

特集：WHO 国際疾病分類第 11 回改訂 (ICD-11) および ICF, ICHI の導入に向けて

<総説>

ICD-11 フィールドトライアルについて

佐藤洋子^{1,2)}, 水島洋²⁾

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

²⁾ 国立保健医療科学院研究情報支援研究センター

Field trial of ICD-11

Yoko Sato^{1,2)}, Hiroshi Mizushima²⁾

¹⁾ Division of Biomedical Engineering, National Defense Medical College Research Institute, National Defense Medical College

²⁾ Center for Public Health Informatics, National Institute of Public Health

抄録

ICD-11(国際疾病分類第11版)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改定プロセスの最終段階に位置付けられ, 国際的に共通のプロトコールで行われることとなった. WHOは2016年のパイロットテストに加え, 2017年にフィールドトライアルを実施し, 我が国を含め世界各国・地域が参加した. 分類粒度が細くなったICD-11ではICD-10よりも精緻な疾患概念の記述が可能になった一方で, 「どこまで細かくコードを付与すべきか」といった, 診療体制に応じた適切なコーディング粒度の検討が必要な点が課題として挙げられた. また, ICD-11で新たに設けられた, 部位, 重症度, 時間軸などを表現するためのエクステンションコードを用いたポストコーディネーションの周知および教育の重要性も指摘されている. 2019年にICD-11が公開されたのち, フィールドトライアルで得られた知見・課題をもとに, 各国でのスムーズな導入・運用に向け, より具体的な課題や指針を明確にするための議論が活発になるであろう. 我が国独自の医療提供体制や公的統計におけるICD-11の円滑な導入・運用に向け, ICD-11フィールドトライアルの追加実施の検討が望まれる.

キーワード: ICD-11, ICD-11フィールドトライアル

Abstract

The ICD-11 (11th revision of the International Classification of Diseases) field trial is positioned at the final stage of the ICD-11 revision process, and its purpose is to evaluate the applicability, reliability, and usefulness of ICD-11. WHO held the field trial in 2017 with a common protocol, in addition to the pilot test in 2016, in several countries and regions around the world, including Japan. Several field trials said that ICD-11 had finer classification granularity, and that it was possible to describe more detailed disease concepts than ICD-10. On the other hand, they showed the necessity of discussing an appropriate coding granularity, such as "To what level should we code." In addition, the importance of education on post-coordination using extension codes that express parts, severity, time axis, and so on was pointed out. After releasing ICD-11 in 2019, discussions to clarify more specific issues and to establish guidelines for the smooth introduction

連絡先: 佐藤洋子
〒359-8513 埼玉県所沢市並木3-2
3-2 Namiki, Tokorozawa, Saitama 359-8513, Japan.
E-mail: ysato@ndmc.ac.jp
[平成30年11月29日受理]

of ICD-11 and operation based on the results of field trials will be active in several countries. Additional implementation of ICD-11 field trials will be desirable for the smooth introduction and operation of ICD-11 in Japan's own medical provision system and public statistics.

keywords: ICD-11, ICD-11 field trial

(accepted for publication, 29th November 2018)

I. はじめに

WHOが勧告している国際統計分類をWHO Family of International Classifications (WHO-FIC) といい、ここには国際疾病分類 (ICD) や国際生活機能分類 (ICF) などが含まれる。WHOは各国・地域にWHO-FIC協力センターを指定し、WHO本部、WHO地域事務局とともにWHO-FICネットワークを組織している。ICD-11改定作業はWHO-FICネットワークで組織された分類改正改訂委員会を中心に世界中の医学、分類、統計の専門家によって行われ、毎年開催されるWHO-FICネットワーク年次会議で各国の国際統計分類担当者による報告や議論が行われている。

フィールドトライアルあるいはフィールドテストとは、一般的にシステム開発などの分野で広く使われている言葉であり、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を意味する。ICDの改定に際して行われる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による評価を行うことで、改定後のICDの適用性、信頼性、有用性などを検討するとともに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。しかし、ICD-11以前の改定プロセスにはフィールドトライアルは含まれておらず、そのほとんどが新分類システムへの移行を促進するための事後措置として各国独自の手法で行われてきた[1]。一方、ICD-10第5章「精神及び行動の障害」のフィールドトライアルはICD-10改定前に実施され、他の分類への影響などを複合的に評価した結果がICD-10に反映された[1]。この経験から、ICD-11ではフィールドトライアルが改定プロセスの最終段階に位置づけられ、国際的に共通のプロトコルで実施されることとなった。

WHOは2017年にフィールドトライアルを実施したが、それに先立ち、2016年にパイロットテストを11か国で行った[2]。我が国では、現在までに、WHOから正式なプロトコルが発表される前のプレテスト実施に加え、パイロットテストとフィールドトライアルに参加した。

本稿では、ICD-11フィールドトライアルの各国および本国の実施概要を報告する。まず始めに、次章においてICD-11フィールドトライアルのプロトコルを示す。3章では、WHO-FICネットワーク年次会議2016、同2017、同2018において報告された各国のICD-11フィールドトライアルの概要を紹介し、続いて4章では、我が国で実施されたICD-11フィールドトライアルについて述べる。5章では、ICD-11フィールドトライアルの展望を述

