

ICD-11の国内実用化検討に向けた最新版ICD-10の日英統合管理システムとそのICD-11対応化に関する研究

研究分担者 今井 健（東京大学 大学院医学系研究科 講師）

研究要旨

疾病及び関連保健問題の国際統計分類(ICD)の第 11 版改定(ICD-11)の国内実用化を検討する上で、最新の第 10 版(ICD-10)情報は必須であるが、国内の ICD 適用と WHO ICD-10 の間には大きな時差があり、その乖離が問題となっていた。そこで平成 26 年度は国内の ICD-10(2003 年版)に対し、2013 年までに累積した修正勧告を適用し、最新の ICD-10 傷病名索引日英対応データベースを構築した。本年度はこの成果を基に、日本語版 ICD-10(2013 年版)準拠索引表の書籍刊行用の Linearization データを作成すると共に、今後日英双方において最新版 ICD-10 と ICD-11 との対応を統合的に管理するための Web 管理プラットフォームを開発した。本研究成果によって傷病名レベルで最新の WHO ICD-10 を国内適用し、ICD-11 との対応を含めて統合的に維持管理することが可能となると共に、今後 ICD-11 との対応分析を行う際に有用な基盤となると考えられる。

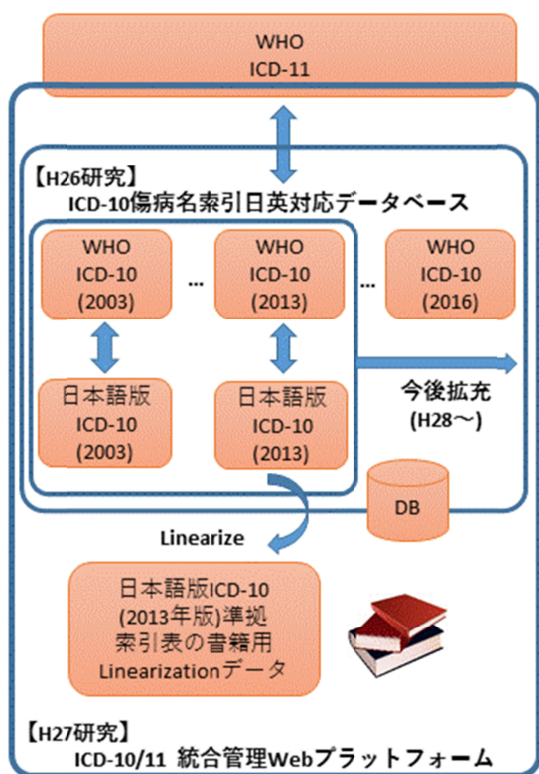
A. 研究目的

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。現在 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、我が国における ICD-11 の実用化に向けた検討を行う必要がある。このような背景の下、本研究事業全体の目的は、ICD 改訂作業をサポートすると共に、我が国の医療において適切な疾病分類をとりまとめること、また WHO 検討の場で行うべき対応に資する基礎資料を作成することであるが、そのために必要な要素の 1 つとして、現状における ICD-10 をはじめとした我が国における疾病分類の整理は重要な課題である。特に最新の傷病

名に基づいた分析を行うためには、ICD の傷病名索引は重要であるが、現在の我が国の ICD 告示は ICD-10 (2003 年版) に基づくものであり、WHO の ICD-10 (2010 年版) からの乖離が大きいという問題が以前から指摘されていた。

このような背景のもと、平成 26 年度は、国内の ICD-10(2003 年版)に対し、2013 年までに累積した修正勧告を適用し、最新の ICD-10 傷病名索引日英対応データベースを構築した。本年度はこの成果を基に、日本語版 ICD-10(2013 年版)準拠索引表の書籍刊行用の Linearization¹データを作成すると共に、今後日英双方において最新

¹ データベース中の構造を“Linearize”し、実際にエンドユーザーが目にする順序を伴ったデータ列に出力した結果。例えば Alphanumeric あるいは読み仮名順でソートした冊子等。



