

厚生労働科学研究費補助金(政策科学総合研究事業(政策科学推進研究事業))

「社会構造の変化を踏まえた保健医療にかかる施策立案に資する

国際統計分類の国内導入のための研究」

分担研究報告書(令和2年度)

## ICD-11 導入における国内標準用語集とのマッピングに関する研究

研究分担者 今井 健 (東京大学大学院医学系研究科)

### 研究要旨

疾病及び関連保険問題の国際統計分類(ICD)の第11版改訂(ICD-11)がWHOによりリリースされ、今後国内適用に向けた準備が急務となっている。本研究班全体の目標はICD-11国内導入に資する各種分析を行い基礎資料として取りまとめることであるが、本分担研究ではその中でも特に、これまでICD-10に準拠して構築され広く診療現場において用いられてきた国内標準用語集とどのように整合性を取るか、を検討することを目的とする。本年度は、標準病名マスターを対象とし、ICD-11の現状の和訳案リストとのカバレッジを調査すると共に、主要なICD-10章の個々の病名を対象に、ICD-11コードとのマッピング調査を行った。結果として、ICD-11自体の和訳のみでは既存の国内用語集と一致させることは難しく、Post-Coordinationを駆使した対応付けが必要で、国内導入にあたってはこの作業が極めて重要であることが判明した。

### 研究協力者

滝澤雅美

国際医療福祉大学 講師

高橋幸恵

国際医療福祉大学 助教

坂本幸平

国際医療福祉大学 助教

での適用に向けたICD-11の妥当性の検討を行う必要がある。研究班全体の目標はICD-11国内導入に資する各種の分析を行い基礎資料として取りまとめることである。これには複数の要素が含まれ、例えば公的統計への影響なども重要であるが、考慮すべきことはそれだけに止まらない。ICD-11では(1)分類体系の構造自体の変化(2)個々の分類概念の詳細化(3)単なる分類体系の提供から用語集の提供への変化(4)Post-Coordinationを用いた任意粒度での概念のコーディング、と言ったこれまでにない特徴があり、国内でこれまで用いられてきたコーディングスキームとは大きく異なっている。国内の電子カルテ環境では厚生労働省標準ともなっているICD-10対応電子カルテ用標準病名マスター(以下、標準病名マスター)が広く導入されており、病名の記録とICD-10コ

### A. 研究目的

疾病及び関連保健問題の国際統計分類(ICD)は、我が国では死亡統計を始め患者調査、医療保険制度、診療情報管理に広く活用されている。WHOが進めているICD-10から11への改訂においては我が国の医療実態を踏まえたより適切な医療情報を将来的に確保するために、我が国

ードへの対応づけが行われているだけでなく、

その ICD-10 コードは DPC など診療報酬請求にも広く用いられている。一方、ICD はこれまで単なる疾患分類のためのカテゴリ体系を提供するものであったが、ICD-11 からは Foundation Component 中に約 13 万語の用語セットが含まれ、大きく拡充されると共に用語集としての性質を帯びてきた。従って、ICD-11 の日本語版を作成し国内導入する際には、これまで国内で標準的に用いられてきた用語集とのダブルスタンダードを避け、円滑な移行を促すため整合性をとる必要がある。

そこで本分担研究では、これまで国内で用いられてきた ICD-10 ベースのコーディングスキーム、特に標準病名マスターに焦点を当て、ICD-11 を国内導入した際にどのように整合性を担保すべきかを検討することを目的とする。

## B. 研究方法

ICD-11 の国内導入に向け、昨年度までの厚生労働省内の作業により、既に各学会から寄せられた ICD-11 Foundation の用語に対する和訳案の集約作業が終了している。現在、この和訳案の確認作業が進められているところで最終確定はしていないが、本研究では一旦集約が行われた和訳案セットを一時的な「ICD-11 日本語版（仮）」とみなし、標準病名マスターとの病名のカバレッジについて調査を行った。

次に、標準病名マスターと一致しなかった事例について調査を行い、対応づけを行うための方策についてパターンを分類し検討を行った。

さらに ICD-11 の Post-Coordination 機能を用いないと既存の標準病名が ICD-11 と対応付けられない事例が多数存在することが判明したため、本年度標準病名マスター中の個々の病名について主要な章を対象に、ICD-11 コードとのマッピング施行を行い、今後の対応付け作業のための知見を得た。