べ、最後に結語を記す。

なお、伝統医学分野のICD-11フィールドトライアルは別途実施されており(本誌渡辺賢治先生著「東洋医学におけるICD-11活用」参照)、本稿で述べるICD-11フィールドトライアルは伝統医学分野を除いた分野を対象にして実施されたものである。

II. ICD-11 フィールドトライアルのプロトコル

本章ではICD-11フィールドトライアルのプロトコルの概要を述べる。

ICD-11フィールドトライアルでは専用のプラットフォームであるICD-Fitが用意され(図1)、ウェブ上で実施された。フィールドトライアルセンター(FTC)が評価者をICD-Fitに登録し、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件とケースサマリー30件の2種類が用意された。評価者はICD-Fitにログイン後、表示されたケースをICD-10およびICD-11でコーディングした。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。ICD-10のコーディングはICD-10web版で、ICD-11のコーディングはICD-11コーディングツールやICD-11βweb版(ICD-11ブラウザ)をそれぞれ用いた。これらの電子ツールについての詳細は本号「ICD-11の情報学的な意味」を参照されたい。

診断用語コーディング(ラインコーディング)では評

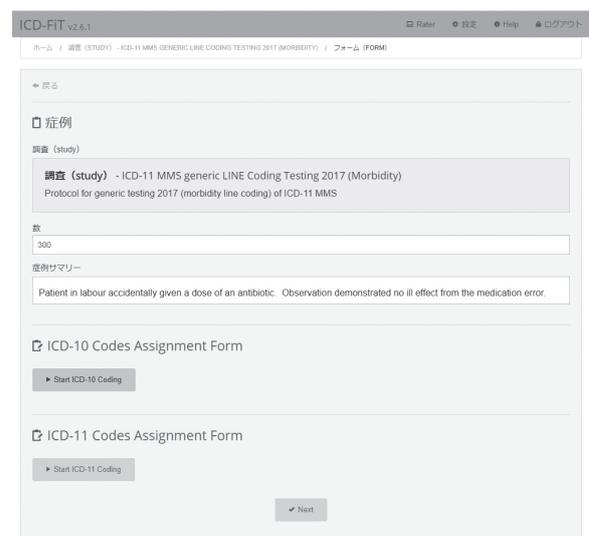


図1 ICD-Fitの画面

評価者はコードを付与したのち、個別評価としてコーディングの難しさ、コードの特異度（詳細さ）、コードのあいまいさを評価した。（図2）ケースサマリーコーディングではケースサマリーと主病名が表示され、その主病名が適切かどうかを判断したのち、主病名のICD-11によるコーディングや難しかったかどうかなどの評価を行った。（図3）

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
・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
・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
・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

図2 診断用語コーディングの個別評価項目

1. 記載された主たる病名は「遭遇の原因か、エピソードの終わりの入院時のものですか？
いいえを選んだ場合
1.1 新たな主病名を記載してください
1.2 主病名を再選択した際に用いた罹患ルール（Morbidity Rule）を選んでください。
・Mb1 主病名に複数の状態が含まれていた
・Mb2 主病名として記録された状況は診断・治療された症状であった
・Mb3 兆候と症状
2. このケースにおいて罹患ルールを適用する際に問題がありましたか。
・はい
・いいえ
2.1 「はい」であればその点を記載してください。
3. 主病名および他の症状に対するICD-11コードを記載してください。必要に応じてポストコーディネーションあるいはクラスターコーディングを使用してください。
4. 主病名コードに割り当てた主条件の幹コードを見つけるのは難しかったですか。
・はい
・いいえ
4.1 「はい」の理由を教えてください。
・コーディングツールで見つからなかった
・候補が多く、正解を探すのが難しかった
・除外事項の問題

・包含/同義語の問題
・補足に関する問題
・そのほか
5. 主病名コードのポストコーディング/クラスターコーディングの際に問題がありましたか・
・該当せず（ポストコーディングは不要）
・はい
・いいえ
5.1 「はい」の場合、理由を教えてください。
・システムコードでポストコーディングのインターフェース機能が使用できなかった
・ポストコーディングのインターフェース機能は使えたが正解を探すことが困難だった
・ポストコードがなかった
・そのほか
6. 他の病名コードに割り当てた主条件の幹コードを見つけるのは難しかったですか。
・はい
・いいえ
6.1 「はい」の理由を教えてください。
・コーディングツールで見つからなかった
・候補が多く、正解を探すのが難しかった
・除外事項の問題
・包含/同義語の問題
・補足に関する問題
・そのほかの問題
7. 他病名コードのポストコーディング/クラスターコーディングの際に問題がありましたか・
・該当せず（ポストコーディングは不要）
・はい
・いいえ
7.1 「はい」の場合、理由を教えてください。
・システムコードでポストコーディングのインターフェース機能が使用できなかった
・ポストコーディングのインターフェース機能は使えたが正解を探すことが困難だった
・ポストコードがなかった
・そのほか

図3 ケースサマリーコーディングの個別評価項目

提供されたすべてのケースのコーディングを終えると、評価者は全体評価の項目（図4）が選択できるようになり、ICD-11の網羅性や詳細度、使いやすさ、ICD-11コーディングツールやICD-11ブラウザなどについて評価した。結果はFTC管理者がcsvでダウンロードでき、上述した評価項目のほか、自動的に記録されたコーディング時間データが取得された。

