

# わが国におけるICD-11フィールドトライアル

## —診断用語コーディングの分野別解析—

サトウ ヨウコ ミズシマ ヒロシ  
佐藤 洋子\*1\*2 水島 洋\*3

**目的** ICD-11（国際疾病分類第11版）が10年以上の改訂作業を経て、2018年に公表された。ICD-11の適用性、信頼性、有用性などを検討するICD-11フィールドトライアルは、ICD-11改訂プロセスの最終段階に位置づけられ、WHOが主導となり国際的に共通のプロトコールとプラットフォームが提供された。わが国のICD-11フィールドトライアルは2017年に実施され、評価者として診療情報管理士378名が参加した。ICD-11フィールドトライアルでは、診断用語とケースサマリーが提供された。本研究では診断用語コーディングにおける分野ごとの解析を行い、分野ごとの特徴・課題を踏まえ、ICD-11のわが国への導入に向けた提言を試みた。

**方法** 評価者はICD-10とICD-11でそれぞれコーディングしたのち、難しかったか（「はい」「いいえ」）、コードの詳細さ（「ちょうどいい」「詳細すぎる」「十分でない」）、コードのあいまいさ（「いいえ」「はい」）について評価した。コーディング時間は自動記録された。各評価項目について、分野ごとのICD-10/11比較と、ICD-10およびICD-11での分野間比較を行った。ICD-10/11比較分析には、マクネマー検定およびウィルコクソンの符号付き順位検定を、分野間比較分析にはボンフェローニ補正をした正規近似による比率の差の検定およびウィルコクソンの符号付き順位検定を行った。

**結果** 評価者378名のうち、診断用語コーディング全症例の評価を終えたのは83名（22.0%）で、合計の回答数は38,654件だった。診断用語は22分野に分けられた。「難しかった」と答えた割合をICD-11とICD-10で比較したところ、19分野ではICD-11の割合の方が高かった（ $p \leq 0.035$ ）。詳細さが「ちょうどいい」と答えた割合のICD-10/11比較においては、13分野でICD-10の値が高かった（ $p \leq 0.044$ ）。一方で、「04免疫機構の障害」と「07睡眠・覚醒障害」はICD-11の割合が高かった（ $p < 0.001$ ,  $p = 0.038$ ）。詳細さが「あいまいである」と答えた割合のICD-10/11比較においては、18分野でICD-11の値が高かった（ $p \leq 0.006$ ）。

**結論** コーディングの難しさ、コードの詳細さ、コードのあいまいさをICD-11/10で比較したところ、総じてICD-10の評価が高い結果となった。理由としては、ICD-11に関する事前の教育不足や英語環境下での評価であったことが大きく影響したと考えられた。しかし新しく追加された免疫疾患や睡眠・覚醒障害の章では、ICD-11の評価が良好だった。ICD-11は2022年に発効予定であり、わが国で今後予定されているICD-11の日本語翻訳作業に並行し、わが国の医療・診療体制に即した独自のプロトコールによる導入のためのフィールドトライアルを行う必要がある。

**キーワード** ICD-11, 国際疾病分類, フィールドトライアル, 診療情報管理士, 診断用語コーディング, 情報共有

\* 1 防衛医科大学校防衛医学研究センター医療工学研究部門助教

\* 2 国立保健医療科学院研究情報支援研究センター客員研究員 \* 3 同センター長

## I 緒 言

死因や疾病の国際的な統計基準として世界保健機関（WHO）が公表している疾病及び関連保健問題の国際統計分類（International Statistical Classification of Diseases and Related Health Problems, ICD）は、国際生活機能分類（ICF）や医療行為の分類（ICHI）とともにWHO国際統計分類（WHO Family of International Classifications, WHO-FIC）の中心分類として位置づけられている。ICDは数字とアルファベットの組み合わせによるコードで構成され、国際比較のためだけでなく各国の医療統計や医療機関における診療記録の管理などにも活用されている。わが国にはじめてICDが導入されたのは1900年（明治33年）であり、人口動態統計や患者調査、社会医療診療行為別統計などの国の統計の他、DPC（診断群分類別包括評価）算定のための基礎資料として用いられており、医療の現場での重要度が非常に高い。ICD-11改訂作業は2007年に始まり、2018年6月に最終段階に達したICD-11が一般公開された<sup>1)2)</sup>。現行のICD-10への改訂（1990年）から約30年ぶりの大改訂となり、今後、2019年5月にWHO世界保健総会で採択され、2022年に発効が予定されている。