(図 1. 本研究全体のスキーム)

版 ICD-10 と ICD-11 との対応を統合的に管理するための Web 管理プラットフォームを開発することを目的とする。

B. 研究方法

B-1) 研究全体のスキーム

本研究全体のスキームを上記図 1 に示す。昨年度までに図 1 中央部分の ICD10 索引日英対応データベースの作成と ID 管理フレームワークの作成は完了している。本年度は次節 B-2) で述べる方法にて書籍用 Linearization を生成すると共に、ICD-11 との対応情報を記述するように枠組みを拡充し、ICD10/11 統合管理 Web プラットフォームとして実装した。

B-2) ICD-10 傷病名索引日英対応データベースと書籍用 Linearization

これまで我が国内では、ICD-10 (2003 年度版) 準拠のデータが国内適用されてきた。

昨年度作成した「ICD-10 傷病名索引日英対応データベース」では、WHO ICD-10 の 2003 年版 (WHO が 2003 年 1 月に集積版として公表したもの) から、2013 年版 (WHO が 2013 年 1 月に集積版として公表した、Approved date が 2012 年 10 月までのもの) までの累積修正勧告、さらにこれに加え、Implementation date が 2014 年 1 月の Minor update 分 (糖尿病など) と、2016 年 1 月である Major update 分 (腐骨形成 円板) を含んだものを、ICD-10 (2013 年版準拠) 国内適用版とするべく、準備を進めてきた。

WHO 修正勧告は英語エントリーに対して成されているため、修正作業も英語を基準として行う。しかしその結果を日本語書籍用に Linearize するためには日本語のヨミガナに基づいてソートし直さなければならない。索引にはインデントによる階層構造があるため、当初、上位階層からヨミガナを引き継いだ「階層パス付きヨミガナ」を振ることで、一意に並べ替えを出来ると考えた。例えば、以下のようなヨミガナ

上位 A (ヨミ A)

- 下位 B (ヨミ B)

- 下位 C (ヨミ C)

に対し、

上位 A (ヨミ A)

- 下位 B (ヨミ A ヨミ B)

- 下位 C (ヨミ A ヨミ C)

とした状態で、() 内の文字列でソートすれば良い。しかし、ここでいくつかの問題が判明した。

- (1) ヨミガナでソートするためには、ヨミガナに直した状態で文字コード順にソートすることになるが、UTF-8 では、例えば「ヴ」が「ン」の後に来ってしまう。「ウ」の後に「ヴ」が来るようにしたい、と言った書籍特有の要望に答えられない。
- (2) UTF-8 の文字コード順では、ひらがな「あ - ん」の次にカタカナ「ア - ン」が

並ぶため、カタカナ語が、全てのひらがな・漢字語の後に並んでしまう。

「あ」行であれば、その中で「初めにカタカナ語、次にひらがな・漢字語と並べたい」という書籍特有の要望に答えられない。

- (3) 同じヨミガナを持つ複数のエントリーがあった場合、区別が出来ない。例えば

上位 A (ヨミ E)

- 下位 B (ヨミ B)

上位 C (ヨミ E)

- 下位 D (ヨミ D)

に対し、

上位 A (ヨミ E)

- 下位 B (ヨミ E ヨミ B)

上位 C (ヨミ E)

- 下位 D (ヨミ E ヨミ D)

と引き継がせると、ヨミガナソート時に

上位 A (ヨミ E)

上位 C (ヨミ E)

- 下位 B (ヨミ E ヨミ B)

- 下位 D (ヨミ E ヨミ D)

と並んでしまう。

- (4) 上位からヨミガナを引き継がせた結果、偶然同じヨミガナになってしまうことがある。

上位 A (ア)

- 下位 B (イス)

上位 C (アイス)