1.1 ICD-11の収載範囲を評価してください
・とても良い
・良い
・ふつう
・悪い
・とても悪い
1.2.1.1の評価の理由を簡単に説明してください。

ICD-11 フィールドトライアルについて

2.1 ICD-11の詳細度を評価してください
・ちょうどいい
・詳細すぎる
・広義すぎる
2.2.2.1の評価の理由を簡単に説明してください。
3.1 ICD-11の使いやすさを評価してください
・非常に使いやすい
・使いやすい
・普通
・使いにくい
・非常に使いにくい
3.2.3.1の評価の理由を簡単に説明してください
(4の設問無し)
5. ICD-11に大きな欠落はありましたか
6. ICD-11に冗長性はありましたか
7.1 コーディングで使用した電子ツール（コーディングツール、ICD-11ブラウザ）を評価してください
・とても良い
・良い
・ふつう
・悪い
・とても悪い
7.2.7.1の評価の理由を簡単に説明してください。
8.1 ICD-11リファレンスガイドに記載されているコーディングガイドランスや説明を評価してください
・とても良い
・良い
・ふつう
・悪い
・とても悪い
8.2.8.1の評価の理由を簡単に説明してください。
9. ICD-11の改善について意見があれば教えてください

図4 全体評価項目

なお、パイロットテストも同様のプロトコルで実施されたが、ICD-Fitデモバージョン上で行われた。また、診断用語はフィールドトライアルで提供されたものと異なる約400件（国によって問題数が異なった）が提供された。ケースサマリーは提供されなかった。

III. 各国のICD-11 フィールドトライアル

この章では、WHO-FICネットワーク年次会議2016（東京開催、2016年10月8-12日）、同2017（メキシコ国開催、2017年10月16-21日）、同2018（大韓民国開催、2018年10月22-27日）において報告されたICD-11フィールドトライアルに関する研究発表（ポスター発表）を抽出し概要を紹介する。

WHO-FIC年次会議ネットワーク2016、同2017、同2018における、ポスターによる研究発表は全部で334件だった。このうちICD-11に関する内容は45件あり、この中でICD-11フィールドトライアルに関する研究は14件だった。報告国の内訳は、オーストラリアから5件、日本から3件、カナダから2件、大韓民国から2件、カリブ諸国7地域から1件、WHOから1件であった。表1に概要を示す。どの国でも評価者として医師かその国の臨床情報管理士にあたる専門家が参加していた。なお、パイロットテストの参加国内訳や、フィールドトライアルの参加国総数および内訳は本稿執筆時点でWHOから公表されていない。

1. オーストラリアにおけるICD-11フィールドトライアル（ニュージーランドおよび北欧諸国を含む）

オーストラリアからはパイロットテストに関する研究が3件[3-5]とフィールドトライアルに関する研究が2件[6,7]報告された。オーストラリアWHO-FIC協力センター（オーストラリア、ニュージーランド）と北欧WHO-FIC協力センター（デンマーク、フィンランド、アイスランド、ノルウェー、スウェーデン）による共同で行われたパイロットテストでは、評価者25名間のコード一致率はICD-10のほうがICD-11より高かったものの、章によってばらつきがみられたと報告されている[5]。2017年4月から6月に実施されたフィールドトライアルにはオーストラリアおよびニュージーランドから49名が参加し、ポストコーディング（追加コードを付与すること）が必要だったケースの多くで、最初に付与したコードの評価者間一致率が2番目に付与したコードの評価者間一致率より高かったことを報告し、ポストコーディングに関する教育を充実させることとともに継続してフィールドテストを実施する必要性を述べている[6]。また追加のフィールドトライアルが2018年3月に行われ、機能改善されたコーディングツールによりコーディングエラーが減少したことが報告されている[7]。

2. カナダにおけるICD-11フィールドトライアル

カナダのフィールドトライアルは評価者8名を対象に行われた[8]。付与されるコードが1つの場合の一致率が79%だったのに対し、付与されるコードが増えると一致率が9%まで低下したことを報告し、ICD-11の一貫性や信頼性を高めるためにはポストコーディングの教育を充実させることが必要だと述べている。また、独自のプロトコルによる追加フィールドトライアルも行っている[9]。あらかじめICD-10でコードした入院サマリー730件の主病名に対して、評価者6名が付与したICD-11コードと、WHOから提供されたICD-10/ICD-11クロステーブルから導かれるICD-11コードとの適合度解析を行い、全サマリーのうち79.0%で両者が一致したことから、ICD-10からICD-11への変換において、実際のコーディングとWHOによる定義の一貫性が高いことを報告した。