ICD-11は死因・疾病統計にとどまらず、基礎医学、臨床医学、そして公衆衛生分野など様々な場面での使用を想定したコード体系として整備され、病名コードだけでなく疾患概念を含めた情報提供システムも兼ねている。またネットワーク環境での使用を前提としたシステムとなっており、ICD-11ブラウザ<sup>3)</sup>と呼ばれるウェブ上での提供が基本となる。さらに関連用語によるコード検索を行うコーディングツール<sup>4)</sup>も提供され、あいまい検索や予測検索など機能が充実している。ICD-11はICD-10から分類項目の詳細化、分類軸の変更などが加わり、コードの数は約4000増加する予定である。新たに免疫系疾患、睡眠・覚醒障害などの章が追加されたことに加え、複合的かつ多様な疾患や

病態を表すための、コードの組み合わせや追加（子クラスタリング、エクステンションコード、ポストコーディネーションなど）が行える。

ICD-11改訂作業においてWHOは各国・地域にWHO-FIC協力センターを指定し、WHO-FICネットワークを組織し、その中で分類改正改訂委員会を中心とした世界中の専門家による議論が行われてきた。ICD-11改訂のプロセスの最終段階には、その適用性、信頼性、有用性などを検討する作業であるフィールドトライアルが位置づけられ、国際的に共通のプロトコルやプラットフォームが提供された。フィールドトライアルとは実際の業務をシミュレートしてシステムの妥当性を評価するプロセスのことである。ICD-11フィールドトライアルでは、コーディング業務に携わる様々な職種（診療情報管理士など）による評価を行い、課題や問題点を抽出し適切にフィードバックすることが求められる。WHOは2017年にフィールドトライアルを実施するためのプロコルヤリファレンスガイド、プラットフォーム（ICD-Fit）を公開し、ICD-11フィールドトライアルが世界各国・地域で実施された<sup>2)</sup>。わが国ではWHO-FIC協力センターである国立保健医療科学院がフィールドトライアルセンター（FTC）となり、2016年のプレテストおよびパイロットテスト<sup>6)</sup>を経て、2017年に日本診療情報管理学会、日本病院会の協力により全国の診療情報管理士を対象にしたフィールドトライアルを実施した。

ICD-11フィールドトライアルの概要を記す。FTCは専用のプラットフォームであるICD-Fitに評価者を登録し、個別にケースを割り振った。コーディングケースとして診断用語298件（診断用語コーディング）とケースサマリー30件（ケースコーディング）の2種類が用意された。評価者はICD-Fitにログインし、診断用語コーディングにおいては、表示されたケースをICD-10およびICD-11でコーディングし、コーディングが難しかったか、コードの詳細さ、コードのあいまいさを評価した。ICD-10とICD-11どちらを先にコーディングするかはケースや評価者によってランダムに設定された。

表1 診断用語コーディングの個別評価項目

このケースにコードを割り当てる際に問題が発生しましたか(難易度)
・はい(はいの場合、どんな問題が起こったか教えてください)
・いいえ
割り当てたコードの当てはまりレベルを評価してください(詳細さ)
・ちょうどいい
・詳細すぎる
・十分でない(広義すぎる)
コードの割り当てを行う際にあいまいさを感じましたか(あいまいさ)
・いいえ
・はい、分類名(コードタイトル)が曖昧だったから
・はい、包含用語が不足していたから
・はい、分類名(コードタイトル)が他のものと区別できなかったから
・はい、コーディングガイドに不備があり不明確だった
・はい、診断用語や記述があいまいだった
・はい、今回の診断をコードする私の知識が不十分だった
・はい、その他の理由