尚、本研究では倫理面への配慮は必要としな

い。

## C. 研究結果

### C-1) ICD-11 Foundation 和訳案と標準病名マスターとのカバレッジ

ICD-11 は疾病・傷病概念のコーディングに必要な全てのエンティティをプールする Foundation Component と、そこから死亡統計、罹患統計、プライマリ・ケアなど各種の目的に応じて適切な粒度を選択し、排他的分類体系として切り出した Linearization Component から成っている。エンドユーザーが使用するのは主に Linearization であるが、各種の Linearization の大元となるデータベースは Foundation Component であり、2020 年 6 月版では総数で 130,151 エンティティが含まれている。内訳を下記表 1 に示す。この中で定義文エンティティが 16,843、注記エンティティが 731 含まれているので、実質的な用語数は 112,577 語（カテゴリ名称約 3 万含む）である。各々の章に含まれる用語数は ICD-10 までと比べて大幅に増加し、Post-Coordination に使用される拡張用コード（X 章）も約 24,000 語を収載している。

表 1 ICD11Foundation の章ごとエンティティ数(2020/6 月)

章	内容(略記)	全数	章	内容(略記)	全数
1	感染症	8,174	15	筋骨格	6,696
2	新生物	8,138	16	泌尿器	4,361
3	血液	1,797	17	性保健	165
4	免疫	1,669	18	妊娠	2,370
5	内分泌	6,155	19	周産期	2,357
6	精神	3,709	20	発達異常	10,172
7	睡眠障害	250	21	症状所見	4,624
8	神経	5,588	22	損傷	7,617
9	視覚	3,759	23	外因	5,438
10	耳	663	24	保健サービス	2,996
11	循環器	4,143	25	特殊	47
12	呼吸器	2,785	26	伝統医学	1,102
13	消化器	5,921	V	生活機能	131
14	皮膚	4,902	X	拡張コード	24,422

この X 章は ICD-11 での大きな特徴であるが、臨床病態をより詳細にコーディングするため

に、重症度、時間的特徴、病因、位置関係、解剖構造、組織病理、損傷や外因の程度、意識レベル、物質・アレルゲン等々の用語が用意されている。21章の症状・所見用語と合わせ、国内でこれまで存在してこなかった領域の標準的な用語集・コードセットとして機能する可能性を秘めている重要なパートである。今後、X章を含めたこれら全体11万語のFoundation Componentの日本語版を整備していく必要がある。この作業は厚生労働省が中心となり、各関連学会との連携のもと進められており、訳語候補の集約作業と事務局での確認作業を経て、現在再度各学会での確認作業が始められているところである。

一方、これまで我が国でも臨床現場で用いられる標準用語集やコード体系の策定が進められてきた。現在、17の厚生労働省標準規格が定められており、その中で用語集・コード体系に関するものとしては、ICD10対応標準病名マスター5（以下標準病名マスター）、医薬品HOTコードマスター、臨床検査マスター、また標準歯科病名マスターや看護実践用語標準マスターといったものが存在する。しかし臨床情報のコーディングに必要な様々な領域の用語という観点では未だカバー範囲は限定されており、症状・所見や解剖用語を始め、アレルギーや化学療法などでも標準的用語・コード体系は存在していない。その中で、電子カルテの普及に伴い、比較的国内普及が進んでいるものの1つが病名である。MEDIS-DCからリリースされている「ICD10対応標準病名マスター」は、我が国における標準病名集で、電子カルテで広く用いられている。また、事実上同じ中身のものが社会保険診療報酬支払基金からも「レセプト電算処理用傷病名マスター」としてリリースされており、レセプトの傷病名欄記載、あるいはDPCでも用いられている。

標準病名マスターはV5.04版(2020年6月)では病名基本テーブル26,420件、修飾語テーブル2,393件、索引テーブル106,834件の規模

になっており、約26,000の病名代表表記と、対応する同義語・類義語、これらにPost-Coordination用の修飾語を組み合わせて用いることができるようになっている。今や用語集としての色合いを濃くしたICD-11の日本語版の国内導入を今後考えていく上では、この現在用いられている標準病名マスターとの適切な統合は必要不可欠である。