表 1 各国におけるICD-11フィールドトライアル実施状況

No	参考文献番号	研究演題	著者	発表国	種類	実施期間	評価者数(職種)	評価対象ケース	備考
1	[2]	Line coding pilot testing ICD-11 MMS	N. Kostanjsek, R. Jakob, C. Celik, A. Schmitter, L. Moskal, L. Best	WHO	パイロットテスト	-	-	-	パイロットテストの概要
2	[3]	ICD-11 Pilot testing - The Australian Experience	Jenny Hargreaves, Brooke Macpherson, Miriam Lum On	オーストラリア	パイロットテスト	Unwritten	11名 (診療情報管理士)	診断用語420問	発表時は実施中
3	[13]	Pretest for the ICD-11 Filed trial in Japan	H Ogata, Y Sato, N Tomita, K Mori, H Mizushima	日本	独自プロトコールによる	2016年2月	7名 (医師4名, 診療情報管理士3名)	ケースサマリー36問	フィールドトライアルに向けたプレテストの位置づけとして実施
4	[4]	ICD-11 Field Testing: Australian Experiences	Brooke Macpherson, James Eynstone-Hinkins, Jenny Hargreaves, James Kattre1 and Sue Walker	オーストラリア	パイロットテスト	2016年8月～2017年2月	11名 (診療情報管理士)	診断用語420問	No.2の実施報告
5	[6]	Phase 2 Morbidity Field Testing of ICD-11 MMS in Australia	Brooke Macpherson, James Kattre and Jenny Hargreaves	オーストラリア, ニュージーランド	フィールドトライアル	2017年4月～2017年6月	49名 (政府関係者, 公共および民間の病院部門, 診療情報専門団体, 関連大学学生)	診断用語298問	
6	[5]	Pilot testing ICD-11: Reasons for disagreement in code allocation using the results from two field trial centres	Olair Stenius, Brooke Macpherson, Maria Boreklev, James Kattre	オーストラリア, ニュージーランド	パイロットテスト	2016年8月	25名 (医師, 診療情報管理士)	診断用語308問	
7	[8]	ICD-11 Field Trials: A Canadian Perspective on Post-Coordination	Denise Cullen, Jilian Speckeen, Lisa Burrelle, Jin Wang, Tobi Henderson	カナダ	フィールドトライアル	Unwritten	8名 (診療情報管理士)	Trial1: 診断用語88問 Trial2: 診断用語298問	
8	[12]	ICD-11 Field Trials in the English-speaking Caribbean	A Hinds, S Quesnel-Crooks, J Martin	ジャマイカ, パナマ, バルバドス, ドミニカ, セントルシア, セントビンセント・グレナディーンズ, トリニダード・トバゴ	フィールドトライアル	Unwritten	Study1: 22名 (診療情報管理士) Study2: 12名 (診療情報管理士)		
9	[16]	Field Test of ICD-11 in Japan	H Mizushima, H Ogata, Y Sato, K Mori	日本	フィールドトライアル	2017年8月～	298名 (診療情報管理士)	診断用語298問 ケースサマリー30問	発表時は実施中
10	[10]	Pre-test for developing model of ICD-11 field trial in Korea	Korean Medical Record Association	大韓民国	独自プロトコールによる	2017年6月11日～20日	16名 (診療情報管理士)	参加大病院の医療記録 (418名分)	フィールドトライアルに向けたプレテストの位置づけとして実施
11	[11]	2017 ICD-11 MMS Field Trial in Korea	Baek Seol Kyung, Hong Joon Hyun, Boo Yoo Kyung, Ham Seung Woo, Yang Ji Hyun, Baek Se Ryon, Kim Na Young, Huh Bo Young, Lee Jong Ran, Han Gyu Cheol, Hong Soon Cheol	大韓民国	パイロットテスト	2016年6月～2017年2月	Unwritten	Unwritten	
11	[11]	2017 ICD-11 MMS Field Trial in Korea	Baek Seol Kyung, Hong Joon Hyun, Boo Yoo Kyung, Ham Seung Woo, Yang Ji Hyun, Baek Se Ryon, Kim Na Young, Huh Bo Young, Lee Jong Ran, Han Gyu Cheol, Hong Soon Cheol	大韓民国	フィールドトライアル	2017年7月	12名 (医師, 診療情報管理士)	診断用語298問 ケースサマリー30問 伝統医学用語40問	発表時は実施中
12	[9]	Main Condition Coding Agreement: Comparing ICD-10 and ICD-11 Codes from Inpatient Records	Chelsea Doktorchik, Danielle A. Southern, Cathy A. Eastwood, William A. Ghali, Hude Quan	カナダ	独自プロトコールによる	Unwritten	6名 (診療情報管理士)	入院サマリー730件の主病名	
13	[7]	ICD-11 MMS Filed Testing for Morbidity Phase 3 Preliminary Findings from Australia	James Kattre, Brooke Macpherson, Jenny Hargreaves	オーストラリア, ニュージーランド	フィールドトライアル	2018年3月	20名 (Unwritten)	診断用語80問	ICD-11コーディングツール改良後, WHO, ICFの最終段階の位置づけとして実施
14	[17]	ICD-11 Filed trial in Japan -An evaluation of a line coding -	Y Sato, H Mizushima, H Ogata, K Mori	日本	フィールドトライアル	2017年8月～2017年9月	298名 (診療情報管理士)	診断用語298問 ケースサマリー30問	

3. 大韓民国におけるICD-11フィールドトライアル

大韓民国は、我が国と同じく、フィールドテストに先立って独自のプロトコールによるプレテストを実施し、16名の評価者を対象に参加病院の418名分の患者医療記録を用いた評価を行っている[10]。ICD-11では解剖学的なコーディングが困難であったことを報告しているが、これはポストコーディングの教育や周知が不足していたことが原因と考えられた。2016年に行ったパイロットテストでは、全体評価において、その他のコード（Y）や不特定コード（Z）分類が不十分な点、同義語検索の改善、コーディング教育や訓練の充実の必要性などの意見が挙げられたことを報告している[11]。フィールドトライアルは2017年に12名の評価者を対象に実施されている[11]。