すべてのケースの評価を終えたのち、全体評価として、ICD-11の適用範囲、詳細さ、使いやすさ、電子ツールやリファレンスガイドの評価に加え、ICD-11の冗長性・欠落、ICD-11の改善点を評価した。ICD-10のコーディングはICD-10web版<sup>5)</sup>で、ICD-11のコーディングはICD-11β web版(ICD-11ブラウザ<sup>3)</sup>)やICD-11コーディングツール<sup>4)</sup>をそれぞれ用いた。一部、日本語翻訳の資料を提供したが、原則英語環境下で実施された。ICD-11フィールドトライアルの結果解析は順次行われており、本研究では診断用語コーディングにおける分野ごとの解析を行い、分野ごとの特徴・課題を踏まえたICD-11のわが国への導入に向けた提言を試みた。

## II 方法

日本診療情報管理学会、日本病院会の協力により全国の診療情報管理士378名が評価者としてICD-11Fitに登録され、診断用語コーディングのための診断用語298件が割り当てられた。評価者はICD-10とICD-11でそれぞれコーディングしたのち、難しかったか(「はい」「いいえ」)、コードの詳細さ(「ちょうどいい」「詳細すぎる」「十分でない」)、コードのあいまいさ(「いいえ」「はい」)について評価した(表1)。あいまいである理由は同時に選択された。ICD-11Fit上ではコーディング時間が自動記録

表2 診断用語コーディングにおける298診断用語の22分野別内訳と回答数

(単位 件)

ICD-11章	ケース数		回答数	
	N	%	N	%
合計	298	100	38 654	100
01 感染症又は寄生虫症	15	5.0	3 265	8.4
02 新生物	26	8.7	4 402	11.4
03 血液および造血器の疾患	7	2.3	1 167	3.0
04 免疫機構の障害	3	1.0	458	1.2
05 内分泌、栄養および代謝疾患	33	11.1	4 792	12.4
06 精神、行動又は神経発達の障害	11	3.7	1 574	4.1
07 睡眠・覚醒障害	6	2.0	856	2.2
08 神経系の疾患	14	4.7	1 864	4.8
09 視覚系の疾患	5	1.7	675	1.7
11 循環器系の疾患	24	8.1	3 086	8.0
12 呼吸器系の疾患	12	4.0	1 471	3.8
13 消化器系の疾患	32	10.7	3 782	9.8
14 皮膚の疾患	6	2.0	581	1.5
15 筋骨格系又は結合組織の疾患	14	4.7	1 419	3.7
16 腎尿路生殖器系の疾患	19	6.4	1 979	5.1
18 妊娠、分娩又は産褥	5	1.7	490	1.3
19 周産期又は新生児期に発生した病態	7	2.3	744	1.9
20 発達異常、先天奇形	2	0.7	223	0.6
21 症状、徴候又は臨床所見で他に分類されないもの	6	2.0	602	1.6
22 損傷、中毒又はその他の外因の影響	44	14.8	4 419	11.4
23 傷病又は死亡の外因	2	0.7	221	0.6
24 健康状態に影響を及ぼす要因又は保健サービスの利用	5	1.7	584	1.5

された。

評価項目のICD-10/11比較についてはマクネマー検定、コーディング時間のICD-10/11比較ではウィルコクソンの符号付き順位検定を行った。分野間の比較はボンフェローニ補正をした正規近似による比率の差の検定およびウィルコクソンの符号付き順位検定を行った。有意水準はすべて5%とした。

## III 結果

評価者378名のうち診断用語コーディング全症例の評価を終えたのは83名(22.0%)で、合計の回答数は38,654件だった。表2に診断用語298件の22分野内訳と回答数を示す。多少のばらつきはあるもののICD-11各章から網羅的に出題されおり、問題はランダムに表示されるため回答数の偏りはケースの偏りに準じていた。

個別評価の分野ごとのICD-10/11の比較を表3に示す。「難しかった」と答えた割合をICD-11とICD-10で比較したところ、「04免疫機構の障害」「09視覚系の疾患」「20発達異常、先天奇形」では差がみられなかったが、それ以

