

一般口演

一般口演22

標準化・EHR・PHR

2018年11月25日(日) 13:40 ~ 15:40 B会場 (4F 409+410)

**[4-B-3-3] 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**○佐藤 洋子^{1,2}, 水島 洋², 緒方 裕光^{2,3} (1.防衛医科大学校, 2.国立保健医療科学院, 3.女子栄養大学)

ICD-11(国際統計分類第11版)フィールドトライアル(FT)では診断用語のコーディングであるラインコーディング(LC)とケースシナリオのコーディングであるケースコーディング(CC)が評価される。ICD-11FTのためのガイドラインとプラットフォーム(ICD-11Fit)がWHOより提供されたことを受け、2017年8月から9月にICD-11FTを実施した。日本健康情報管理学会、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に登録された。LCでは診断用語298件が提供され、評価者はケースごとにICD-10とICD-11でダブルコーディングしたのち、難易度、コードの詳細度、コードの曖昧さについて評価した。コーディング時間は自動記録された。WHOよりゴールドスタンダード(GS, 標準コード)が提供された19件において、GS一致率、GSのコード数、評価項目、コーディング時間の解析を行った。ICD-11コードがGSと一致しているかどうかに影響している因子として、コードの数(Odds: 0.03, 95% Confidence interval: 0.03 - 0.04)、ICD-10コードの一致率(1.55, 1.22 - 1.96)、難易度(0.43, 0.28 - 0.65)、詳細度(1.91, 1.49 - 2.45)が同定された。ICD-10コードの一致率との関係についてはケースによる違いが大きく、「MRSA敗血症(1B46&MJ81.11, A41.0&U82.1)」ではICD-11が高く(31.2% vs 11.9%, $p < 0.001$)、「尿路感染による敗血症(GS58.Z&MJ75.1, A41.9&N39.0)」ではICD-10が高かった(1.7% vs 16.8%, $p < 0.001$)。詳細な質的評価を行ったうえでのICD-11の臨床現場での適用性評価が必要であると考えられる。

我が国における 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 -

佐藤洋子^{*1,2}、水島洋^{*2}、木村映善^{*2}、西大明美^{*3}、緒方裕光^{*4,2}

*1 防衛医科大学校、*2 国立保健医療科学院、*3 東京医療保健大学、*4 女子栄養大学

ICD-11 Field Trial in Japan

- Analysis of Gold standard in diagnostic term coding -

Yoko Sato ^{*1,2}, Hiroshi Mizushima ^{*2}, Eizen Kimura ^{*2}, Akemi Nishio ^{*3}, Hiromitsu Ogata ^{*2,4}

*1 National Defense Medical College, *2 National Institute of Public Health,

*3 Tokyo Healthcare University, *4 Kagawa Nutrition University

In the ICD-11 (International Disease Classification 11th Edition) Field Trial (FT), two types of evaluations are performed: a line coding (LC) that is the coding of diagnostic terms and a case coding (CC) that is the coding of case scenarios. Following a guideline and platform for ICD-11FT (ICD-11 Fit) offered by WHO, ICD-11FT in Japan was implemented from Aug to Sept 2017. With the cooperation of the Japan Health Information Manager Association and Japan Hospital Association, 378 health information managers were registered in ICD - 11 Fit. In LC, 298 diagnostic terms were provided. The raters coded each case with both ICD-10 and ICD-11 then the difficulty of coding, the specificity of the code and the ambiguity of the code were evaluated. We analyzed 19 cases that gold standard codes were provided from WHO. Firstly, we found that the number of codes (Odds: 0.03, 95% Confidence Interval: 0.03 - 0.04), a concordance rate of ICD-10 (1.55, 1.22 - 1.96), a difficulty (0.43, 0.28 - 0.65), and a specificity (1.91, 1.49 - 2.45) were factors influencing the concordance rate of ICD-11. With an association with the concordance rate of ICD-11 and ICD-10, there was a difference depending on the case. For example, the rate of ICD-11 was higher in "Methicillin-Resistant Staphylococcus Aureus Septicemia (1B46 & MJ81.11, A41.0 & U82.1)" (31.2% vs 11.9%, $p < 0.001$), while "Sepsis due to urinary tract infection (GS 58. Z & MJ 75.1, A 41.9 & N 39.0)" was higher in ICD-10 (1.7% vs 16.8%, $p < 0.001$). For evaluating the applicability of ICD-11 in the clinical setting, detailed evaluations are required for each case.