標準病名マスターでの病名表記は代表表記・同義語・類義語を合わせて107,760語存在するが、この中にはカタカナ表記の同義語(48,966語)も大量に含まれており、異字体による同義語(4,280語)も除くと、実質的な病名は50,042語である。これらについては適切な形でICD-11日本語版と対応付けられることが必要となる。

今回の調査では、2020年3月時点でのICD-11 Foundation日本語版作成作業中間データ(候補用語約10万語)と標準病名マスター中の約5万語とで文字列が一致するものは約8,600語程度であった。

但し、完全一致ベースでの検索であることと、日本語版用語が完全に定まっていない中での暫定的な結果であるため、実態を必ずしも反映していない可能性があることには注意が必要である。例えば、標準病名マスター中の「BL型ハンセン病」は、ICD-11にも「BL(境界らい種)型ハンセン病」、「境界らい種型ハンセン病」、「ハンセン病,境界らい種型」、(訳はいずれも仮案)などの用語エントリーが存在しており、完全一致はしていないが、細かな表記の揺れを考慮すれば対応付けが可能となる。このような限界はあるが、あくまでも速報値として、標準病名マスター中でICD-11日本語版作成作業中間データに対応付けができなかった病名数について、ICD-10での章の上位5つを以下に示す。

- ・5,129 S章 (損傷・中毒)
- ・4,653 T章 (損傷・中毒)
- ・4,125 C章 (新生物)

・ 3, 488 D 章 (新生物)

・ 3, 483 M 章 (筋骨格、結合組織の疾患)

上位を占めたのは S, T (損傷・中毒) と C, D (新生物) である。この上位 5 つは標準病名マスターにもともと収載されている語数が多い章の順位とも一致しており、自然だと考えられる。

一方、標準病名マスターのこれらの章は、詳細な部位を予め病名に含めたものが数多く収載されているという特徴を持つ。例えば「第 3 足趾末節骨開放骨折」は、ICD10 では S925 (Fracture of other toe) の .1 (open) である S9251 に分類される。ICD11 では、これに直接対応する語は無いが、ND13.5 Fracture of other toe に対する Post-Coordination として表現されるべきものである。従って ICD-11 の日本語版と直接の完全一致用語が存在せずとも妥当である。実際、ND13.5 Fracture of other toe では、左右性、骨折のサブタイプ、開放・閉鎖の別、詳細な解剖部位 を Post-Coordination で表現するため、「第 3 足趾末節骨開放骨折」も ND13.5 と、XA38Q1 (Distal phalanx of third toe) と XJ7YM (Open fracture) を組み合わせることで「完全に同一概念粒度で」コーディングされる。

すなわち、標準病名マスターと、ICD-11 日本語版との単純な用語一致数を考慮してもあまり意味を持たず、現在の標準病名マスター中の用語が、ICD-11 日本語版でどのように Post-Coordination で表現されるか、を調査することが重要であることが判明した。

## C-2) 標準病名マスター中の病名の ICD-11 コーディング

C-1) の結果を受け、次に、標準病名マスターの個々の病名に対し、必要に応じて Post-Coordination を用い、ICD-11 コードにて表現する作業(ICD-11 コーディング)を行った。全体としては非常に膨大な作業であり、効率化のため本年度は標準病名約 5 万語のうち、同義語ではなく「代表表記」と区分されている約 2 万語、

さらにその中で ICD-10 の主要な 7 章に含まれているもののみを対象とした。以下に ICD-11 コーディング対象数と作業完了数を示す。(作業完了数/コーディング対象数)

・ 1 章(感染症) (1, 029 / 1, 209)

・ 2 章(新生物) (130 / 3, 115)

・ 4 章(内分泌) (870 / 870)

・ 6 章(神経) (650 / 650)

・ 9 章(循環器) (712 / 712)

・ 10 章(呼吸器) (371 / 371)

・ 13 章(筋骨格) (1, 690 / 1, 690)

