厚生労働行政推進調査事業費補助金（厚生労働科学特別研究事業）
総括研究報告書

患者調査の傷病符号付けプロセスの最適化と効率化に資する研究

研究代表者　荒井　康夫　　北里大学 未来工学部データサイエンス学科 准教授

研究要旨

本研究は、統計法（平成 19 年法律第 53 号）第 2 条第 4 項に規定する基幹統計である患者調査を対象に、傷病名から ICD-10 準拠統計基準のコード（ICD-10 コード）への変換（傷病名符号化）プロセスの最適化および効率化に資する基礎的知見を得ることを目的として、令和 5 年患者調査において得られた傷病名を試料として実施した。患者調査における傷病名符号化プロセスは、辞書照合（一次工程）、形態素解析を活用した手付け符号付け（二次工程）、ならびに手付け符号付け（三次工程）の 3 工程で構成される。

本研究は、当該プロセスを前提とし、その正確性、一貫性、再現性および判定根拠の明示性を高めるための補助的検証手段ならびに基礎的知見を提供することを志向して、5 課題の分担研究を相補的に実施した。具体的には、(1)標準病名マスターを照合源とする符号照合アルゴリズム（PCMA）および照合指数（PCMI）の試作と評価、(2)AI を用いた類似度判定の試行、(3)ICD 大分類別にみた手付けによる ICD 符号付与の出現要因の分析、(4)辞書照合で符号化されなかった病名の傾向分析、(5)クラスタリングによる不一致パターン分析に基づく既存符号付けアルゴリズムの論理的改善策の整理、(6)患者調査と標準病名マスターの符号化方針相違事例における診療情報管理士による盲検下符号化の予備的検証を実施した。

主要な結果として、PCMA の主指標対象 111,984 語における第 1 位候補と患者調査の確定符号との 3 桁分類単位の一致率は 85.09％であった。また、辞書照合で処理されなかった語 75,068 語を 3 層構造に層別化することにより、用手的確認における確認対象の重点化に資する構造を示した。

本研究は、患者調査の確定符号の自動決定を主張するものではなく、AI や機械学習による自動分類を主要方法とするものでもない。本研究の成果は、患者調査における傷病符号付けの適正化、辞書ファイル更新、傷病符号付けチェックプログラムの精緻化、ならびに次期統計処理システム更新方針の検討に資するものであり、基幹統計としての継続性と信頼性を損なうことなく、符号化判断の再現性、説明可能性、ならびに確認可能性を高めるための基礎資料を提供したものである。

研究分担者

西本　寛　　JA 長野厚生連　佐久総合病院
　　総合医療情報センター長

阿南　誠　　川崎医療福祉大学
　　医療福祉マネジメント学部

桜澤　邦男　国立大学法人東北大学
　　東北大学病院メディカル IT センター
　　医療秘書学科 特任教授

准教授
來島　裕太　山口県立総合医療センター
医療企画室　室長補佐
森藤　祐史　大阪府立病院機構
大阪急性期・総合医療センター
医療情報部診療情報管理室 総括主査
石割　大範　国立健康危機管理研究機構
システム基盤整備局 係長

研究協力者

小泉　麻美　北里大学病院　医療の質評価室
診療情報管理士
村上　沙織　北里大学病院　診療情報管理室
診療情報管理士
香取　麻紀子　北里大学病院　診療情報管理室
診療情報管理士

A．研究目的

（1）患者調査の位置づけ

患者調査は、統計法（平成 19 年法律第 53 号）第 2 条第 4 項に規定する基幹統計である患者統計を作成するための調査であり、患者調査規則第 1 条および第 2 条に基づき、医療施設を利用する患者について、その傷病の状況等の実態を明らかにすることを目的として実施される。本調査は、病院および診療所を利用する患者の傷病状況を明らかにする我が国唯一の網羅的な疾病統計であり、その集計結果は、保健医療政策の立案・評価における基礎資料として用いられる。

患者調査では、医療機関から無作為抽出された施設で調査日に受療した患者の傷病情報を、世界保健機関（WHO）が定める疾病及び関連保健問題の国際統計分類第 10 版（ICD-10）に基づき集計する。傷病名から ICD-10 準拠統計基準のコード（ICD-10 コード）への変換（以下「傷病名符号化」という。）の正確性、一貫性、再現性および判定根拠の明示性は、本基幹統計の品質を担保する基盤である。