Keywords: ICD-11, ICD-10, Field trial

1. 緒論

2018年6月に開発の最終段階に達したICD-11(国際疾病分類第11版)が加盟国に対して公開された。ICD-11は死亡・疾病統計だけでなく臨床現場や研究など様々な場面での利活用に向け、より多様な病態を表現できるようコード体系が整備され、免疫系疾患や睡眠・覚醒障害など新しい章が追加された。また重症度や部位指定など診断の詳細を補足するためのエクステンションコードの章も加わっている。さらに ICD-11 は電子ツールでの利用が想定されており、ウェブサイト上での分類情報の提供や、用語を入力してコードを検索できるコーディングツールなどが WHO から提供される。

ICD-10(国際疾病分類第10版)からICD-11の改訂におけるフィールドトライアルではICD-11の適用性、信頼性、有用性などを検討したうえで、抽出された課題について解決・検討することが求められている⁽¹⁾。フィールドトライアルは実際の医療現場でICDコーディングを行う専門家による評価であり、改訂が統計データに与える影響を検討する際に有用な情報となるだけでなく、改訂そのものの合理性を高めるためにも必要不可欠である⁽²⁾⁽³⁾⁽⁴⁾。国際的に共通のフィールドトライアルを実施するためにWHOからガイドライン(ICD-11リファレンスガイド)やウェブベースのプラットフォーム(ICD-Fit)、および診断用語とケースシナリオが提供された。

我が国のICD-11フィールドトライアルは2015年度、2016年度のプレテストを経たのち、2017年度に実施された。フィールドトライアルで得られる結果は我が国におけるICD-11の導入・実装における課題抽出や提言などへの利活用が期待さ

れている。

2. 目的

2017年度に我が国で実施されたICD-11フィールドトライアルでは診断用語のコーディングであるラインコーディング(診断用語コーディング)とケースシナリオのコーディングであるケースコーディングが評価された。本研究では、ラインコーディングの評価結果のうち、WHOから標準コードとしてのゴールドスタンダード(GS)が提供されたものについて、ICD-11とICD-10の一致率や評価結果を比較し、考察を加えることを目的とした。

3. 方法

我が国のフィールドトライアルは2017年8月から9月にかけて実施され、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名が評価者として参加した。評価者はフィールドトライアル用のWebサイトであるICD-Fitにログインし回答を行った。ICD-11リファレンスガイドやICD-Fitの一部は日本語翻訳を提供したが、コーディング作業は英語環境下で行われた。ラインコーディングでは診断用語298件が提供され、評価者はケースごとにICD-10とICD-11のそれぞれでコーディングしたのち、難易度(「難しかった」「難しくなかった」)、コードの詳細度(「ちょうどいい」「詳細すぎる/大略である」)、コードの曖昧さ(「曖昧である」「曖昧でない」)について評価した。コーディング時間は自動記録された。

298 件の診断用語中、WHO より GS が提供された診断用語 19 件において、GS 一致率、GS コード数、評価項目、コーディング時間の集計を行った。比率の比較はマクマナー検定、コーディング時間の比較は Wilcoxon 符号順位検定を行った。有意水準は $p < 0.05$ とした。ICD-11 の GS と一致したかどうかを目的変数とし、コードの数、ICD-10 が GS と一致したかどうか、難易度、詳細度、曖昧さ、コーディング時間を説明変数としたロジスティック回帰分析を変数減少法 ($p < 0.05$) で実施した。なお、今回 WHO から提供された GS で提示されている ICD-11 コードは 2017 年 8 月時点のバージョンのものであり、改訂が進んだ結果、現在公表されているコードと一致しないものもある。

4. 結果

WHO より GS が提供された診断用語 19 件の全回答数は 3,959 件だった。19 件の分野内訳は感染症分野(第1章)が 14 件、新生物分野(第2章)が 5 件となった。感染症分野の内訳は敗血症分野が 6 件、HIV 分野が 5 件、下痢分野が 3 件だった。ICD-11 の GS 一致率と ICD-10 と GS 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 6 件のうち Cytomegalovirus Colitis (サイトメガロウイルス性大腸炎) の ICD-10、ICD-11 の GS はそれぞれ 1A24、B25.8 だった。GS 一致率は ICD-11 が 92.9%、ICD-10 が 37.1% であった ($p < 0.001$)。ICD-11 ではサイトメガロウイルス性大腸炎として 1 つのコード 1A24 が付与されるようになっている。一方、ICD-10 の GS として設定されていた B25.8 はサイトメガロウイルス性腸炎だけでなく、脊髄炎、脳炎、網膜炎など様々なサイトメガロウイルスによる炎症病態を表すコードである。ほかの回答コードとして A08.3「その他のウイルス性腸炎」が 15.2% を占めていた。評価者が付与したコードの種類(コードのばらつき)は ICD-11 では 9 種類、ICD-10 では 12 種類だった。このことから、ICD-11 は、ICD-10 では分類の粒度が荒く、複数の解釈に誘導されるような曖昧さがある分類形式で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より正確に反映でき、コードの分散も抑えられることが示唆された。

