーである。また、約 39 万の各々の概念の間には Is-A, Associated_morphology, Causative_agent, Finding_site, Has_specimen, Severity など 65 種類の意味関係が付与され、その意味関係リンクの総数は約 142 万に及ぶ。

ICD11 content model においても現在のところ、一部 (Functioning Properties における ICF の利用等) を除く多くの主要な項目において SNOMED-CT を用いた記述あるいは SNOMED-CT エンティティへのリンク作成が予定されている。SNOMED-CT 利用についての詳細は未だ WHO-TAG HIM 内でも議論中であるが、ICD11 において SNOMED-CT へのリンクが重要である 1 つの理由は ICD11 改訂の当初の構想において SNOMED-CT における意味関係リンクをオントロジーとしてのバックボーンとして利用しようとしたことにある。例として、別添図 4 に SNOMED-CT における “Acute glomerulonephritis (急性糸球体腎炎)” の意味関係を示す。この場合、下位概念へのリンク以外、つまり 3 つの Is-A ならびに Clinical_course, Associated_morphology, Finding_site で表現される疾患特性のセットが SNOMED-CT における本疾患の定義であり、“Defining Relationships” と呼ばれる。SNOMED-CT 内ではこれら “Defining Relationships” に基づいて、疾患同士の上下関係が無矛盾に定義されていることを “Description Logic (記述論理)” の計算により自動的にチェックしている。これらの意味関係は定義上必要最低限の意味関係を付与するという方針であり、ある疾患に関する教科書的な臨床知識全般を記述しているわけではない。例えば疾患が呈する可能性がある症状や疫学的知識、観測される検査値異常などの所見を網羅的に記述していない。しかし少なくとも定義された範囲内で、オントロジーとしての概念間の上下 (Is-A) 関係の整合性を担保することを目的としている。一方で ICD11

では、SNOMED-CT ではカバーされていない広範囲な臨床上的特性を記述しようとしており、そこにおける分類体系としての疾患カテゴリの上下階層関係は、必ずしもオントロジーとしての上下関係ではない。過剰に深い階層となることを防ぐために無理に複数の疾患概念をまとめて 1 つのカテゴリとしていることもあり得る。つまり、ICD11 は SNOMED-CT における疾患特性記述とは異なる目的とカバー範囲を持ったもので、ICD11 の上下階層関係はオントロジーとしての厳密な意味での上下関係ではないが、オントロジカルな整合性を担保するバックボーンとして SNOMED-CT を利用する、ということが当初の構想には含まれていた。これがどの程度実現されるが未だ不明であるが、少なくとも ICD11 content model と SNOMED-CT での疾患特性記述モデルに整合性がなければならず、それは ICD11 と親和性を持った臨床オミックス情報モデルの構築においても同様に重要である。

別添表 3 に SNOMED-CT の疾患特性記述モデルを、また表 4 に ISO 25720 GSVML Disease_description 部分から見た SNOMED-CT の疾患特性記述モデルとの対応を示す。(ISO 25720 GSVML の Clinical_annotation 部分には Disease_description には含まれない Family_history 等の患者臨床情報記述に必要な項目が含まれているが、今回は ICD11 を中心とした整合性に注目しているため、そのような個人・個別事例に特化した記述項目は検証から除外した)

SNOMED-CT の疾患特性記述モデルは基本的に上位概念を示す “Is-A” と、表 3 における 13 個の特性項目から成る。Associated_with は「患者内での単純な共起を超えた 臨床上的’関連’を表現するが、causal もしくは sequential な関係を主張するものでもどちらかを除外するものでもない」という意味であり、もっと意味が特化された以下の 3 つを包含しており、それら詳細区分も含めると 16 個の特性項目が存在する。

- Associated_with::After

ある臨床上的所見が他の臨床上的所見や手技の後に起こることを表現。causal relationship を主張するものでも除外するものでもない。単純に sequence of events を強調するもの

- Associated_with::DueTo

臨床上的所見を直接 ‘原因’ と関連づけるもの。もしある所見が、直接的に単に他の疾患の「傾向を与える」or「悪くさせる」のであれば、もっと一般的な ASSOCIATED WITH が使われる

- Associated_with::CausativeAgent

疾患の直接的な causative agent。媒介生物は包含しない

また表 3 で Episodicity は過去のバージョンとの互換性、Post Coordination の目的のために残されており、Pre-coordinated definition についてはもう使用されていない。表 4 の対応表において★印は「直接的に特性が対応している

わけでは無いが、他の方法で表現されているもの」を示す。また“(inverse) Is-A”は Is-A の逆関係、すなわち下位概念への参照を使って表現されていることを示す。