（2）患者調査における傷病名符号化プロセスの現状と課題

患者調査における傷病名符号化プロセスは、3 つの工程で構成されている。一次工程は、過去調査の蓄積を反映した患者調査にもとづく辞書（以下「傷病名辞書」という。）との照合により ICD-10 コードを付与する工程（以下「辞書照合」という。）である。二次工程は、辞書照合で符号化されなかった語のうち、形態素解析により候補が提示された語に対して手付けを行う工程（以下「形態素解析を活用した手付け」という。）である。三次工程は、形態素解析により候補が提示されなかった語に対して手付け符号化を行う工程（以下「手付け符号化」という。）である。

令和 5 年患者調査において得られた傷病名の種類数約 11 万 2 千語のうち、辞書照合により約 33%（約 3 万 7 千語）、形態素解析を活用した手付けにより約 53%（約 6 万語）が符号化された一方、約 14%（約 1 万 6 千語）については手付け符号化を要した。

患者調査における傷病名は、医療機関が調査票の傷病名欄に自由記述形式で記入する。傷病名には、表記揺れ（異字・誤字・略語・全角半角混在）、修飾語、部位、病態、重症度および複合的記載、複数傷病名の併記が含まれる。また、修飾語のみ、部位のみ、診療科名、後置語、なし表現など、傷病名として成立しない語（非傷病名）も一定数含まれる。これらの多様性および曖昧性が、辞書照合で符号化されない語の発生要因として作用している。

手付け符号化を要する語の発生要因の分析、患者調査用辞書ファイルの要件整理、および自動符号付けアルゴリズムと傷病符号付けチェックの要件

整理は、傷病名符号化プロセスの基礎的検討において継続的に求められる課題である。

（3）国内外の動向と先行研究

傷病符号付けを支援するプログラムの研究開発は、国内外で進展しており、人工知能（AI）や機械学習の活用、クラウドベースのソリューション、自然言語処理、インターフェースの改善など、多様な技術的アプローチが提案されている。これらの動向は、医療情報の標準化と相互運用性の向上に寄与し得るものとして注目されている。

本邦においては、医療機関における傷病名の ICD-10 コード付与に関し、医療情報システム開発センター（MEDIS-DC）が管理する ICD10 対応標準病名マスター（厚生労働省標準規格 HS005、以下「標準病名マスター」という。）の利用が、事務連絡等により推奨され普及してきている。標準病名マスターは、レセプト電算処理システム用傷病名マスターと共通化が図られ、電子カルテおよび電子レセプトに使用されている。また、内閣府の戦略的イノベーション創造プログラム（SIP）「AI ホスピタル研究」では、約 40 万語の同義語を含む病名・症状の用語集が作成されている（標準病名集は約 3 万語）。同用語集の一部は商用利用されつつ、今後も拡張・修正が継続される予定であり、符号付け精度の向上に貢献することが期待されている。一方で、ICD-11 への移行に際しては、WHO が ICD-10 を ICD-11 に変換するテーブルを提供しているものの、付与病名そのものや変換テーブルの精度に課題があると考えられている。

先行研究（厚生労働科学研究費補助金 政策科学総合研究「患者調査の効率的な実施手法の確立に資する研究」、21AB1001）[1][2]では、効率化および精度向上の観点から、現状の符号化プロセスの課題分析を踏まえた情報通信技術（ICT）の活用手法が検討されている。同先行研究の検証によれば、ICT ツールは一定程度の表記揺れに対応可能である一方、一部の表記揺れには対応できず、結果として検索精度の低下が確認されている。また、同先行研究では、患者調査の本来の趣旨が臨床病名での調査であることに鑑み、電子カルテやレセプトの病名を ICT ツールに参考病名として取り込むことは可能であるが、調査票上の傷病名候補として直接表示することは望ましくないとの考察がなされている。

公的統計においては、再現性、説明可能性および判定根拠の明示性が、その品質を担保する基盤として求められる。生成 AI や教師あり機械学習は、学習データ、推論過程、出力の安定性および判定根拠の説明可能性の各観点において、公的統計実務への適用に際して検討すべき課題が残されている。

（4）本研究の目的

本研究の目的は、患者調査における傷病符号付けプロセスを前提として、当該プロセスの最適化および効率化に資する基礎的知見を得ることである。

本研究において「最適化」とは、患者調査における傷病名符号化プロセスを前提として、適合度の高い符号候補を選択するための判断材料を提示し、符号化の正確性、一貫性、再現性および判定根拠の明示性の向上に資することを指す。「効率化」とは、用手的確認における確認対象の重点化、および符号候補の選択に伴う判断負荷の軽減可能性を検討することを指す。なお、所要時間の短縮は本研究の検討対象に含めない。

本研究は、患者調査の傷病名符号化プロセスを変更・代替するものではなく、当該プロセスを尊重しながら、独立した照合ロジックを並行して提供することにより、明示性および再現性のある第二の判定根拠を提示する補助的検証手段としての位