表3 章別の個別評価

ICD-11章	難しさ			詳細さ					
	「はい」の割合			「ちょうどいい」の割合			「詳細すぎる」の割合		
	ICD-11	ICD-10	p値	ICD-11	ICD-10	p値	ICD-11	ICD-10	p値
01 感染症又は寄生虫症	11.0	4.7	<0.001**	69.2	71.7	<0.001**	3.52	2.17	<0.001**
02 新生物	7.6	3.1	<0.001**	71.3	74.3	<0.001**	2.50	1.54	<0.001**
03 血液および造血器の疾患	7.1	1.8	<0.001**	72.1	78.2	<0.001**	1.46	0.77	0.077
04 免疫機構の障害	9.2	7.4	0.268	69.0	59.0	<0.001**	3.93	3.28	0.607
05 内分泌、栄養および代謝疾患	5.7	2.4	<0.001**	73.6	76.0	<0.001**	1.52	0.73	<0.001**
06 精神、行動又は神経発達障害	5.1	2.0	<0.001**	72.2	74.1	0.022*	1.52	0.83	0.043*
07 睡眠・覚醒障害	4.1	2.6	0.035*	72.8	70.3	0.038*	2.57	1.87	0.308
08 神経系の疾患	9.2	1.4	<0.001**	67.0	74.2	<0.001**	2.84	1.23	<0.001**
09 視覚系の疾患	2.8	1.9	0.327	75.4	74.7	0.620	2.07	0.59	0.013*
11 循環器系の疾患	5.3	2.5	<0.001**	72.0	73.1	0.057	2.88	1.46	<0.001**
12 呼吸器系の疾患	5.5	2.5	<0.001**	72.2	73.6	0.073	1.90	1.09	0.029*
13 消化器系の疾患	3.8	2.2	<0.001**	73.2	73.1	0.730	1.56	0.98	0.005*
14 皮膚の疾患	11.2	3.8	<0.001**	56.8	64.2	<0.001**	6.54	2.41	<0.001**
15 筋骨格系又は結合組織の疾患	9.7	3.3	<0.001**	58.4	66.1	<0.001**	4.09	2.75	0.004*
16 腎尿路生殖器系の疾患	8.2	2.7	<0.001**	64.4	68.6	<0.001**	3.49	2.12	<0.001**
18 妊娠、分娩又は産褥	9.6	3.9	<0.001**	61.6	69.0	<0.001**	6.33	5.31	0.180
19 周産期又は新生児期に発生した病態	5.6	2.4	<0.001**	73.9	76.1	0.044*	2.28	1.21	0.057
20 発達異常、先天奇形	2.7	0.4	0.063	77.1	78.9	0.455	2.24	0.90	0.375
21 症状、徴候又は臨床所見で他に分類されないもの	8.6	3.7	<0.001**	67.3	69.9	0.068	3.49	2.82	0.424
22 損傷、中毒又はその他の外因の影響	12.0	4.2	<0.001**	59.8	65.6	<0.001**	5.61	3.28	<0.001**
23 傷病又は死亡の外因	18.1	6.8	<0.001**	53.8	66.5	<0.001**	8.14	4.98	0.119
24 健康状態に影響を及ぼす要因又は保健サービスの利用	15.9	10.3	<0.001**	59.2	61.8	0.110	7.19	4.79	0.009*

注 \*p<0.05 in McNemar test, \*\*p<0.001 in McNemar test, +p<0.05 in Wilcoxon signed-rank test, ++p<0.001 in Wilcoxon

外の19分野ではICD-11の割合の方が高かった (p ≤ 0.035)。割合の差のオッズ比が10を超えたものは、「22損傷、中毒又はその他の外因の影響」(オッズ比 (OR) : 11.12, 95%信頼区間 (CI) : 7.82-16.29), 「03血液および造血器の疾患」(11.33, 4.95-31.96), 「19周産期又は新生児期に発生した病態」(13.00, 3.25-113.02), 「16腎尿路生殖器系の疾患」(14.63, 7.18-34.67), 「18妊娠、分娩又は産褥」(15.00, 3.81-129.54), 「08神経系の疾患」(37.25, 14.25-138.50) だった(表4)。これらの章をICD-11間で比較したところ、「22損傷、中毒又はその他の外因の影響」が難しいと答えた割合(12.0%)が最も高く、「3血液および造血器の疾患」(7.1%, p < 0.001), 「16腎尿路生殖器系の疾患」(8.2%, p = 0.001), 「19周産期又は新生児期に発生した病態」(5.6%, p < 0.001)と差がみられた。