表 4 の対応分析結果から、SNOMED-CT での疾患特性記述モデルでは、ISO 25720 GSVML の Disease_description と対応しない項目がかなり多いことが判明した。また SNOMED-CT においては疾患定義に必要最小限のもののみを記述するという方針のため、これら全ての特性項目が使用されているわけではない。従って実際の疾患特性の記述内容では、ISO 25720 GSVML の Disease_description 内の項目と対応がつくものはさらに少なくなると考えられる。しかし、本研究で構築する臨床オミックス情報モデルは、複数のリソースと親和性を持つ中間参照情報モデルであるため、ISO 25720 GSVML 側の記述項目の方が多い場合については、問題にならないと考えられる。

4. ICD11 対応オミックスサブ情報モデル(α版)

別添図 5 に素描した ICD11 オミックスサブ情報モデルの α 版 (ICD11 Omics Sub Information Model α 版: iCOS α) を示す。エントリーポイントは、オミックスデータカテゴリーである。オミックスデータは内部連携を通じて ICD Annotation に連携される。ICD Annotation のサブカテゴリーである Disease Description は、ICD11 のメインパラメーターとほぼ同等の項目配置となっており、一対一対応が可能である。この Disease Description は、また、ISO 25720 GSVML の Indirect Annotation 中にある Clinical Annotation の中の Disease description と等価であり、NCK の Archetype と完全な互換性がある。従って、近年統合データベースなどで重要性が指摘されているデータ知識リンクが可能である。この Disease description の下位構造に Classification Annotation を設定した。この Classification Annotation を通じて、ICD Category, ICD Entity あるいは SNOMED Code に対応することができる。Omics Annotation カテゴリーは、オミックスデータをアノテートするために必要な情報群が収納される。これは、実際にオミックスデータの質、データ取得背景などを評価するために必須の項目である。

別添図 6 に、ICD11 コンテンツモデルと iCOS α との情報連携関係を示す。ICD11 コンテンツモデルの中の Causal Properties の中の Genomics Linkages に格納されたオミックス情報は、iCOS α サブモデルの Omics Data にリンクする。Omics Data のサブ構造は、そのオミックスの種類や特性を必要に応じて記述できるようになっている。

Omics Data からの情報は、iCOS α 内の内部リンクを通じて ICD Annotation に連結され、その下位構造となる Disease Description に連結される。これらの連結は、確率統計表現を介して連結定義することも可能である。この Disease

Description の細項目は、ICD 11 メインパラメーターの大部分に一対一対応する。従って、重複を避けるため、iCOS α 側から ICD 11 コンテンツモデルの等価項目を指し示す必要があり、サブモデルから親モデルをポイントする、いわば、逆転写型情報モデルという構造であることが必要となる。

さらに、この Disease Description は、Classification Annotation に一対一で連結する。Classification Annotation には、ICD Entity title, Classification properties, あるいは ICD Code を格納する。従って、Classification Annotation は、ICD11 Contents Model における ICD Entity title, Classification Properties とほぼ等価な内容を包含することとなり、iCOS α 側から ICD 11 Contents Model 側をポイントし、親モデルに逆転写するような構造が必要となる。その他の付加情報についても、国際標準的な範囲の内容は、Omics Annotation、ICD Annotation に記述することができる。

D. 考察

現在の ICD11 の Content model は “Genomic linkage” の特性項目の詳細が決まっておらず、GO (Gene Ontology) や他のリソースと連携をする上では記述粒度の点で不十分である。ISO 25720 GSVML のジェノミクス情報記述構造の活用は、ICD11 における疾患特性とジェノミクス情報との親和性を高めるために非常に有用だと考えられる。また疾患とジェノミクス情報の再帰的な相互参照のためには、ISO 25720 GSVML の疾患特性記述モデルと ICD11 content model との整合性が重要であるが、若干の修正が必要なものの、ほぼ content model と対応付け可能であった。ただし一部 1:1 に対応していない項目が存在しているため、Etiology_description, Diagnostic_differential, Symptoms_description など詳細化した Reference model を構築する必要があり、今後の課題である。

一方、ISO 25720 GSVML の Disease_description と SNOMED-CT の疾患特性記述モデルとの対応については、SNOMED-CT 側に対応しない項目が多いことが判明した。しかし、仮に ISO 25720 GSVML を基にして構築される臨床オミックス情報モデルで疾患特性を記述した場合、SNOMED-CT 内に対応項目が少ないことはあまり問題にならない。なぜなら SNOMED-CT を ICD11 のオントロジカルなバックボーンとして活用する場合、ICD11 で記述される予定の臨床的特性を全て用いてオントロジカルな整合性のチェックをする訳ではなく、SNOMED-CT 内で定義されている必要最小限の特性のみを用いて行えば十分だからである。しかし、もしより正確にマッピングを行うのであれば、SNOMED-CT の Permissible Value の活用が考えられる。例えば SNOMED-CT 側で非常に広範囲な意味関係をカバーする特性として用いられている