4. カリブ海諸国におけるICD-11フィールドトライアル

カリブ海諸国7地域（ジャマイカ、バハマ、バルバドス、ドミニカ、セントルシア、セントビンセント・グレナディンズ、トリニダード・トバゴ）において、10年以上の経験をもつ診療情報管理士22名によるフィールドトライアルが行われた[12]。フィールドトライアルの事前教育としてバーチャルセッションを行っており、便利かつ経済的なので有用であったという一方で、ICD-11導入時の教育では直接の対面教育（face to face）との併用が望まれると述べている。

IV. 我が国における ICD-11 フィールドトライアル実施の経緯と概要

我が国のICD-11フィールドトライアルはWHO-FIC協力センターである国立保健医療科学院がFTCセンターとなり、筆者を含む国立保健医療科学院の研究者で構成された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）「国際統計分類ファミリーに属する統計分類の改善や有用性の向上に資する研究」において実施された。研究班の初年度（平成27年度）は予備的シミュレーションとしてプレテスト（予備プレテスト）を行った。平成28年度にWHOから要請された少人数でのパイロットテストを実施し、平成29年度に本番のフィールドトライアルを実施した。（表2）

1. 予備的シミュレーションとしてのプレテスト（予備プレテスト）（平成28年2月実施）

平成28年2月に3医療機関7名（医師4名、診療情報管理士3名）を対象にICD-11フィールドトライアルの予備的シミュレーションとしてのプレテストを実施した。3医療機関の特性に応じて、生活習慣病、がん疾患、小児疾患の評価をそれぞれに依頼した[13,14]。この時点ではWHOからICD-11ブラウザ（β版）、ICD-Fitやコーディング用ケースは提供されていなかったため、ドラフト版のICD-11フィールドトライアルガイドラインを参考にして実施手順を決定し、質問紙形式で行った。そのため、II章で示したICD-11フィールドトライアルのプロトコールとは異なる手順での実施となった。ケースサマリーは「ICDコーディングトレーニング第2版（医学書院）」から生活習慣病、がん疾患、小児疾患の12問ずつをそれぞれランダムに選出し転載許諾を得たのち使用した。評価者にはICD-11β版やリファレンスガイドを電子ファイル（英語）で提供した。評価者は12問のブリッジコーディング（1つのケースサマリーに対してICD-10とICD-11でコーディングを行う）と信頼性評価（1つのケースサマリーに対して同一医療機関の評価者全員がICD-11でコーディングを行う）、ICD-11の総合評価（網羅性、詳細度など）、ICD-11改訂に関する意見を得た。ブリッジコーディングではICD-11コーディングが難しかったかどうか、使いやすさ（非常に使いやすい、使いやすい、ふつう、使いにくい、非常に使いにくい）、コードの自信（とても低い、低い、ふつう、高い、とても高い）についての項目を設定した。

ブリッジコーディングの評価として合計84件が得られた。難しかったと答えた割合は44.0%、使いやすさについて「非常に使いやすい、使いやすい」と答えたのが25.0%だった。コーディングの自信については「高い、とても高い」と答えたのは11.9%にとどまった。コード体系の理解不足、英語での実施、コード検索の煩雑さが評価理由として挙げられた。総合評価では分類が詳細すぎることが指摘され、詳細情報がカルテから読み取れず、かえって詳細不明が多くコーディングされる恐れがあることや、臨床適用においては十分な議論が必要であることが示唆された。また詳細度は分野（章）ごとの差が大きいことが予想されるため、ICD-11の評価は分野ごとに行うべきではないかという意見も挙げられた。本予備プレテストは予備的シミュレーションとしての位置づけ

表2 我が国におけるICD-11フィールドトライアル

	実施期間	対象者	コーディング問題
1) シミュレーションとしてのプレテスト（予備プレテスト）	H27年度（H28.2）	診療情報管理士、医師7名	既存の問題集から選出し3領域（生活習慣病、がん疾患、小児疾患）各12ケース
2) パイロットテスト	H28年度（H28.12～H29.1）	診療情報管理士7名	WHO提供の診断用語308件
3) ICD-11フィールドトライアル（ICD-11 FT）	H29年度（H29.8～9）	診療情報管理士319名	WHO提供の診断用語298件+ケースシナリオ30件

となったが、このように重要な指摘を得ることができた。反省点としてはICD-10での評価項目の設定を行わなかったため、ICD-10との比較ができなかったことである。

2. パイロットテスト(平成28年12月～平成29年1月実施)

平成28年度にWHOからICD-Fitのデモバージョンがリリースされた。本番のフィールドトライアルを控え、WHOより評価者5名以上によるパイロットテストの実施要請を受け、平成28年12月から平成29年1月に実施した[15]。評価者は日本診療管理学会所属会員を中心に、診療情報管理士指導者、ICD-11の翻訳作業を担当したもの、教育関係者、病院経験を有するもの、計7名を対象とした。コーディングツール、ICD-11ブラウザ(β版)、リファレンスガイドはICD-Fitからアクセスできる状態であった。評価者間のばらつきが少なくなるよう事前に共通認識を持つための機会を設けた。また作業開始後にも意見交換を行えるようメーリングリストを作成した。ICD-Fitの一部は日本語で情報提供したが、コーディング作業は英語環境下で行われた。ケースは308件の診断用語が提供され、感染症、血液学、免疫、循環器、内分泌、周産期など広範囲にわたる分野が網羅されていた。本プレテストではケースサマリーは提供されなかった。評価項目は個別評価(図2)と全体評価(図4)が設定された。結果の解析では、個別評価、全体評価の集計に加え、同一ケースについて評価者のコーディング結果にばらつきが生じたかどうかを指標とした解析を行った。