またコーディング作業においては、各々の標準病名に対し以下の 2 種類の対応付けが必要であることが判明した。

(1) Foundation 中のエンティティの URI(固有の ID)との対応付け

(2) MMS (Mortality and Morbidity Statistics Linearization) の粒度における対応付け

MMS は Foundation 中の限られた概念のみが選択されており、粒度としては粗いものである。従って、仮にある標準病名が Foundation 中のある用語と一致したとしても、それが MMS に含まれていなければ、より粗い粒度の MMS 用語と対応付けざるを得ず、もしその標準病名の概念を完全に MMS で表現しようとするれば、Post-Coordination を利用する必要がある。

例えば、

「**A 群連鎖球菌敗血症**」(標準病名) は Foundation 中の用語としては

**“Sepsis due to Streptococcus, group A without mention of septic shock”**

(<http://id.who.int/icd/entity/126483251>)

と完全一致している。()内は URI である。

しかし、MMS ではこの語は採用されておらず、

**1G40 “Sepsis without septic shock”**

の粒度までしか存在しない。従ってこの標準病名を MMS の粒度レベルで表現しようとする

**XN6LP “Streptococcus, group A”**

という X 章における就職語と組み合わせて、**1G40 & XN6LP** という Post-coordination 表現にする必要がある。つまり、

**A 群連鎖球菌敗血症 (ICD10: A400)** は  
= <http://id.who.int/icd/entity/126483251>  
(ICD-11 Foundation レベル)  
= **1G40 & XN6LP**  
(ICD-11 MMS レベル)

の 2 通りのコーディングとして表現される。元来この病名に付与されている ICD-10 コード A400 はこの病名概念が含まれる「より粒度の粗いカテゴリ」のコードであるが、ICD-11 になることにより、もっと詳細な病名自体が持つ概念粒度と等価なコーディングを行うことができる。ただしこの方法が Foundation レベルと MMS レベルで 2 通りある、ということである。

一方、MMS レベルの Post-coordination、あるいは粒度の細かい Foundation の用語を用いても、標準病名の意味粒度の方が細かく、対応付けできない事例も存在した。このような場合は ICD-11 の表現力をもって等価な表現はできず、病名より粒度が粗い表現に留めざるを得なかった。

例えば、**痛風発作 (ICD10: M1009)** は  
(ICD-11 Foundation レベル)  
<http://id.who.int/icd/entity/526468865>  
(ICD-11 MMS レベル)

**FA25.2Y Other specified gout without specification whether primary or secondary**  
までとの対応付けが限界であり、これ以上細かなコーディングは行えない。このような事例も少数ではあるが存在した。

#### D. 考察

本年度の研究で、現時点での ICD-11 Foundation の日本語版作成作業中間データ(約 10 万語)と、標準病名マスター中の病名(約 5 万語)では完全一致するものが約 8,600 語程度であり、今後これらの対応づけが ICD-11 の円滑な国内導入に向

けて極めて重要であること、また疾患領域によっては、ICD-11 側では Post-Coordination で表現することが前提となっているため、1 対 1 対応で両者に対応づけることは困難で、標準病名の ICD-11 コーディングが必要であることが判明した。

標準病名の ICD-11 コーディングについては代表表記区分であるものに限定し、主要な 7 章から作業を行っているが、代表表記の病名に対する ICD-11 コーディングが可能であれば、同義語についてもほぼ同様のコーディング結果となると考えられる。本年度の作業で ICD-11 との対応付けが行えたものは想定作業対象病名の 1/3 程度であるが、今後の本研究班の全研究期間にて、標準病名マスター病名と ICD-11 との対応付けが一通り完了する見通しである。

国内の電子カルテにて広く利用されている標準病名マスターと ICD-11 との対応付けが完了すれば、ICD-11 の国内導入に向けて大きく前進すると考えられる。ただし、現状検討しているのは標準病名マスターから ICD-11 への対応付けであり、逆方向、すなわち ICD-11 Foundation の和訳(約 10 万)の中で、標準病名マスター(約 5 万)に含まれない疾患用語も大量に存在する。これをどのように標準病名マスターに取り込み、統合していくかについては別途検討が必要である。