詳細さの評価で、「ちょうどいい」と答えた割合のICD-10/11比較においては13分野でICD-10の値が高く(p ≤ 0.044), このうち差が大きい分野としてオッズ比が0.5以下だった

のは「23傷病又は死亡の外因」(OR : 0.18, 95%CI : 0.06-0.43), 「18妊娠、分娩又は産褥」(0.25, 0.12-0.48), 「03血液および造血器の疾患」(0.26, 0.16-0.40), 「08神経系の疾患」(0.33, 0.24-0.44), 「15筋骨格系又は結合組織の疾患」(0.41, 0.31-0.53), 「14皮膚の疾患」(0.43, 0.27-0.65), 「22損傷、中毒又はその他の外因の影響」(0.44, 0.37-0.52), 「16腎尿路生殖器系の疾患」(0.49, 0.37-0.65) だった(表4)。一方で、「04免疫機構の障害」と「07睡眠・覚醒障害」はICD-11の割合が高かった(p < 0.001, p = 0.038)。また、「詳細すぎる」と答えた割合のICD-10/11比較では、14分野でICD-11の割合が高く(p ≤ 0.029), そのうちオッズ比が5以上だったのは「09視覚系の疾患」(OR : 6.00, 95% CI : 1.34-55.20), 「14皮膚の疾患」(5.80, 2.22-19.19), 「22損傷、中毒又はその他の外因の影響」(5.68, 3.59-9.39), 「11循環器系の疾患」(5.40, 2.72-11.89) だった(表4)。

あいまいさの評価で「あいまいである」と答えた割合のICD-10/11比較においては18分野

表4 割合 (ICD-11/ICD-10) の差のオッズ比

あいまいさ			コーディング時間		
「はい」の割合			中央値		
ICD-11	ICD-10	p値	ICD-11	ICD-10	p値
18.7	13.3	<0.001**	109.0	54.0	<0.001++
17.2	10.4	<0.001**	52.0	22.0	<0.001++
15.9	8.0	<0.001**	40.0	22.0	<0.001++
18.3	24.9	0.003*	65.0	36.0	<0.001++
13.4	8.4	<0.001**	36.5	19.0	<0.001++
12.6	10.3	0.006*	29.5	16.0	<0.001++
12.5	11.4	0.228	31.0	13.0	<0.001++
21.1	10.2	<0.001**	36.0	14.0	<0.001++
8.1	8.6	0.788	23.0	16.0	0.018+
14.0	10.7	<0.001**	25.0	13.0	<0.001++
12.3	8.8	<0.001**	33.0	14.0	<0.001++
9.8	8.1	<0.001**	24.0	12.0	<0.001++
28.6	15.5	<0.001**	65.0	18.0	<0.001++
27.6	14.2	<0.001**	47.0	19.0	<0.001++
23.3	13.7	<0.001**	39.0	18.0	<0.001++
26.5	13.3	<0.001**	68.0	24.0	<0.001++
13.6	7.0	<0.001**	30.0	19.0	<0.001++
8.1	7.6	0.791	23.0	14.0	0.080
21.8	12.6	<0.001**	46.0	23.5	<0.001++
30.2	15.8	<0.001**	59.0	19.0	<0.001++
30.3	14.0	<0.001**	74.0	19.0	<0.001++
28.9	23.3	<0.001**	75.0	27.0	<0.001++

signed-rank test.