置づけで設計した。本研究は、患者調査において最終的に確定された ICD-10 コード（以下「確定符号」という。）を否定するものではなく、当該符号の確認可能性を高めることに資することを意図している。本研究は、確定符号の自動決定を主張するものではなく、AI や機械学習による自動分類を主要方法とするものでもない。

これらの目的に対応するため、本研究班は 5 課題の分担研究を相補的に配置した。具体的には、（１）標準病名マスターを照合源とする符号照合アルゴリズム（Patient-survey Coding Match Algorithm；患者調査符号照合アルゴリズム、以下「PCMA」という。）および照合指数（Patient-survey Coding Match Index；患者調査符号照合指数、以下「PCMI」という。）の試作と評価、（２）AI を用いた類似度判定の試行、（３）ICD 大分類別にみた手付けによる ICD 符号付与の出現要因の定量的分析、（４）辞書照合で符号化されなかった病名の傾向分析、（５）クラスタリングによる不一致パターン分析に基づく既存符号付けアルゴリズムの論理的改善策の整理、（６）患者調査と標準病名マスターの符号化方針相違事例における診療情報管理士による盲検下符号化の予備的検証、を実施した。

これら 5 課題の分担研究は、令和 7 年度研究計画書に示された 3 つの整理項目――手付け符号化が必要となった傷病名の事例分類と発生要因の分析、患者調査用辞書ファイルの要件整理、自動符号付けアルゴリズムと傷病符号付けチェックの要件整理――を相補的に網羅する構造として設計されている。

（５）期待される効果

本研究の成果は、(1)コーディングの最適化・効率化に向けた基盤構築：PCMA の試作により、標準病名マスターを補助的参照基盤として用い、形態素解析とは異なる照合原理に基づく補助的判定根拠を提示する基盤を構築した。(2)辞書ファイル等を含む傷病符号付けに関する更新の方向性提示：傷病名の表記揺れにすべて辞書登録で対応することには限界がある。一方、近似照合により高い正答率が得られる語については、辞書登録だけに依存せずに符号候補を提示できる可能性が示された。(3)傷病符号付けチェックプログラム精緻化への貢献：PCMA で実装した概念点検の考え方を既存のチェックプログラムに応用し得ることを示すとともに、不一致病名の上位分類およびクラスタ構造を整理し、チェック対象を重点化するための基礎資料を得た。(4)統計調査公表の効率化に向けた基盤形成：PCMA により、手付け符号化を要する語の量的縮減可能性を示し、傷病符号付け作業の効率化を検討するための基礎資料を得た。なお、作業時間は測定していない。(5)次期統計処理システム導入への準備：PCMA を試行的に構築し、標準病名マスターを照合源とする傷病符号照合アルゴリズムの統計処理への適用可能性を示した。本研究は、患者調査における傷病符号付けの検証可能性と再現性を高める技術的基盤を提示した点に意義がある。今後の課題は、独立参照基準の構築および過年度データによる外的妥当性検証である。

以上の目的のもとで、本研究班は令和 5 年患者調査により得られた傷病名を試料として、5 課題の分担研究を実施した。次章以降、研究方法、研究結果、考察および結論を順に述べる。

Ｂ．研究方法

(1) 研究全体の基本方針

本研究は、患者調査における傷病名符号化プロセスを前提として、当該プロセスの正確性、一貫性、再現性および判定根拠の明示性を高めるため

の補助的検証手段ならびに基礎的知見を得る研究として設計した。本研究は、患者調査の既存プロセスを否定、変更または代替することを意図したものではない。

公的統計には、再現性、説明可能性および判定根拠の明示性が求められる。本研究班は、これらの要件と整合的な方法論として、ルールベースの並行照合手段を中核に据えた。生成 AI および教師あり機械学習は、学習データ、推論過程、出力の安定性および判定根拠の説明可能性の各観点において公的統計実務への適用に際して検討すべき課題が残されていることから、本研究では主要方法としては採用せず、補助的試行に限定して位置づけた。

本研究班は 5 課題の分担研究を相補的に配置し、令和 7 年度研究計画書に示された 3 つの整理項目――手付け符号化が必要となった傷病名の事例分類と発生要因の分析、患者調査用辞書ファイルの要件整理、自動符号付けアルゴリズムと傷病符号付けチェックの要件整理――にそれぞれ対応させた。

(2) 患者調査の傷病名符号化プロセスの位置づけ

本研究では、患者調査における傷病名符号化プロセスを、A 節(2)で述べた 3 工程――一次工程としての辞書照合、二次工程としての形態素解析を活用した手付け、三次工程としての手付け符号化――として固定的に位置づける。本研究は、これら 3 工程のすべてを尊重したうえで、各工程と相補的に並行する基礎的知見を提供する設計である。