評価者7名全員が308症例をコーディングし、個別評価2,156件、全体評価7件を得た。「コーディングは難しかった」と答えた割合はICD-11で11.3%、ICD-10で13.9%、「コードの当てはまり(詳細さ)はちょうどいい」と答えた割合はICD-11で87.7%、ICD-10で83.3%、「コーディングの際にあいまいさを感じなかった」はICD-11で74.7%、ICD-10で68.9%となった。またコーディング時間はICD-11が有意に短くなっていた(ウィルコクソンの順位和検定, $p=0.017$)。上記の評価項目では差は小さいもののICD-11のほうが高い評価を示した。全体評価ではICD-11の適用範囲がよいと答えたのが71.4%、詳細度が「ちょうどいい」と答えたのが85.7%と高い評価であった一方で、使いやすさは「非常に使いやすい」「使いにくい」が0%、「中程度」が71.4%、「使いにくい」が28.6%とやや低い評価となった。使いやすさの評価の理由を自由記載で尋ねたところ、コーディングツールの検索システムの不備として、検索ワードが少ない、複数語検索ができないなどの意見が挙げられた。

パイロットテストでは各ケースのゴールドスタンダード(GS)が提供されなかったため、正解率の解析が不可能であった。そこで、一つのケースに対して評価者が付与したコードのばらつきを評価した。1つのコードに収束した割合はICD-10で46.8%、ICD-11で14.6%となり、ICD-11ではコードのばらつきが生じやすいことが示唆

された。またICD-10でコードがばらついていたがICD-11でコードが1つに収束した21症例は血液学、免疫、周産期の分野が有意に多かった(カイ二乗検定, $p<0.001$)。一方、ICD-10でコードが1つに収束していたがICD-11ではばらついていた症例は133あり、内分泌、眼科、消化器、皮膚分野に偏っていた(カイ二乗検定, $p<0.001$)。これらの結果は、ICD-11は章によって特徴が異なることを示唆しており、予備プレテストでも指摘された通り、章や分野ごとの評価が重要であることを裏付けるものとなった。

プレテストでは本番のフィールドトライアルと同じウェブプラットフォーム(ICD-11 Fit)、コーディングツール、ICD-11βweb版、リファレンスガイドが提供され、ICD-10との比較項目も追加された。ICD-11の優位性が示唆され、コーディングの際のばらつきや曖昧さを減らす可能性が示された。一方で、コーディングツールなどの電子ツールのシステムの精度や利用者側の習熟について課題が残った。

3. 本番フィールドトライアル(平成29年8～9月実施)

平成29年8月から9月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に参加を募り、合計378名がICD-Fitに登録された。評価者に診断用語298件、ケースサマリー30件が割り当てられた[16,17]。フィールドトライアルに関する資料やICD-Fitの一部は日本語で情報提供したが、コーディング作業は英語環境下で行われた。評価項目は個別評価(図2,3)と全体評価(図4)が設定された。本フィールドトライアルの結果解析は現在進行中である。以下に現時点での解析結果および今後の展望をまとめる。

評価者378名のうち、診断用語コーディング全症例を終えたのは83名(22.0%)、ケースサマリーコーディング全症例を終えたのは109名(28.8%)だった。診断用語コーディングの個別評価で得られた回答38,654件の集計では、「コーディングは難しかった」と答えた割合はICD-11で7.7%、ICD-10で3.1%、「コードの当てはまり(詳細さ)はちょうどいい」と答えた割合はICD-11で68.9%、ICD-10で71.8%、「コーディングの際にあいまいさを感じなかった」はICD-11で65.0%、ICD-10で70.9%となった。またコーディング時間はICD-11が有意に長くなっていた(ウィルコクソンの順位和検定, $p<0.001$)。以上から、パイロットテストの結果に反し、全体的にICD-10が優位な結果となった[16]。これは診断用語の難易度やエクステンションコードの有無などが関与していることが示唆される。エクステンションコードとはICD-11で新たに加えられた章であり、重症度、時間軸情報、部位の別など詳細な補足情報を追加できるコードである。パイロットテストで提供されたのは短い診断用語であり、その語句をコーディングツールの検索欄にそのままコピーアンドペーストをすれば容易にコードを検索できるケースが多かった。しかし、フィールドテストでは修

飾語を伴う診断用語が多く出題され、ポストコーディングが必要なものも多くなっていた。コーディングツールの曖昧検索や複数語検索機能が十分でなかった点に加え、単語を区切って検索する必要がある場合にどの単語を検索語として使用するかについて言語学的な問題が大きく影響した可能性も考えられた。

診断用語コーディングの個別評価をICD-11の章ごとで解析してみたところ、難しかったと答えた割合はほとんどの章でICD-11が高かった。比率のオッズ比を求めたところ、周産期、血液、腎尿路生殖器、神経、妊娠、損傷分野で特に大きな差がみられ、これらの章は特にICD-11適用・導入においてコーディングルールの整備などを慎重に行う必要があることが示唆された[17]。章ごとの解析はさらに詳細な検討を行う予定である。