“Associated_with” は ISO 25720 GSVML の Disease_description 側では複数の箇所に該当する可能性があるが、SNOMED-CT 内では Associated_with 特性のとり値として許されている概念カテゴリの種類が限定されている (表 3 参照)。従って、「Associated_with + 対象概念カテゴリ」の組み合わせを詳細に指定して区別することで、より正確なマッピングが可能となると考えられる。

一方、SNOMED-CT 内で記述されている特性が臨床オミックス情報モデルに存在しない場合は不整合を引き起こす問題となりえる。すなわち臨床オミックス情報モデル側の記述粒度が粗く SNOMED-CT 内の特性に正確にマッピングできないと、SNOMED-CT 内でのオントロジカルな整合性計算で支障を生じる可能性がある。ターミノロジーである SNOMED-CT では、用語をベースとするため、ISO 25720 GSVML よりも記述粒度が細かい事例が見られた (例えば、検査所見を「<Interprets> “Muscle tone” + <Has_interpretation> “Decreased”」で表現)。このような事例については ISO 25720 GSVML の疾患特性記述モデルを部分的に詳細化する必要がある、今後の課題である。しかしその際においても、上と同様に「<特性> + “Permissible Value”」の組み合わせ情報を活用することで効率的にマッピング規則が作成できると考えられる。

現在、Causal Properties - Genomics Linkages に設定されているオミックス情報は、病因に関する情報の一つとしての利用が想定されている。しかし、一般的には、病気とオミックス要因が一对一对応になることはむしろまれで、通常、多対多の対応となることが多い。つまり、ある病気からオミックス情報をポイントしたときに、そのオミックス情報は同時に他の病気における原因でもありうる。従って、ICD11 において、オミックス情報を表現可能とするためには、多対多の対応を許容し、それを表現可能としつつ、論理一貫性を失うことのない情報モデルが必要である。このためには、ICD11 コンテンツモデルと iCOS α との間でエントリーポイントをオミックス情報の一点だけにはっきりと限定すること、そのオミックス情報に対応する臨床情報は、逆に一对多の対応となる場合が多く、それを許容した上で論理破綻の無い構造にすることが必要である。今回素描した逆転写型情報モデルは、ICD11 パラメータの内部連携をより現実性のある形で可能にし、高度化する情報モデルとして意義付けることができる。

ICD11 のミッションとして、国際標準であることと電子化対応できることが言われている。これを迅速に実現するためには、すでに国際標準電子交換技術となっている ISO 25720 GSVML と互換性を持つサブ構造を設計する手法は、簡便でかつ意義が大きい。今回 ISO 25720 GSVML の内部構造を応用して別体のサブモデルとして素描し、その連携方式を逆転写型とした情報

構造は、ICD11 をより現実的な形で国際標準としうるサブモデル構造という事ができる。

E. 結論

多方面からの解析を行った結果、まず、ISO 25720 GSVML におけるジェノミクス情報記述モデルは、現在詳細が決まっていない ICD11 content model における同記述項目に当てはめる記述モデルとして原理的に応用可能であることが判明した。今後 1 対 1 対応していない項目のマッピングのために Reference Model を作成することにより、ICD11 content model と親和性の高い臨床オミックス情報モデルへと発展させることが可能と考えられる。特に ICD11 content model と臨床オミックス情報モデルにおける疾患特性記述部分との対応をとることは、「疾患 \leftrightarrow 関連ジェノミクス情報」の双方向探索 (ある疾患が関連するジェノミクス情報、さらにそのジェノミクス情報に関連する複数の疾患...) の整合性を保つために非常に重要なポイントである。

一方、ICD11 のオントロジカルバックボーンとして利用が見込まれている SNOMED-CT との対応については、ISO 25720 GSVML の疾患特性記述部分 (Disease_description) の方が豊富な項目を持っている部分については不整合の問題を引き起こさないが、SNOMED-CT の方のみ存在する項目については不整合の原因となり得る。このようなケースについては、今後 SNOMED-CT の「疾患特性と Permissible Value の組み合わせ」を用いて、マッピング規則を作成していく必要があると考えられる。

多方面からの解析結果に基づいて、逆転写型情報モデル α 版 (iCOS α) を素描した。iCOS α の内部構造を、ISO 25720 GSVML を応用した別体のサブモデルとして素描し、その連携方式を逆転写型とした情報構造は、ICD11 をより現実的な形で国際標準としうるサブモデル構造である上に、ICD11 パラメータの内部連携を高度化しより現実性のある表現を可能にする付加情報モデルとして意義付けることができる。