PCMA および PCMI は、患者調査の確定符号を自動的に決定するものではなく、傷病名に対し、独立した照合ロジックに基づき適合度の高い ICD-10 符号候補を提示し、用手的確認を支援する第二の判定根拠として位置づける。

(3) 対象データ

本研究班は、令和 5 年(2023 年) 患者調査により得られた傷病名を主たる対象とした。各分担研究の対象データの範囲および分母の構成は、(5)各分担研究の方法において個別に示す。

(4) 5 課題の分担研究の構成

(4-1) 4 群構成

本研究班は、5 課題の分担研究を単に並列的に配置するのではなく、研究計画書の目的に沿った相補的構造として、以下の 4 群に整理した。

第 1 群は、標準病名マスターを参照軸とする補助的検証を担う。荒井・西本研究が、標準病名マスターを照合源とする並行照合手段の試作と評価を実施する。

第 2 群は、患者調査の傷病符号付けプロセスにおける課題の可視化を担う。桜澤研究、來島研究、森藤研究の 3 課題が、ICD 大分類別の傾向、ICD1 桁・3 桁別の分布、表記の特徴、ならびに辞書判定コードと手付けコードの不一致パターンの観点から、手付け符号化を要する語の発生構造を多面的に明らかにする。

第 3 群は、自動符号付けアルゴリズムと傷病符号付けチェックの要件整理を担う。荒井・西本研究、阿南研究、森藤研究の 3 課題が、ルールベース照合の試作、AI を用いた類似度判定の試行、ならびに既存符号付けアルゴリズムの論理的改善策の提示を通じて、要件整理に資する基礎的知見を提供する。

第 4 群は、患者調査の符号化方針と標準病名マスター側 ICD-10 コードとの関係整理を担う。荒井・石割研究が、診療情報管理士による盲検下符号化の予備的検証を実施し、両体系の符号化方針差を、文字列照合の結果と専門家評価の双方の観

点から検討する。

これら 4 群は相互に独立ではなく、複数の分担研究が複数の群にまたがる相補的構造となっている。

(4-2) 研究計画書の 3 整理項目との対応

令和 7 年度研究計画書に示された 3 つの整理項目と 5 課題の分担研究の対応関係は、以下のとおりである。

整理項目 1(患者調査用辞書ファイルの要件整理)には、荒井・西本研究、來島研究、森藤研究の 3 課題が対応する。

整理項目 2(手付け符号化が必要となった傷病名の事例分類と発生要因の分析)には、桜澤研究、來島研究、森藤研究の 3 課題が対応する。

整理項目 3(自動符号付けアルゴリズムと傷病符号付けチェックの要件整理)には、荒井研究、阿南研究、森藤研究の 3 課題が対応する。

これにより、5 課題が研究計画書の目的に沿って相補的に実施されたことを示す。

(5) 各分担研究の方法

(5-1) PCMA および PCMI の試作と評価(荒井・西本研究)

本分担研究は、本研究班の中核成果に位置づけられる。標準病名マスターを照合源とする PCMA を試作し、令和 5 年患者調査の傷病名符号化が完了したデータを対象に、患者調査の確定符号との一致率および工程別の振る舞いを評価した。

対象データは、令和 5 年患者調査において得られた種類数（type 数。重複を除いた傷病名の種類数）を単位とする。本分担研究では、評価対象を主指標対象 111,984 語と参考指標対象 110,670 語の 2 分母として構成した原データ 112,133 語からデータ例外 14 語および非傷病名 135 語を除外した 111,984 語を主指標対象とし、ここからさらに ICD 未設定または候補なし 116 語、参照基準不明 15 語(確定符号不明 9 語および索引一致・確定符号不明 6 語)、ならびに符号化方針差 1,184 語(索引一致・参照基準相違 746 語、左右差不問・参照基準相違 438 語)を除外した 110,670 語を参考指標対象とした[注 1]。

[注 1]：分担研究 1 原典の表 2 における左右関連の区分について、本総括報告書では「左右差不問・参照基準相違」に統一する。

照合源には、MEDIS-DC が管理する ICD10 対応標準病名マスター Ver.5.17 を用いた。標準病名マスターを照合源として採用した理由は 3 点ある。第一に、当該マスターは厚生労働省標準規格 HS005 に位置付けられた標準的病名資源であること、第二に、医療情報システムにおいて広く利用される標準的資源であること、第三に、患者調査の過去実績に基づく傷病名辞書とは独立した外部資源であり、第二の判定根拠を提供しうる位置にあることである。