また本フィールドトライアルでは診断用語の一部（19件）のGSがWHOから提供された。ICD-10とICD-11でGSとの一致率を比較したところ、ICD-11一致率のほうが高かったものが6件、ICD-10一致率のほうが高かったものが7件、差がみられなかったものが6件だった。ICD-11一致率のほうが高かったものの症例の特徴を考察したところ、分類粒度の粗いICD-10では複数の解釈に誘導されるような曖昧さがあったために評価者の判断によって複数のコードに分散してしまっていた病態をICD-11ではより正確に反映でき、コードの分散も抑えられる可能性が示唆された。一方、ICD-10の一致率のほうが高かった症例では、エクステンションコードや複数コードが必要なケースが多く見られた。GS一致率が低いケースでは「難しかった」と答えた割合が高くなっており、ポストコーディングが必要になる場合において困難さを感じた評価者が多かったことが示唆された[18]。WHOからの追加のGS提供を要望するとともに、エクステンションコードや複数コードが必要になるケースの詳細な評価が望まれる。

診断用語コーディングにおける全体評価では予備プレテストで指摘されていた点と同じく、分類が詳細すぎるため、必要なコードを探すことが困難になる可能性や、コーディングするための疾患情報が不足する事象が多くなる可能性への懸念が挙げられた。また、パイロットテストと同じく、ICD-11電子ツールの使いやすさの評価では検索システムの改善などが指摘された。

V. ICD-11 フィールドトライアルの展望

1. ICD-11導入に向けた我が国独自のフィールドトライアルに向けて

我が国における予備プレテストを含むフィールドトライアルはすべて原則英語環境下で実施された。この点については予備プレテストの段階から「英語での評価が困難である」「日本語で評価をしたい」との意見が非常に多く挙げられた。現在、厚生労働省の社会保障審議会統計分科会疾病、障害および死因分類部会を中心に、ICD-

11の日本語翻訳作業のロードマップが検討されている。

ICD-11は臨床や研究など多様な実態に即した表現に適しており、ICD-10よりも精緻な疾患概念の記述が可能となる。一方で詳細になりすぎるが故の問題点も指摘されている。つまり、コードが詳細になりすぎること、かえってコードを検索する手間がかかることや、診療体制によってどれくらい詳細にコードする必要があるのかの共通認識がないと、運用において混乱を生じる可能性があることが挙げられる。前者においては、基本的にはコーディングツールの検索システムの精度向上でカバーされるものであるが、英語と日本語の言語学的な違いを考えると現在WHOから提供しているコーディングツールをそのまま日本語に訳すだけでは、複数語検索や曖昧語検索に問題が生じることが予想される。後者については、カルテなどから詳細情報が読み取られなければ、結局、詳細不明コードが増えることになり、急性期医療、慢性期医療など診療体制に応じて、どこまで詳細にコードすべきなのかについて共通認識を定めることが重要であることは言うまでもなく、詳細度は章ごとによっても異なるため、章ごと疾患ごと詳細な検討が重要となる。

今後、日本でのICD-11導入に向け、我が国の医療・診療体制に即した独自のプロトコールによるフィールドトライアルが必要となるだろう。フィールドトライアル後もコード変更が加わっているため、追評価は必須である。すでに一部の国ではその国独自のプロトコールによるフィールドトライアルが行われており、我が国でも、日本語翻訳作業と並行し、同様の検討を始めることが望ましいと考えられる。

2. 改定後フィードテストへの期待

ICD-11フィールドトライアルはWHOが主導となり、全世界で共通のプロトコールおよびプラットフォームで実施された。そ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や、経時的な比較解析が期待された。しかし、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なかったため、各国の報告体系にばらつきが生じている。GSの解釈の基準（GSの一致率を5段階で評価する基準）も一部で示されているものの[2]、本稿執筆時点ではコーディング対象となったすべての診断用語やケースサマリーのGSの提供には至っていない。また、WHOがフィールドトライアルの結果をどのような方法で解析及び解釈をし、ICD-11改訂に反映させたかというプロセスも本稿執筆時点では公表されていないため報告が待たれるところである。フィールドトライアルで用いられたプラットフォームであるICD-Fitは、今後は主に教育ツールとして利用されることとなるが、今回の改定前フィールドトライアルで得られた知見や課題を活かし、ICD-11改定後の改訂（リビジョン）、一部改正（アップデート）のプロセスにおける利活用が望まれる。

VI. 結語

ICD-11は2019年に公開される。その後、フィールドトライアルで得られた知見をもとに、各国でのスムーズな導入・運用に向けた、より具体的な課題や指針を明確にするための議論が活発になるであろう。我が国においてもICD-11をはじめとした関連資料の日本語翻訳作業の遂行とともに、我が国独自の医療および診療体制におけるICD-11の円滑な導入・運用に向けた検討が行われることとなる。同時に、我が国独自のプロトコールによるICD-11フィールドトライアルの追加実施の検討が望まれる。