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 ICD 疾病分類に組み込むための要に位置づけられる研究であり、研究、臨床を含めた医学全体の発展に貢献することができる。臨床オミックス情報モデルの提示は、ポストゲノム時代に入って常識となったオミックス情報の医学研究応用をデータ交換の効率性を向上するといった観点から加速することができる。また、この情報モデルが ICD11 への親和性を持つことにより、国際的な標準を担保した形での効率化が可能である。このことは、医学、医療におけるさまざまな局面での効率の安定的向上を促し、コスト低減、時間短縮、効率向上などの観点から、将来にわたって大域的に国民の保健・医療・福祉の向上の一助となることができる。また、ICD11 に対応したオミッ

クス医学情報モデルを、日本が主導的に開発検討提案することは、今後大きく発展を遂げると思われるオミックス医学の発展に日本が大きく寄与し国際的にも主導的な役割を果たすことを意味する。長い ICD の歴史の中で、電子化への転換の節目となる ICD11 において、日本がこのような形で主導的な役割を果たすことは、日本が世界に貢献する歴史的な成果となる可能性がある。

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) 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.). 79(2). 130-142. PMID: 19969503.
- B) Samson W. Tu, Olivier Bodenreider, Can Çelik, Christopher G. Chute, Sam Heard, Robert Jakob, Guoquian Jiang, Sukil Kim, Eric Miller, Mark M. Musen, Jun Nakaya, Jon Patrick, Alan Rector, Guillermo Reynoso, Jean Marie Rodrigues, Harold Solbrig, Kent A Spackman, Tania Tudorache, Stefanie Weber, Tevfik Bedirhan Üstün. (2011). A Content Model for the ICD-11 Revision. Journal of American Medical Informatics Association.

(In Submission).

- C) Michio Kimura, MD PhD, Jun Nakaya, MD PhD, Kazuhiko Ohe, MD PhD, Ryuichi Yamamoto, MD PhD, Mihoko Okada, PhD, Naoki Nakashima, MD PhD, Hideshige Hasegawa, Kazuyuki Nakayasu, Hiroshi Watanabe, MD, Shigeki Tani, PhD, Hiroshi Tanaka, PhD. 2011 "A survey aimed at general citizens about their attitudes toward electronic medical data handling", Journal of American Medical Informatics Association. In submission
- D) Jun Nakaya, Michio Kimura, Riichiro Mizoguchi, Kouji Kozaki, Kaei Hiroi, Keisuke Ido, and Hiroshi Tanaka. 2011. Semantics in the Japanese Integrated BioMedical Database. Journal of American Medical Informatics Association. In submission
- E) Shigeki Tani PhD1, Shinichi Higuchi1, Goshi Fujimoto1, Jun Nakaya MD, PhD2, Michio Kimura MD, PhD. 2011 More powerful search engine invalidates anonymity guidelines for case reports. Methods of Information in Medicine (Methods Inf Med.). In submission
- F) Imai T, Kajino M, Sato M, Ohe K. Development of structured ICD-10 and its application to computer-assisted ICD coding. Studies in Health Technology and Informatics. 2010;160(Pt 2):pp.1080-4.
- G) Yamada E, Aramaki E, Imai T, Ohe K. Internal structure of a disease name and its application for ICD coding. Studies in Health Technology and Informatics. 2010;160(Pt 2):pp.1010-4.
- H) 藤田伸輔, 今井 健. SNOMED-CT と ICD-11 に見る医学・医療分野の Ready to Use Ontology. 人工知能学会誌 2010;25(4):pp.501-508.
- I) 中谷 純. 2010. 温泉の百科事典、ゲノム・ポストゲノム. 丸善. in Press

2. 学会発表

招待講演

- A) 医学用語について. 中谷 純. 平成 22 年度日本温泉気候物理医学会専門医講習会. 2010. 8
- B) 用語標準化の最前線 (ISO TC215 WG3、WHO ICD11 の現状と未来). 中谷 純. 平成 22 年度医療情報学会シンポジウム. 2010
- D) レセプトナショナルデータベース活用に向けたレセプト情報モデルの検討. 広井嘉栄、中谷純、井戸敬介、岡本悦司、田中博. 2010 医療情報学連合大会. 浜松、2010

H. 知的財産権の出願・登録状況

1. 特許取得 なし

以上

(別添図表あり)

(別添図表)

Content Model Parameter	Rationale
1. ICD Concept Title	1. Needed for human understandable title label for any concept 2. Useful for translators 3. Possible use of Fully Specified Name (if agreed 4.4)
1.1. Fully Specified Name	1.To indicate the most concise title name that will completely represent the entity as a stand-alone term 2. This will be useful to link to SNOMED FSN Set
2. Classification Properties	
2.1. Parents	1. to generate the linearizations: Hierarchical relations will be built accordingly 2. Useful for Navigational purposes 3. Comparison with other ontologies /hierarchies: e.g. SNOMED CT
2.1.0 Current ICD-10 code <i>(read-only field for unchanged categories)</i>	1. Linkage to existing ICD-10 2. Legacy linearization prototype 3. Helpful Identify the new ICD entries with a new code
2.2 Type	1. To identify the nature of the entity – the type of the classification unit <ul style="list-style-type: none"> • Disease • Disorder/Syndrome • Sign/ Symptom • External Cause • Injury • Reason for Encounter 2. To precisely identify the shared characteristics in terms of definitions. This will have implication for indicating the chapter as well as naming convention