PCMA は、Python によって実装したルールベースの照合アルゴリズムである。生成 AI および教師あり機械学習を主要方法として採用しなかった理由は、(1)に述べた説明可能性および再現性の要請による。PCMA の処理は、以下の 3 段階で構成される。第一段階は正規化である。傷病名に対し、NFKC による文字種正規化、全角・半角の統一、異字・誤字の変換、ならびに修飾語の識別と除去を行う。第二段階は索引用語との照合である。正規化された入力語と標準病名マスターの索引用語との間で照合を行い、上位候補を順位付けする。第三段階は概念点検である。疾病、傷害及び死因の統計分類提要 ICD-10 準拠 第 1 巻（内容例示表）に基づき、候補 ICD-10 コードの医学概念ならびに章・サブ分類との整合性を判定する。

複合的に記述された傷病名については、句読点等を区切りとして並列分割し、先頭パートを優先する簡略ルールを用いた。なお、第 18 章および第 21 章の症状・所見または保健サービス利用の理由が先頭にあり、後続に第 2 章悪性新生物等の特定章の傷病名を含む場合の例外規定を併設した。

PCMA は、傷病名ごとに上位 5 候補を ICD-10 コードを出力する設計とした。これは、用手的確認における確認範囲の絞り込みを支援する意図に基づく。PCMI は、候補の適合度を示す照合指数であり、0 から 1 の範囲の値をとる。PCMI は、8 観点の重み付き線形結合――共通文字率、非共通文字ペナルティ、語順整合度、索引用語長係数、2-gram 一致率、形態素類似度、修飾語タイプ整合度、ならびに修飾語除去後文字集合共通率――を正規化したうえで、ICD-10 章およびサブ分類との整合性を表す互換性係数を指数として適用し、補正項を加えて算出される。PCMI は候補順位付けを目的とした指標であり、候補が患者調査の確定符号と一致する確率を示すものではない。

PCMI の算出に用いる重み係数は、基本観点 8 個、互換性係数指数 1 個および補正項 2 個の計 11 次元である。これらの重み係数は、Optuna に実装された CmaEsSampler による CMA-ES 最適化により決定した。再現性を確保するため、最適化過程の乱数シードは 42 に固定した。なお、CMA-ES は重み係数のバランスの微調整に限定して用いており、アルゴリズム上の構造的問題の解決手段としては用いていない。

参照基準は、患者調査において最終的に確定された ICD-10 コード(以下、参照基準と確定符号を併用する)とした。評価単位は、ICD-10 基本分類表の 3 桁分類とした。PCMI 上位候補の ICD-10 コードと確定符号が同一の 3 桁分類に属する場合を、参照基準と一致した語と定義する。本分担研究では、分担研究 1 原典の用法に従い、本研究上の評価基準に対する一致を意味する語として「正答」「正答率」を限定的に使用するが、これは独立した医学的・分類学的絶対解との一致率を意味するものではない。統計手法として、形態素解析との比較には McNemar 検定(Yates 連続補正、有意水準 $\alpha=0.05$)を用い、各群の一致率に対する 95%信頼区間(以下「95%CI」という。)は Wilson 法により算出した。本分担研究では、文字列照合の結果として候補が選択されたことを意味する「一致」と、当該候補の ICD-10 コードが参照基準の 3 桁分類と同一であることを意味する「正答」を区別して用いる。

(5-2) AI を用いた類似度判定の試行(阿南研究)

本分担研究は、自動符号付けアルゴリズムと傷病符号付けチェックの要件整理を目的として、AI を用いた類似度判定を試行した。対象とした病態は、慢性 C 型肝炎、慢性 B 型肝炎およびアルコール性肝炎の 3 疾患である。これら 3 疾患は、いずれも慢性肝疾患であるという共通点を有しつつ、原因に基づく特有の特徴をも併せ持つことから、類似度判定の試行対象として選定した。

使用した解析ツールは、Gemini および MeCab である。MeCab による形態素解析の後、最多単語数の傷病名との類似度を比較することにより、コーディング精度の向上可能性を試行的に検証した。

(5-3) ICD 大分類別にみた手付けによる ICD 符号付与の出現要因(桜澤研究)

本分担研究は、令和 5 年度患者調査の傷病名について、辞書ファイル等により符号付けが行われた群と、手付けによる符号付けとなった群に区分したうえで、両群の傷病名リストを統合し、分析用データセットを構築した。なお、本分担研究では分析単位を「件」として集計している[注 2]。

[注 2]：本総括報告書では、PCMA 研究文脈での分

析単位を「語」(得られた種類数)とし、本分担研究および後述の來島研究の分析単位を「件」とする。

解析は、ICD 大分類別の集計、Chi-square test、ならびに単変量および多変量ロジスティック回帰分析により行った。解析ソフトウェアは IBM SPSS Ver.30 を用いた。