に対し、単純にヨミガナを引き継ぐと下位 B の階層付きヨミガナは (アイス) となり、C と区別が付かない。

これらの問題に対処するため、以下のようなルールで生成される「書籍用 Linearization 文字列」を用いることとした。

- ・ 「ア」は「ア1」、「あ」は「ア2」等と置換することで、カタカナ語、ひらがな・漢字語の並び順を書籍特有の要望に基づき微調整できるようにした。

「ヴ」は「ウ4」と置換することで、

「ウ」行の最後にすることが出来る。

これで問題 (1), (2) に対応できた。

- ・ 上位からヨミガナを引き継ぐ際に、セパレータ文字「」を挟むこととした。これにより問題(4)に対応した。
- ・ ヨミガナの後に、セパレータ「`@`」を挟み、元々の文字列本体を付加することとした。これにより問題(3)に対応した。

以上をまとめると、

【「ア-ン + 数字」ペアの列で表現したヨミガナ+`@`+ エントリー文字列本体】を、上位階層から「」で引き継がせた階層パス付き文字列

を「書籍用 Linearization 文字列」とした。

これをプログラム中で文字コード順ソートすることで目的を達成することが出来る。

この Linearization 方式を用いて日本語用のソート ID (JID) を付与し、最終的に、日本語版 ICD-10(2013 年版)準拠索引表の書籍刊行用 Linearization データを作成した。

B-3) ICD10/11 統合管理 Web プラットフォームの構築

既に研究分担者は、平成 26 年度厚生労働科研特別研究「WHO ICD-10 一部改正に伴う分類比較とコーディング支援に関する研究(研究代表者: 大江和彦)」において、従来 Excel で維持管理されてきた ICD10 の日英対応表を、Web アプリケーションとデータベースで管理するためのプラットフォームについてプロトタイプ実装を行っている。しかし管理対象は ICD-10 のみであり、ICD-11 との対応分析を行うことができなかった。そこで、本研究ではこの成果を発展させ、各エントリーが ICD-11 ではどう変わるのか管理できるように拡充した上で、任意の時点での最新版 ICD-10 と ICD-11 の対応を、日英対応と合わせ統合的に管理す

るためのプラットフォームを構築した。データベースは MySQL を採用し、データベースインポートモジュールについては C#、Web アプリケーションについては、Java Play Framework 並びに Scala sbt を用いて実装した。

C. 研究結果

C-1) 日本語版 ICD-10(2013 年版)準拠索引表の書籍刊行用 Linearization データの作成
昨年度までに作成した「ICD-10 傷病名索引日英対応データベース」に対し、B-2)で述べた手法を用いて付与された日本語ソート ID (JID2016) を基にした、日本語版 ICD-10(2013 年版)準拠索引表の書籍刊行用 Linearization データを作成した。詳細の一部を別添資料に示す。

別添資料は JID にてソートされているが、JID は絶対位置・相対位置が決まっている特殊な事例を除き、上位エントリから B-2)で述べた手法で引き継がれた「書籍 Linearization 用文字列」を UTF-8 文字コードに基づき自動的に並べ替えた順序により決定されている。

また、全体は (1)「疾病および傷害の性質」55,483、(2)「外因」2,938、(3)「薬物および化学物質」5,296 エントリからなり、2016 年 2 月現在、印刷刊行に向け校正が行われている。

C-2) ICD10/11 統合管理 Web プラットフォームの構築

今後、傷病名単位で WHO ICD-10 改定のキャッチアップを継続すると共に、ICD-11 との対応分析を行うため、ICD10/11 統合管理 Web プラットフォームを構築した。概要を別添資料に示す。

これは図 1 (別添資料。以下同様) に示すようなユーザー権限管理が成された Web アプリケーション形式であり、従来 Excel で

管理されてきた ICD-10 索引の日英対応情報について表形式での GUI インターフェイスにて閲覧・編集が可能となっている。現在では ICD-11 に関する情報は入力されていないものの、データベース上は準備されており、ICD-10/11 の表示を切り替えることで対応を見ることが可能である(図 2)。また検索機能 (図 3, 4)、WHO 勧告を随時入力する機能 (図 5~9)、また WHO 勧告と無関係に、データベースの修正をする直接編集機能 (図 10, 11)などを備えている。