2.3. Use and Linearization(s)	<p>1. To identify what level of granularity of linearization will be applied with inclusion, exclusion and residual codes.</p> <ul style="list-style-type: none"> - <u>Clinical</u> <ul style="list-style-type: none"> - primary care (simple categorizations – fewer number) - specialty adaptations (referring to common medical services) - <u>Research</u> (more detailed equivalent to foundation component) <p>2. To align with use cases to identify the purpose of the linearization: e.g. Mortality 1- main tabular list; 3 alternative views, Morbidity 1 common list, alternatives (e.g. DRG, Admission, Discharge, Event)</p> <p>3. To allow also the Specialty Adaptations with implied inclusion, exclusion and residual codes.</p>
2.3.1. Linearization Parent	
3. Textual Definition(s)	<p>1. Key description and guidance on what the meaning of the entity/category is, in natural language</p> <p>2. Useful for translation</p> <ul style="list-style-type: none"> a. Concise definition for printing b. Detailed definition for online viewing
4. Terms	
4.1. Base Index Terms	To indicate the index entries which enable coders to search for the right code These include the 4.1.1 Synonyms and 4.1.2 children.
4.1.1. Synonyms	To indicate similar terms that are commonly used for the same entity – to enable coders and translators to specify the term
4.1.2. Narrower Terms <i>(to become → Children in the future)</i>	To indicate the terms which are different from the synonyms. In the final instance, these should be children.
4.2. Inclusion Terms	To understand conceptual space of the entity through a subset of index terms which provide convenient examples
4.3. Exclusions	To understand the boundaries of the conceptual space of the entity through convenient examples of other entities. Exclusion 1: elsewhere classified

	<p>Exclusion 2: residual category</p> <p>Exclusion 3: not included here</p>
5. Body Structure Description	
5.1. Body System(s)	To identify the physiological grouping of the entities (<i>which physiological system does this entity belong to?</i>)
5.2. Body Part(s) [Anatomical Site(s)]	To identify the anatomic grouping of the entities (<i>which anatomic grouping set does this entity belong to?</i>)
5.3. Histopathology	To identify the cellular type or morphological appearance of the entity (<i>usually used for tumors, skin lesions, etc</i>)
6. Manifestation Properties	
6.1. Signs & Symptoms	<ol style="list-style-type: none"> 1. To enable the functional description of the disease or disorders as they are seen constellation of certain sets of signs and symptoms. 2. This will be useful to link to SNOMED and other terminologies through codes. 3. This will be useful to express Diagnostic Criteria (<i>13 below</i>) digitally, as necessary.
6.2. Investigation findings	To enlist laboratory, imaging and other type of established diagnostic interventions
7. Causal Properties	
7.1. Etiology Type	<p>To indicate the basic grouping causal factors underpinning the entity to group</p> <ul style="list-style-type: none"> - Infectious - Nutritional - Physical forces - Chemical substances - Neoplastic - Vascular - ... <p><i>{ This does not apply to Chapter 20 External Causes}</i></p>
7.2. CP - Agents	<p>To enlist the detailed agents: microbiological, physical, chemical, etc. to enable different sorting of the causal agents</p> <p><i>{ This does not apply to Chapter 20 External Causes}</i></p>
7.3. CP- Causal Mechanisms	<p>To enlist the mechanism by which an agent causes disorder or disease</p> <p><i>{ This does not apply to Chapter 20 External Causes}</i></p>
7.4. Genomic Linkages	<ol style="list-style-type: none"> 1. To enlist the genes (and genetic mechanisms) to see whether a specific class should be assigned to a specified term 2. To seek similarities in grouping similar disease/disorders