(5-4) 辞書照合で符号化されなかった病名の傾向分析(來島研究)

本分担研究は、令和 5 年患者調査において辞書照合で符号化されなかった病名 75,218 件を対象に、ICD1 桁(アルファベット)および 3 桁分類別の出現状況を集計し、出現頻度の高い分類への傾向を分析した。本分担研究の分析単位も「件」である。また、新型コロナウイルス感染症関連の傷病名については、表記揺らぎの実態を個別に検証した。

(5-5) クラスタリングによる不一致パターン分析(森藤研究)

本分担研究は、最新の傷病名辞書によるマッチングを経て、形態素解析による参考コードが付与された 59,436 語のうち、手付け符号付け作業により ICD コードが変更された 23,498 語を対象とした。辞書コードと手付けコードを 2 次元ベクトルとしてラベルエンコーディングし、K-means 法によりクラスタリングを行った。クラスタ数の決定には、エルボー法およびシルエット係数に基づく数量的判断と、実務上の診療領域分離状況を踏まえた解釈的判断を併用し、最終的にクラスタ数 8 を採用した。解析はレコード単位 30,020 レコードで実施した。

本分担研究におけるクラスタリングは、実務上の既知の不一致傾向を論理的に整理することを目的とするものであり、自動学習モデルの構築を目的とするものではない。

(5-6) 専門家による盲検下符号化の予備的検証(荒井・石割研究)

本分担研究は、本研究班の前報により抽出された符号化方針相違事例 1,184 件を母集団とし、ICD-10 章別の構成比を反映するように抽出した 124 症例(標本割合 10.6%)を検証対象とした。評価者は、診療情報管理士 4 名(3 組、うち 1 組は 2 名共同付与)である。評価用ワークシートには、入力語、索引用語、正規化後表記、病名表記、ならびに参考 ICD-A および参考 ICD-B を併記し、確認結果列に ICD-10 コードを記入する構成とした。参考 ICD-A と参考 ICD-B は、患者調査 ICD と標準病名マスターICD に対応するが、いずれが患者調査由来であるかは評価者に開示しない盲検評価とした。評価者は、標準病名マスターを参照しない条件下で ICD-10 コードを付与した。主要評価指標は、事例単位(n=124)における 3 組の多数決一致先である。3 組のうち 2 組以上が同一の参考符号と 3 桁レベルで一致した場合に多数決成立とみなし、95%CI は Wilson 法により算出した。副次評価指標として、4 桁完全一致、3 桁一致および ICD-10 章一致の 3 水準で延べ観測を行い、事例クラスタを考慮したノンパラメトリック・ブートストラップ法(2,000 回反復)により分布を推定した。専門家間一致度の指標として、ペアごとの Cohen の κ 係数および 3 組間の Fleiss の κ を算出した。本分担研究では、独立評価者対である A-B ペアの Cohen の κ を主要解釈の対象とし、A-C および B-C ペアならびに Fleiss の κ は参考値として扱う。これは、評価者群 C が 2 名共同付与であることに起因する一致度の偏りを考慮したものである。本分担研究は、仮説検定を行わず、効果量と信頼区間による記述統計的提示にとどめる予備的検証として位置づけた。

(倫理面への配慮)

本研究で取り扱うデータには、個人を直接識別できる情報は含まれていない。本研究は、人を対象とする生命科学・医学系研究に関する倫理指針との関係において、介入を伴わず、かつ個人を直接識別できる情報を含まない既存資料を用いる研究として実施したものであり、当該指針における倫理面の問題は生じない。MEDIS-DC が管理する標準病名マスター Ver.5.17 については、その利用条件に従って研究目的での参照に用いた。分担研究 6 においては、診療情報管理士 4 名（3 組）による盲検下符号化を実施したが、評価者個人を特定可能な形での結果記述を避けるため、専門家 A・B・C の匿名表記とした。

以上が、本研究班の研究方法である。次節「C. 研究結果」では、これら 5 課題の分担研究の結果を、第 1 群の PCMA/PCMI の試作と評価、第 2 群の課題可視化、第 3 群の AI 類似度判定の試行、第 4 群の診療情報管理士による盲検下符号化の順に統合して述べる。

C．研究結果

（1）患者調査の傷病符号付けプロセスにおける課題の可視化

桜澤研究、來島研究および森藤研究の 3 課題は、それぞれの観点から、辞書照合で符号化されなかった語の発生構造を明らかにした。本節では、これら 3 課題の結果を、規模および分布、ICD 大分類別の傾向、ICD1 桁および 3 桁分類別の分布、表記の特徴、ならびに辞書判定コードと手付けコードの不一致パターンの観点から、統合的に示す。