また、本年度の研究成果で非常に大きな知見が、1 つの標準病名に対する ICD-11 コーディングが複数レベルで行える、という点である。

Post-Coordination(コードの合成)を用いても標準病名の粒度が表現できない事例もあったが、仮にできたとしても、Foundation レベル、MMS レベルと複数の表現が可能である。このような複数のレベル・粒度のコーディングはその目的に応じて使い分けられるべきであり、今後ユースケースに基づいた粒度設定の議論が必要と考えられ

た。例えば、Foundation レベルの (URI を用いた) コーディングは最も粒度が細かいが、これは国の公的統計 (死亡・罹患) として用いられることは想定しづらく、細かな粒度が必要な研究用途に用いられると考えられる。また、MMS レベルのコーディングについても、目的に応じて必要な粒度を選定し、それ以上に細かな Post-Coordination については捨象する、という線引きが必要と考えられる。例えば、「**第3足趾末節骨開放骨折**」は、ICD-10 では **S9251** (その他の趾<足ゆび>の骨折・開放骨折) であった。これが今や ICD-11 では病名自身が持つ概念粒度を、**ND13.5 & XA38Q1 & XJ7YM** (Fracture of other toe & Distal phalanx of third toe & Open fracture) として Post-Coordination により完全に等価に表現可能である。しかし、国の公的統計 (死亡・罹患) においてここまでの粒度は必ずしも必要ではないと考えられ、その場合は例えば **ND13.5** 程度に丸めて集計する、という方針になると考えられる。

このように、ユースケースごとのコーディング粒度設定は、今後 ICD-11 を我が国に導入する際には非常に重要である。本研究では、複数のレベルでコーディングすることによって、後に研究・公的統計・診療報酬請求など様々な用途で異なる粒度のコードが選択できるようにしている点で極めて貴重なリソースとなると考えられる。

さらに、本厚労科研班全体の目標には、公的統計への影響分析が含まれている。ICD-10 と ICD-11 のカテゴリは必ずしも 1 対 1 対応ではなく、多対多対応であることも多い。従って、ICD-10 ベースで集計された統計データが ICD-11 ベースでどのように変化するか、を分析するためにはカテゴリベースの集計分析では不十分で、病名単位の粒度での解析が必要である。その場面においても、本分担研究の成果は活用可能であると考えられ

る。例えば、現在 ICD-11 導入の死亡統計への影響を調査するため統計法 33 条に基づいた死亡票情報の提供を受け分析準備中であるが、年間 130 万件に及ぶ死亡票を全て ICD-11 ベースで人手解析し直すことは現実的ではない。その際に、すでに病名ベースで ICD-11 コード付けが行えているものについてはこれを適用することで、ICD-10 ベースの確定原死因コードを ICD-11 ベースに自動変換することができ、大幅な効率化が図れると考えられる。このように本研究成果は、ICD-11 の国内診療現場への導入を円滑化するために必要不可欠なリソースであるのみならず、研究班全体の目標の効率化にも活用可能な基盤リソースとなると期待される。

## E. 結論

本年度研究では、ICD-11 の国内導入における既存標準用語集との整合性確保のため、標準病名マスターと ICD-11 Foundation 日本語版中間作業データとのカバレッジ調査を行った。完全一致するものは非常に少数であったが、一致しないものを対象に Post-Coordination を用いた ICD-11 コーディングを行った。結果として主要な ICD-10 7 つの章の標準病名について、MMS レベル、Foundation レベルなど複数の表現レベルでのマッピング作業が完了した。本研究により標準病名マスターの ICD11 化に向けた重要な知見が得られただけでなく、その成果は公的統計への影響分析を加速化する目的にも活用可能な貴重なリソースである。今後、全研究期間において標準病名マスターと ICD-11 との対応付けを完了させ、国内診療現場への導入準備を整えていく予定である。

## F. 健康危険情報

なし

#### G. 研究発表

1. 今井 健. 用語集として見た ICD-11 と国内適用に向けた課題. 医療情報学 40(Suppl.): 324-327, 2020.
2. 小川俊夫, 今井 健, 滝澤雅美, 小松雅代,

今村知明. ICD-11 国内適用に向けた日本語化の検討. 医療情報学 40 (Suppl.):328-331, 2020.

#### H. 知的財産権の出願・登録状況

なし