7.5. Risk Factors	To enlist factors that have increased associations with the occurrence of a condition. Their association may be grouped to concur with WHO Comparative Risk Assessment framework.
8. Temporal Properties	
8.1. Age of Occurrence & Occurrence Frequency	To identify the Pediatric, Adult or Geriatric Specialty Adaptation
8.2. Development Course/Stage	{ when a FSN uses a temporal qualifier} To bring operational definitions to loosely used terms as temporal qualifiers (acute, subacute, chronic; immediate-onset, late-onset etc)
9. Severity of Subtypes Properties	
Option 1: No Severity sub-classification	DEFAULT
Option 2: If the classification entity employs terms MILD, MODERATE, SEVERE for the disease name.	{ when a FSN uses a severity qualifier} To bring operational definitions to loosely used terms as severity qualifiers (i.e.: mild, moderate, severe)
Option 3: Custom scale – if the category name, employs the severity or staging terms.	To capture existing severity scales that are commonly used with the ICD entities/categories (e.g. TNM for Neoplasms).
10. Functioning Properties	
10.1. Impact on Activities and Participation	1.To identify functional properties associated with the ICD entities 2.To differentiate from severity criteria, when necessary
10.1.1. Understanding	Ditto
10.1.2. Communication	Ditto
10.1.3. Mobility	Ditto
10.1.4. Self-care	Ditto
10.1.5. Interpersonal Relations	Ditto
10.1.6. Life Activities (Household, School, Work & Economic Life, Life Management)	Ditto
10.1.7. Social Participation	Ditto
10.1.8. Children and Youth	Ditto
10.2. Contextual factors	To link mainly Chapter Z codes to ICF contextual factors, as necessary
10.3. Body functions	To link signs – symptoms and other features with ICF body functions as necessary

11. Specific Condition Properties	
11.1 Gender	To identify gender related categories for validity checking of codes (e.g. women cannot have prostate cancer, men cannot be pregnant...)
11.2. Life-Cycle Properties	To identify certain ICD categories related to pregnancy, delivery, abortion etc.
12. Treatment Properties	<ol style="list-style-type: none"> 1. To identify treatment related diagnostic labeling (e.g. Insulin-resistant diabetes) 2. To identify linkage to WHO treatment guidelines, ICHI, ICTM etc (e.g. Hypertension) 3. To identify use in Primary Care Version to link Diagnostic Modules with Management Modules (e.g. PC version, Map of Medicine) 4. To complete the Reasons for Encounter entities
13. Diagnostic Criteria	<ol style="list-style-type: none"> 1. To enable the compilation of the core diagnostic information necessary and sufficient to describe a category 2. To enable the digital representation of the diagnostic algorithms / aggregation logic using SNOMED- CT and other elements as appropriate
14. External Causes	<p>To enable the coding of the External Causes and enable identification distinct components such as:</p> <ul style="list-style-type: none"> EC- Agents (object - substance) EC- Mechanism EC- Intent EC - Place EC- Activity EC - Substance use

(表 1 Content Model with Rationale – WHO-TAG HIM 2010 年 8 月 27 日版より)

ICD11 Content model	GSVML
5. Body Structure Description	Striking_body_area(Disease_epidemiology 内)
6. Manifestation Properties	
6.1 Signs & Symptoms	Symptoms (詳細項目有り)
6.2 Investigation findings	Laboratory_findings, (詳細項目有り) Pathological_findings (詳細項目有り)
7. Causal Properties	
7.1 Etiology Type	Etiology_description
7.2 CP-Agents	Etiology_description
7.3 CP-Causal Mechanisms	Etiology_description. Etiology_condition
7.4 Genomic Linkages	Variation_data
7.5 Risk Factors	Etiology_condition, Family_history
8. Temporal Properties	
8.1 Age of Occurrence & Frequency	Diagnostic_differential
8.2 Development Course / Stage	Diagnostic_differential
9. Severity of Subtypes Properties	Diagnostic_differential, Pathological_findings_expression_probability, Symptom_expression_probability,
10. Functioning Properties	
10.1 Impact on Activities and Participation	Symptoms_description
10.2 Contextual factors	Symptoms_condition
10.3 Body functions	Symptoms_description
11. Specific Condition Properties	
11.1 Gender	Gender (Disease_epidemiology > Differences 内)
11.2 Life-Cycle Properties	Striking_age (Disease_epidemiology 内)
12. Treatment Properties	Therapy (詳細項目有り)
13. Diagnostic Criteria	Diagnostic_criteria (詳細項目有り)
14. External Causes	None (ターゲットとしていない)

(表 2: ICD11 Content model から見た GSVML の臨床的疾患特性部分の対応)

Defining Attribute	Permissible Values (Category)
Finding Site	<i>Anatomical structure, Acquired body structure</i>
Associated Morphology	<i>Morphologically abnormal structure</i>
Associated With:	<i>Clinical finding, Physical object, Procedure, Physical force, Event, Pharmaceutical/biologic product, Organism, SNOMED-CT concept, Substance</i>
-CausativeAgent	<i>Organism, Physical force, Substance, Pharmaceutical/biologic product, Physical object, SNOMED-CT concept</i>
-DueTo	<i>Clinical finding, Event</i>
-After	<i>Clinical finding, Procedure</i>
Severity	<i>Severities</i>
Clinical Course	<i>Courses</i>
Episodicity	<i>Episodicities</i>
Interprets	<i>Observable entity, Evaluation procedure, Laboratory procedure</i>
Has Interpretation	<i>Findings values, Result comments</i>
Pathological Process	<i>Autoimmune</i>
Has Definitional Manifestation	<i>Clinical finding</i>
Occurrence	<i>Period of life</i>
Finding Method	<i>Procedure</i>
Finding Informer	<i>Performer of method (person), Subject of record (person), Provider of history other than subject (person), Subject of record or other provider of history (person)</i>