（1－1）規模および分布

桜澤研究は、令和 5 年患者調査の対象全体 112,133 件を、辞書照合により符号化された群（36,915 件、32.9%）と、辞書照合後に手付けを要した群（75,218 件、67.1%）に群別したうえで、両群間の分布特性を分析した。本分担研究では分析単位を「件」として集計した[注 1]。

[注 1]：桜澤研究および來島研究の分析単位は「件」、PCMA 研究および森藤研究の分析単位は「語」である。両者の対象範囲は近接するが、評価単位および抽出条件が異なるため、各分担研究の文脈で扱う。

辞書照合後に手付けを要した群の中で、出現頻度の高い分類は、新生物 19,858 件（17.7%）、損傷および中毒等 18,752 件（16.7%）、筋骨格系および結合組織の疾患 10,760 件（9.6%）、ならびに循環器系の疾患 9,747 件（8.7%）であった。

（1－2）ICD 大分類別の傾向

桜澤研究は、ICD 大分類別に辞書照合後に手付けを要した割合を集計し、Chi-square test および単変量・多変量ロジスティック回帰分析を行った。

対象全体における辞書照合後に手付けを要した割合（67.1%）よりも有意に高い分類として、U コード、Z コード、呼吸器系の疾患、皮膚および皮下組織の疾患、循環器系の疾患、ならびに消化器系の疾患が示された。一方、有意に低い分類として、周産期に発生した病態、先天奇形、耳および乳様突起の疾患、眼および付属器の疾患、神経系の疾患、損傷および中毒等、ならびに新生物が示された。

多変量ロジスティック回帰分析では、辞書照合後に手付けを要する確率を高める影響要因として、U コード（オッズ比 5.963）、Z コード（同 2.594）、呼吸器系の疾患（同 1.231）、皮膚および皮下組織の疾患（同 1.223）、ならびに循環器系の疾患（同 1.153）が導出された。周産期に発

生した病態のオッズ比は 0.601 であり、対象全体の中で最も低い値を示した。

（1－3）ICD1 桁および 3 桁分類別の分布

來島研究は、辞書照合で符号化されなかった病名（アンマッチ病名）75,218 件を対象に、ICD1 桁分類（アルファベット）および 3 桁分類別の出現状況を集計した。本分担研究の分析単位も「件」である。

ICD1 桁分類では、C コード（悪性新生物）が 9,395 件（12.5%）と最も多く、続いて S コード（損傷および中毒等のうち外傷に該当するコード、8,203 件、10.9%）、M コード（筋骨格系および結合組織の疾患、7,248 件、9.6%）、ならびに循環器系の疾患（6,867 件、9.1%）が上位を占めた。これら上位 4 分類で全体の 42.16%を構成した。

ICD3 桁分類では、I63（脳梗塞）が 1,442 件（1.9%）、C34（肺癌）が 1,198 件（1.6%）、C50（乳癌）が 865 件（1.1%）、L30（湿疹）が 760 件（1.0%）、ならびに D48（骨腫瘍）が 750 件（1.0%）の順に多く確認された。

「癌」を含む病名は 6,339 件確認され、そのうち 5,956 件（93.9%）が C コード（悪性新生物）に該当した。C コードに該当しなかった病名の内訳としては、上皮内癌・非浸潤癌 102 件（1.6%）、複数病名による別疾患への分類 177 件（2.8%）、ならびにその他 104 件（1.7%）が含まれた。

（1－4）表記の特徴

來島研究では、辞書照合でアンマッチ病名における側性表現の出現頻度として、「右」12,555 件（16.7%）、「左」12,239 件（16.3%）、「両」5,894 件（7.8%）、「両側」1,925 件（2.6%）、「左側」662 件（0.9%）、「右側」569 件（0.8%）、「左右」128 件（0.2%）、「ミギ」111 件（0.1%）、ならびに「ヒダリ」102 件（0.1%）が観察された（いずれも辞書照合でアンマッチ病名全体に対する割合、重複を許容して集計）。

新型コロナウイルス感染症に関連するアンマッチ病名表記は 916 件確認され、そのうち 715 件（78.1%）が U コードに符号化されていた。U コード以外への符号化は、感染後の病態が 54 件（7.6%）、関連する多系統炎症性症候群が 2 件（0.3%）であった。U コード内訳では、U07.1 が 641 件、U07.2 が 18 件であった。表記には、「コヴィッド 19」、「COVID」、「コロナウイルス感染症 2019」、「新型コロナウイルス感染症」など、略語、カタカナ表記、ならびに数字の組み合わせによる多様な揺らぎが認められた。