WHO の改訂勧告はそれが実際に適用される implementation date が規定されており、それは必ずしも現時点ではなく数年後のこともある。一方我が国でもそれを踏まえ改定を集積し、一定の間隔を経た後、国内告示を行っている。従ってある時点で何が日・英それぞれ ICD の最新情報なのか(どの変更が有効化されているのか)を調べる機能は統計分析上、極めて重要である。また、以前の修正履歴を全て見たい場合、あるいは複数回の修正を経た最終的な結果を見たい場合、今後行われる予定の修正を見たい場合も考えられる。そこで、指定した年月日に対し、「それ以前」あるいは「未来」の改訂について「全て表示する」か「マージした最終結果を表示する」か「表示しない」かを選択し、修正履歴の表示モードを切り替える機能を実装した(図 12, 13)。さらに現在閲覧してるモードで TSV 形式データをダウンロードすることが可能である。また、日本語版、英語版ではそれぞれ書籍用 Linearization が異なるため、その切り替え機能(図 14)、「疾病及び傷害」「外因」「薬物および化学物質」の切り替え機能(図 16)、登録されている有効開始年月日のリスト表示(図 17)、Excel から列を指定することで、データベースへとインポートする機能(図 18)などを備えており、継続的

な管理を可能とするための機能が実装されている。

D. 考察

従来国内では、ICD-10 (2003 年版準拠) においても、索引語の階層構造や書籍での並び順に基づいたソート用文字列が管理されてこなかった。英語版と日本語版では並び順が異なり、また階層構造も一部異なっていることから、WHO の修正勧告を順次適用する際に管理が非常に煩雑になる問題があった。本研究で構築した Web 管理プラットフォームにより、いつでも英語版、日本語版それぞれの構造・並び順を切り替えながら管理を行うことができ、また本研究で導入した書籍用 Linearization 方式により、いつでも書籍用データが整合性を持った規則で自動生成できることになる。作業労力の大幅な軽減とエラー防止の観点から大いに有用と考えられる。

また、今回は日本語版 ICD-10(2013 年版) 準拠を対象としたが、この国内書籍刊行を待たずして、既に現在次の改定バージョンである WHO ICD-10(2016 年版) が発表されており、今後 2013 年版から 2016 年版に向けてのアップデートを国内に適用しなければならない。この管理においても本研究で構築した Web 管理プラットフォームは大いに有効であると考えられ、次年度では厚生労働省大臣官房統計情報部企画課国際分類情報管理室の協力を得て、このアップデート作業を通じた Web 管理プラットフォームの実運用試験を行う予定である

また、同プラットフォームに、まだ現時点では ICD11 情報は入力されていないが、管理するための枠組みは構築済みであり、今後 WHO の ICD-11 改訂情報が入手でき次第入力が可能である。今後この作業を進めると共に、蓄積した情報を元に傷病名単位での ICD コード分析を支援する基盤である

ICD10/11 コーディング支援ツールの開発へ発展させて行きたいと考えている。²

E. 結論

本年度研究では、昨年までの成果を発展させ、日本語版 ICD-10(2013 年版) 準拠索引表の書籍刊行用の Linearization データを作成すると共に、今後日英双方において最新版 ICD-10 と ICD-11 との対応を統合的に管理するための Web 管理プラットフォームを構築した。本研究成果によって傷病名レベルで最新の WHO ICD-10 を国内適用し、ICD-11 との対応を含めて統合的に維持管理することが可能となると共に、今後 ICD-11 との対応分析を行う際に有用な情報基盤となると考えられる。

F. 健康危険情報

なし

G. 研究発表

なし

H. 知的財産権の出願・登録状況

なし

² WHO は現在 ICD-11 用のコーディングツールを作成している。

(http://icd11ct.cloudapp.net/ct/icd11beta_jlmms/en/current#/) この日本語版の作成は今後の重要な課題である。