参考文献

- [1] 緒方裕光, 研究代表者. 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)「国際統計分類ファミリーに属する統計分類の改善や有用性の向上に資する研究」(H27-政策-一般-005)平成27年度総括・分担研究報告書. 2015. Ogata H, kenkyu daihyosha. Research on Policy Planning and Evaluation. Health, Labour and Welfare Policy Research Grants. [Kokusai tokei bunrui family ni zokusuru tokei bunrui no kaizen ya yuyosei no jojo ni shisuru kenkyu. (H27-seisaku-ippan-005) Heisei 27 nendo sokatsu / buntan kenkyu hokokusho.] 2015. (in Japanese)
- [2] Kostanjsek N, Jakob R, Celik C, Schmider A, Moskal L, Best L. Line coding pilot testing ICD-11 MMS. WHO-FIC Network Annual Meeting; 2016.10.8-12; Tokyo. Poster. C302.
- [3] Hargreaves J, Macpherson B, On ML. ICD-11 Pilot testing: The Australian Experience. WHO-FIC Network Annual Meeting; 2016.10.8-12; Tokyo. Poster. C304.
- [4] Macpherson B, Eynstone-Hinkins J, Hargreaves J, Kattel J, Walker S. ICD-11 Field Testing: Australian Experiences. WHO-FIC Network Annual Meeting; 2017.10.16-21; Mexico. Poster. C303.
- [5] Steinum O, Macpherson B, Boreklev M, Katte J. Pilot testing ICD-11: Reasons for disagreement in code allocation using the results from two field trial centres. WHO-FIC Network Annual Meeting; 2017.10.16-21; Mexico. Poster. C305.
- [6] Macpherson B, Katte J, Hargreaves J. Phase 2 Morbidity Field Testing of ICD-11 MMS in Australia. WHO-FIC Network Annual Meeting; 2017.10.16-21; Mexico. Poster. C304.
- [7] Katte J, Macpherson B, Hargreaves J. ICD-11 MMS Filed Testing for Morbidity Phase 3 Preliminary Findings from Australia. WHO-FIC Network Annual Meeting; 2018.10.22-27; Korea. Poster. C309.
- [8] Cullen D, Speckeen J, Burelle L, Wang J, Henderson T. ICD-11 Field Trials: A Canadian Perspective on Post-Coordination. WHO-FIC Network Annual Meeting; 2017.10.16-21; Mexico. Poster. C306.
- [9] Doktorchik C, Southern DA, Eastwood CA, William A, Quan GH. Main Condition Coding Agreement: Comparing ICD-10 and ICD-11 Codes from Inpatient Records. WHO-FIC Network Annual Meeting; 2018.10.22-27; Korea. Poster. C304.
- [10] Korean Medical Record Association. Pre-test for developing model of ICD-11 field trial in Korea. WHO-FIC Network Annual Meeting; 2017.10.16-21; Mexico. Poster. C30.
- [11] Kyung BS, Hyun HJ, Kyung BY, Woo HS, Hyun YJ, Ryon BS, et al. 2017 ICD-11 MMS Field Trial in Korea. WHO-FIC Network Annual Meeting; 2017.10.16-21; Mexico. Poster. C310.
- [12] Hinds A, Quesnel-Crooks S, Martin J. ICD-11 Field Trials in the English-speaking Caribbean. WHO-FIC Network Annual Meeting; 2017.10.16-21; Mexico. Poster. C307.
- [13] Ogata H, Sato Y, Tomita N, Mori K, Mizushima H. Pretest for the ICD-11 Filed trial in Japan. WHO-FIC Network Annual Meeting; 2016.10.8-12; Tokyo. Poster. C305.
- [14] 緒方裕光, 研究代表者. 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)「国際統計分類ファミリーに属する統計分類の改善や有用性の向上に資する研究」(H27-政策-一般-005)平成28年度総括・分担研究報告書. 2016. Ogata H, kenkyu daihyosha. Research on Policy Planning and Evaluation. Health, Labour and Welfare Policy Research Grants. [Kokusai tokei bunrui family ni zokusuru tokei bunrui no kaizen ya yuyosei no jojo ni shisuru kenkyu. (H27-seisaku-ippan-005) Heisei 28 nendo sokatsu / buntan kenkyu hokokusho.] 2016. (in Japanese)
- [15] 水島洋, 研究代表者. 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)「国際統計分類ファミリーに属する統計分類の改善や有用性の向上に資する研究」(H29-政策-一般-001)平成29年度総括・分担研究報告書. 2018. Mizushima H, kenkyu daihyosha. Research on Policy Planning and Evaluation. Health, Labour and Welfare Policy Research Grants. [Kokusai tokei bunrui family ni zokusuru tokei bunrui no kaizen ya yuyosei no jojo ni shisuru kenkyu. (H29-seisaku-ippan-005) Heisei 29 nendo sokatsu / buntan kenkyu hokokusho.] 2018. (in Japanese)
- [16] Mizushima H, Ogata H, Sato Y, Mori K. Field Test of

ICD-11 フィールドトライアルについて

- ICD-11 in Japan. WHO-FIC Network Annual Meeting; 2017.10.16-21; Mexico. Poster. C308.
- [17] Sato Y, Mizushima H, Ogata H, Mori K. ICD-11 Filed trial in Japan: An evaluation of a line coding. WHO-FIC Network Annual Meeting; 2018.10.22-27; Korea. Poster. C3100.
- [18] 佐藤洋子, 水島洋, 木村映善, 西大明美, 緒方裕光. 我が国におけるICD-11フィールドトライアルー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ー. 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):800-802.
- Sato Y, Mizushima H, Kimura E, Nishio A, Ogata H. [Waga kuni ni okeru ICD-11 field trial: shindan yogo coding ni okeru gold standard no kaiseki.] 38th Japan Association for Medical Informatics. (19th Nihon Iryo Joho Gakkai Gakujutsu Taikai); 2018.11.22-25; Fukuoka. An Official Journal of the Japan Association for Medical Informatics. 2018;38(suppl):800-802.