(表 3: SNOMED-CT における疾患特性記述モデル)

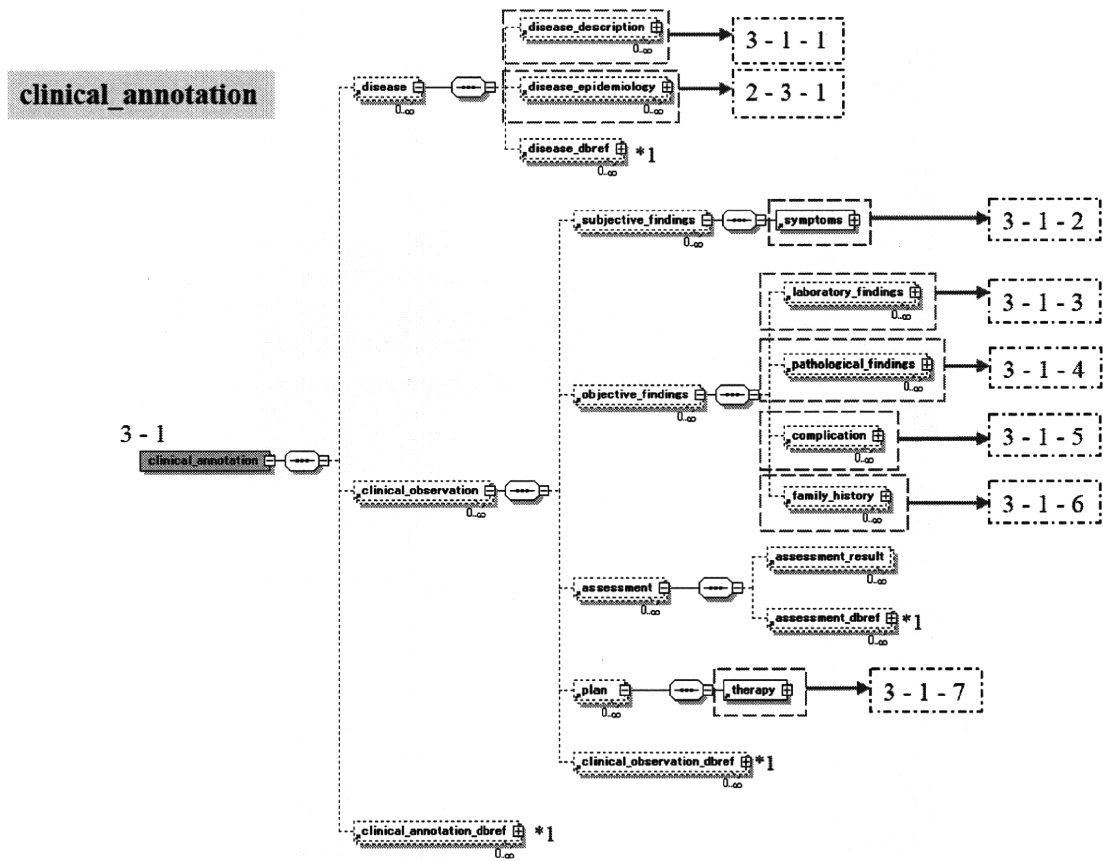
GSVML Disease_description	SNOMED-CT 疾患特性記述モデル
Name	★Description_table に概念表記ラベル有り
Synonym	★Description_table に概念表記ラベル有り
Definition_concept	Is-A
Classification	(inverse) Is-A
Etiology	
-Etiology_description	Associated_with::Causative_agent Associated_with::Due_to Associated_with::After
-Etiology_condition	Associated_with (上記3つの抽象化特性)
-Etiology_expression_probability	None (★上記3つの細分化 Associated_with の区別によって表現されている場合もあり)
Laboratory_findings	
-Laboratory_finding_type	Finding method
-Laboratory_findings_description	Associated_morphology, Has_definitional_manifestation, Has_interpretation, Interprets
-Laboratory_findings_condition	Finding_method, Finding_informer
-Laboratory_findings_expression_probability	None
Pathological_findings	
-Pathological_findings_description	Associated_morphology, Pathological_process, Has_definitional_manifestation, Has_interpretation, Interprets
-Pathological_findings_condition	Finding_method, Finding_informer
-Pathological_findings_expression_probability	Severity
Symptoms	
-Symptoms_description	Has_definitional_manifestation, Finding_site, Associated_with, Severity
-Symptoms_condition	Associated_with
-Symptoms_expression_probability	Severity
Diagnostic_criteria	