一方、ICD-10 の GS 一致率のほうが高かった 7 件のうち、ICD-11 においてエクステンションコードや複数コードが必要なケースの GS 一致率はすべて 5% 未満だった。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。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、複数コードやエクステンションコードの付与につ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。

しかし複数コードに関してはさらなる詳細な検討が必要である。例えば同じ敗血症でも Methicillin-Resistant Staphylococcus Aureus Septicemia (MRSA 敗血症) では ICD-11GS 一致率のほうが高く(31.2% vs 11.9%, $p < 0.001$)、Sepsis due to urinary tract infection(尿路感染による敗血症) では ICD-10GS 一致率のほうが高かった(1.7% vs 16.8%, $p < 0.001$)。両者 ICD-11、ICD-10 とともに 2 つのコードで表されるものであるが、逆の結果となり、複数コードが必要であるからといって必ずしも ICD-11 の一致率が下がるとは限らない。後者の ICD-11 のコード選択では 2 つのコードのうちメインコードのみを選

択した割合が 63.0% と高かったため、複数コードに関する詳細な検討を行うことで、ICD-11 の有用性をより高めることができると考えられた。

GS が提供された症例の全回答数 3,959 件において、ICD-11 の GS と一致したかどうかによる影響を調べるため、ICD-11GS コードの数、ICD-10 が GS と一致したかどうか、難易度、詳細度、曖昧さ、コーディング時間を説明変数として評価した。その結果、ICD-11 コードが GS と一致することに影響を与える因子として、GS コードの数が少ないこと(Odds: 0.03, 95% Confidence interval: 0.03 - 0.04)、ICD-10 コードの一致率の高いこと(1.55, 1.22 - 1.96)、難しくなかったと答えたこと(2.33, 1.53 - 3.53)、詳細度がちょうどいいと答えたこと (1.91, 1.49 - 2.45) が同定された。

5. 考察

2017 年度に我が国で実施された ICD-11 フィールドトライアルにおいて、診断用語のコーディングであるラインコーディングの解析を行った。評価者が付与したコードと WHO が標準コードと定めた GS コードの一致率などを比較することで、ICD-11 の適用性や有用性の評価を試みた。ICD-10 でコードした際と ICD-11 でコードした際の GS 一致率の比較において、ICD-10 では的確なコードが存在しないため表現することが困難で GS 一致率が低かった病態に対し、ICD-11 ではより多くの評価者が GS コードを選択できていた例があった。このような例では評価者が付与したコードのばらつきも ICD-11 の方が小さくなっていった。ICD-11 では基本的に病態とコードが 1 対 1 の関係になるようなコード体系となっており、実際にコードの数も ICD-10 の約 14,000 から ICD-11 は約 32,000 個と増えている。以上のことから 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、有用性が高いことが示唆された。

しかし GS にエクステンションコードや複数コードが含まれるケースでは ICD-11 の GS 一致率が低くなっていった。コードのばらつきが ICD-11 のほうが大きくなっていったものも多かった。ICD-11 で GS と一致するかどうかによる影響を与える因子としては GS コードの数が少ないことが同定されたことから、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評価はさらに広範に詳細に行う必要がある。

また今回のフィールドトライアルは一部の資料を除き英語環境下での実施であったため、その影響が大きかったと考えられる。短い診断用語であれば、その語句をコーディングツールの検索システムに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が、そうでない場合は単語を区切って検索する必要があり、どの単語を検索語として使用するかどうかは言語学的な問題が大きく影響することが考えられた。

6. 結論

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になり、またコーディング品質が向上することが見込まれる点において有用性が高いことが示唆された。一方で、我が国における ICD-11 導入に向けて、さらなる GS の評価、複数コードやエクステンションコードなどコーディングルールの整備、教育、そして日本語翻訳を行った上で、あらためて日常のコーディング業務に近い環境下を想定した詳細な評価が重要である。

7. 参考文献

- 1). Ellis PM, Welch G, Purdie GL, Mellsop GW. Australasian Field Trials of the Mental and Behavioural Disorders Section of the Draft ICD-10. Aust New Zeal J Psychiatry . 1990;24(3):313-21.
- 2). Donada M, Kostanjsek N, Della Mea V, Celik C, Jakob R. Piloting a Collaborative Web-Based System for Testing ICD-11. Stud Health Technol Inform. 2017 ;235:466-70.
- 3). Wamboldt M, Cordaro A, Clarke D. Parent-Child Relational Problem: Field Trial Results, Changes in DSM-5, and Proposed Changes for ICD-11. Fam Process . 2015;54(1):33-47.
- 4). Raghuram R, Shamasundar C. Icd-10 field trials in India - a report. Indian J Psychiatry. 1992;34(3):198-221.