	オッズ比 (OR)	95%信頼区間 (CI)
難しさ		
「はい」の割合(ICD-11/ICD-10)の差*		
22損傷、中毒又はその他の外因の影響	11.12	7.82- 16.29
03血液および造血器の疾患	11.33	4.95- 31.96
19産期又は新生児期に発生した病態	13.00	3.25-113.02
16腎路生殖器系の疾患	14.63	7.18- 34.67
18妊娠、分娩又は産褥	15.00	3.81-129.54
08神経系の疾患	37.25	14.25-138.50
詳細さ		
「ちょうどいい」の割合(ICD-11/ICD-10)の差**		
23傷病又は死亡の外因	0.18	0.06- 0.43
18妊娠、分娩又は産褥	0.25	0.12- 0.48
03血液および造血器の疾患	0.26	0.16- 0.40
08神経系の疾患	0.33	0.24- 0.44
15筋骨格系又は結合組織の疾患	0.41	0.31- 0.53
14皮膚の疾患	0.43	0.27- 0.65
22損傷、中毒又はその他の外因の影響	0.44	0.37- 0.52
16腎路生殖器系の疾患	0.49	0.37- 0.65
「詳細すぎる」の割合(ICD-11/ICD-10)の差***		
09視覚系の疾患	6.00	1.34- 55.20
14皮膚の疾患	5.80	2.22- 19.19
22損傷、中毒又はその他の外因の影響	5.68	3.59- 9.39
11循環器系の疾患	5.40	2.72- 11.89

注 1) \*表3で割合の有意な差 (p < 0.05) がみられたもののうち、オッズ比が10以上のものを抜粋  
 2) \*\*表3で割合の有意な差 (p < 0.05) がみられたもののうち、オッズ比が0.5以下のものを抜粋  
 3) \*\*\*表3で割合の有意な差 (p < 0.05) がみられたもののうち、オッズ比が5以上のものを抜粋

でICD-11の値が高かった (p ≤ 0.006)。「あいまいである」と答えた理由としてはどの分野においても「今回の診断をコードする知識が不十分だった」が高かった。一方、「04免疫機構の障害」ではICD-10の割合が高かった (p = 0.003)。

コーディング時間は「20発達異常、先天奇形」を除いたすべての分野でICD-11のほうが長かった (p ≤ 0.018)。ICD-11においてコーディング時間が最も長かったのは「01感染症又は寄生虫症」であり、これは「23傷病又は死亡の外因」以外のすべての章より有意に長かった (p ≤ 0.001)。

#### IV 考 察

ICD-11フィールドトリアル診断用語コーディングにおける分野別の解析において、コーディングの難しさ、コードの詳細さ、コードのあいまいさをICD-11とICD-10で比較をしたところ、総じてICD-10の評価が高い結果となった。「22損傷、中毒又はその他の外因の影響」

において、難しいと答えた割合はICD-11のほうがICD-10より大幅に高く、またICD-11の中の他の分野との比較でも高かった。また、コードの詳細さの評価において「詳細すぎる」と答えた割合がICD-10に比べて高かった。つまり、ICD-10に比べてコード体系が詳細になりすぎたことでコーディングの難しさを感じる割合が高くなったと考えられる。ICD-11はICD-10からコードが大幅に増加しただけでなくコード体系も変わり、より詳細な病態をコーディングできるようになっている。しかし、コードが詳細になりすぎることによって、ICD-11の構造をよく理解していないとコードを検索する手間がかかってしまう問題が生じる。実際にコーディング時間はICD-11のほうが長く、この傾向はほとんどの章でみられた。また、詳細な分類体系であるがゆえに、医師が記載するカルテなどから詳細情報を読み取られなければ、詳細不明コードが増えてしまう事態が考えられる。ICD-11は臨床の様々な場面での適用が可能になっているが、急性期医療、慢性期医療、救急医療など診療体制に応じたコードの深度（どこまで詳細に

コードすべきなのか)の共通認識を設定しておくことが不可欠となるだろう。本解析でも明らかのように、詳細さは章ごとによっても異なるため、章ごと、疾患ごとなど詳細な検討が重要となる。

総じてICD-10の評価が高い結果となった理由としては、ICD-11に関する事前の教育不足や英語環境下での評価であったことが大きく影響したと考えられる。ICD-11のコード体系やコーディングルールについては、WHOからICD-11リファレンスガイドが提供された。全体評価でのリファレンスガイドの評価では「具体的診断指針が示されていてわかりやすかった」「コードが合っているかを確認するのに便利だった」という意見が得られた一方で、「提供された日本語翻訳が分かりにくかった」という意見も挙げられた。英語環境下での実施については「日本語で評価をしたい」という意見が数多く挙げられた。ICD-11では関連語からコードを検索できるコーディングツールが提供されるが、現在は英語バージョンのみの運用となっている<sup>3)</sup>。英語と日本語の言語学的な違いを踏まえると、システムを単に日本語翻訳しただけでは、部分検索や曖昧語検索に問題が生じることが考えられる。検索システムのプログラム構造そのものを日本語適応させる必要があるだろう。