-Diagnostic_standard	None
-Diagnostic_modify	None
-Diagnostic_differential	Occurrence, Clinical_course
Therapy	
-Conservative	Associated_with
-Pharmaceutical	None
-Responder_side_effects_ Description	None
-Responder_side_effects_causer	None
-Responder_side_effects_condition	None
-Responder_side_effects_express_ probability	None
-Physical	None
-Surgery	Associated_with
-Radiation	Associated_with
Complication	
-Complication_description	Associated_with, Associated_with::After
-Complication_condition	Associated_with
-Complication_expression_probability	None
Prophylaxes	
-Prophylaxes_description	None
-Prophylaxes_condition	None
-Prophylaxes_expression_probability	None

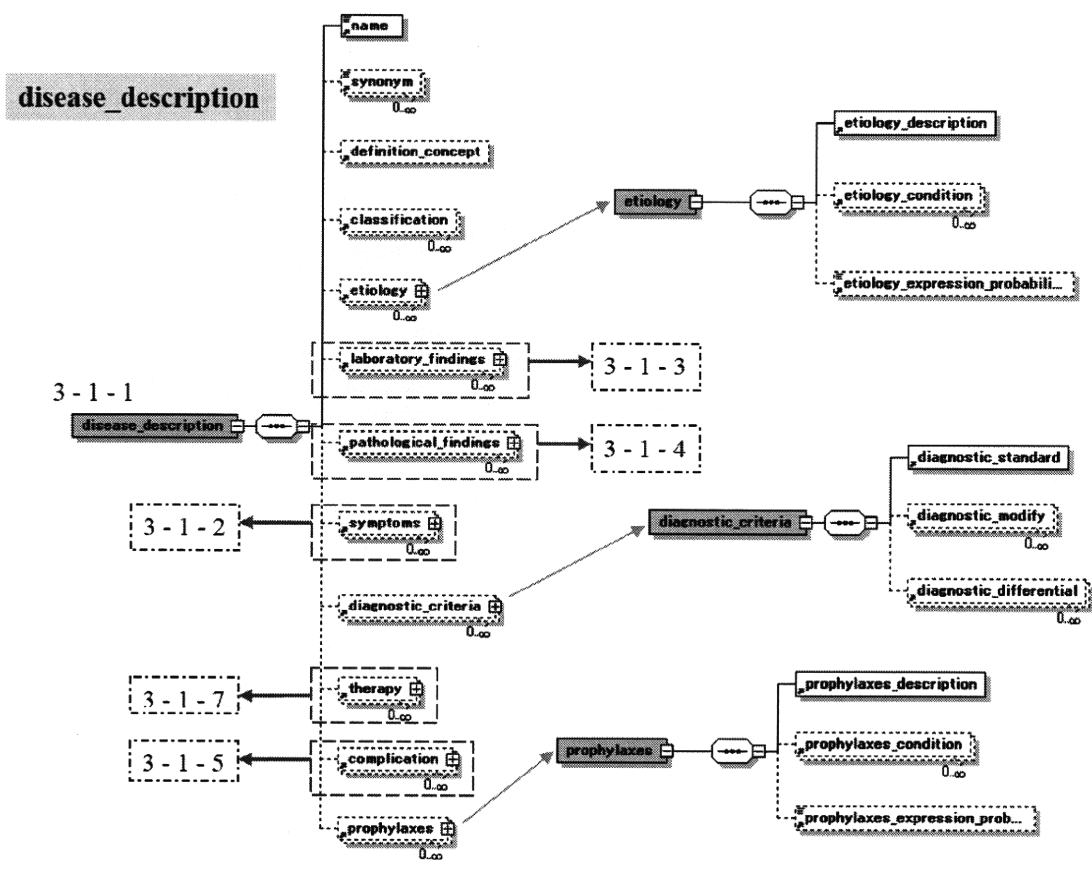
(表 4: GSVML Disease_Description から見た SNOMED-CT 疾患特性記述モデルの対応)

```
GSVMLroot - Variation_data
           - Variation_type
           - Location
           - Variation_att
           - ... etc.
           - Direct_annotation
             - Whole_genome_sequence
             - Mendelian_segregate
             - Homozygote_detect
             - Somatic_mutation
             - Experiment_analysis
             - Epidemiology
             - ... etc.
           - Indirect_annotation
             - Personal_info
             - Phenotype
             - Omics_annotation
             - Environmental_condition
             - Clinical_annotation
```

(図 1: GSVML 高位構造 (ISO25720 GSVML より一部抜粋))



(図 2: Clinical_annotation (ISO25720 GSVML より))



(図 3: Disease_description (ISO25720 GSVML より))