ICD-11の評価が良好だった分野もあった。「04免疫機構の障害」は難しいと答えた割合はICD-10のほうが高く、「ちょうどいい」「あいまいである」と答えた割合はICD-11が高かった。また、「07睡眠・覚醒障害」も、詳細さが「ちょうどいい」と答えた割合はICD-11が高かった。これらの章はICD-11で新たに導入されたものである。ICD-10ではコーディングが困難だった疾患がICD-11では適切にコーディングされる可能性が期待される。

また、WHOが定めた標準コード(Gold standard)との一致率を検討する解析<sup>6)</sup>においては、コードと病名の関係がICD-10では「1対複数」だったため一致率が低かったものの、ICD-11で「1対1」の関係となったことで一致率が上

昇したケースがあった。これはICD-11が、より正確に病態を表現できるようになったことを反映していると考えられる。一方で、ダブルコードやポストコーディングなど組み合わせるコードの数が増えるケースにおいてはICD-11の一致率が低下する傾向がみられた。今後、分野ごとの解析などを含め、正解率の低下につながる要因を詳細に評価する必要がある。

ICD改訂のプロセスにおいてフィールドトライアルが実施されたのは今回の改訂が初めてであり、WHOが主導となり国際的に共通のプロトコルが提供された。これにより、質的にコントロールされた結果が得られることとなり、各国や地域間の比較解析などが期待されていたが、評価項目の集計以外の解析手法に関する統一的な指針が示されず、各国の報告内容にばらつきが生じている<sup>5)</sup>。またWHOが各国・地域のフィールドトライアルの結果をどのような方法で解析および解釈をし、ICD-11改訂に反映させたかの報告が待たれるところである。

ICD-11は2022年に発効予定である。わが国では現在、ICD-11の日本語翻訳作業が始まろうとしている。リファレンスガイドやコーディングツールなどの翻訳と並行し、わが国の医療・診療体制に即した独自のプロトコルによる導入のためのフィールドトライアルが必須となる。評価の基本単位としては今回のような分野(章)が望ましいと考えられ、ICD-11導入における課題抽出を医療従事者や診療情報管理者が情報共有を行いながら、詳細な検討を行うことが重要である。

## 謝辞

本研究は、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)「社会構造の変化を反映し医療・介護分野の施策立案に効果的に活用し得る国際統計分類の開発に関する研究」で実施した。ICD-11フィールドトライアル実施にあたり多大なご尽力をいただいた一般社団法人日本病院会および日本診療情報管理学会の関係者の皆様へ感謝を申し上げる。本論文の内容は第77回日本公衆衛生学会総会に

て発表した内容を改変した。

#### 文 献

- 1) 厚生労働省ホームページ. 国際疾病分類の第11回改訂版 (ICD-11) が公表されました (<https://www.mhlw.go.jp/stf/houdou/0000211217.html>) 2019.4.22.
- 2) ICD-11 (<https://icd.who.int/en/>) 2019.4.22.
- 3) ICD-11ブラウザ (<https://icd.who.int/browse11/f-m/en>) 2019.4.22.
- 4) ICD-11コーディングツール ([https://icd.who.int/ct11\\_2018/icd11\\_mms/en/release#/](https://icd.who.int/ct11_2018/icd11_mms/en/release#/)) 2019.4.22.
- 5) ICD-10Version : 2016 (<https://icd.who.int/browse10/2016/en>) 2019.4.22.
- 6) 佐藤洋子, 水島洋. ICD-11フィールドトライアルについて. 保険医療科学 2018 ; 67(5) : 508-17.
- 7) 佐藤洋子, 水島洋, 木村映善, 他. わが国におけるICD-11フィールドトライアル-診断用語コーディングにおけるゴールドスタンダードの解析-. 第38回医療情報学連合大会 (第19回日本医療情報学会学術大会) ; 2018.11.22-25 ; 福岡. 医療情報学 2018 ; 38(suppl) : 800-2.