なお、標準病名マスター（Ver.5.16）における左右表記に起因する ICD コード変更事例は 24 件にとどまっていた。

（1－5）辞書判定コードと手付けコードの不一致パターン

森藤研究は、最新の傷病名辞書によるマッチングを経て参考コードが付与された 59,436 語のうち、手付け符号付け作業により ICD コードが変更された 23,498 語（約 39.5%）を対象に、辞書コードと手付けコードを 2 次元ベクトルとしてラベルエンコーディングし、K-means 法によるクラスタリングを行った。エルボー法およびシルエット係数（SC=0.52）に基づく数量的判断と、診療領域の分離状況を踏まえた解釈的判断により、クラスタ数 8 を採用した。レコード単位での解析対象は 30,020 レコードであった。

各クラスタの構成は以下のとおりであった。すなわち、Cluster 0（循環器疾患）5,363 語、Cluster 1（外傷）5,841 語、Cluster 2（感染

症）3,864 語、Cluster 3（泌尿器・産科関連）3,321 語、Cluster 4（眼科・精神疾患）2,807 語、Cluster 5（皮膚・筋骨格疾患）5,322 語、Cluster 6（妊娠・管理関連）1,800 語、ならびに Cluster 7（その他）1,702 語である。

クラスタ別の代表的な不一致パターンとしては、循環器疾患領域における「Z 付きコード」の基本コードへの置換、外傷領域における骨折部位の詳細分類による不一致、感染症領域における原因菌・ウイルス種の詳細差、産科領域における妊娠合併症の文脈依存的判断、眼科・精神疾患領域における年齢等の条件分岐を伴う詳細分類の選択、皮膚・筋骨格疾患領域における分類体系および脊椎詳細部位の選択差、妊娠・管理関連領域における R コード（症状・徴候）から特定疾患コードへの置換、ならびに、その他領域における呼吸器疾患の詳細差および COVID-19 関連への修正が観察された。

変更理由のコメントが付与された 784 語の分析からは、「年齢条件（16 歳未満等）によるコード分岐」、「初筆（最初に記載された傷病名）の優先判断」、「部位の明記不足による局所性判断の要否」、ならびに「『？』を含む不明瞭な記載に対する目視確認」が、不一致を生じさせる要因の上位として確認された。

（1-6）群間で共通する所見および相補的所見

3 課題の結果から、以下の所見が共通的および相補的に確認された。

第一に、U コードおよび COVID-19 関連の傷病名は、桜澤研究（多変量ロジスティック回帰における U コードのオッズ比 5.963）、來島研究（U07.1 単独で 641 件、関連表記 916 件）の双方において、辞書照合で符号化されにくい領域として一致して観察された。

第二に、Z コードについても、桜澤研究（オッズ比 2.594）と森藤研究（Cluster 0 における「Z 付きコード」の基本コードへの置換）の双方において、辞書判定と手付けの間に体系的な不一致が観察された。

第三に、悪性新生物（C コード）については、桜澤研究（出現件数 19,858 件、対象全体の 17.7%）、來島研究（辞書照合で符号化されなかった病名における上位、肺癌・乳癌）、ならびに森藤研究（Cluster 2 における腫瘍コード詳細化）の各観点から確認された。「癌」を含む病名の 93.9%が最終的に C コードに分類された一方で、上皮内癌・非浸潤癌の判別、複数病名による別疾患への分類、境界悪性病変、ならびに表記誤りなど、辞書照合に到達しない要因が並列的に存在する構造が示された。

第四に、側性表現（右・左・両等）については、來島研究で多数確認された一方で、標準病名マスター上での ICD コード変更事例は 24 件にとどまっており、左右情報の ICD コーディング全体への影響は限定的である可能性が示された。

これらの所見は、辞書ファイル更新方針および傷病符号付けチェックプログラム精緻化における重点領域として、D 節（考察）において論ずる。

（2）PCMA および PCMI の試作と評価

荒井・西本研究は、本研究班の中核成果として、標準病名マスターを照合源とする PCMA および PCMI の試作と評価を実施した。本節において、分担研究 1 原典の用法に従い「正答」「正答率」を用いる場合があるが、本研究における正答とは、参照基準（患者調査の確定符号）と 3 桁分類単位で一致した候補を指すものであり、独立した医学的・分類学的絶対解との一致を意味するものではない。

（2-1）主指標と工程別の参照基準との一致率

主指標対象 111,984 語における PCMA 第 1 位候補と確定符号との 3 桁分類単位での一致率は 85.09%（95,288／111,984、95%CI 84.88-85.30）であった。参考指標対象 110,670 語における同一致率は 86.10%（同 95%CI 85.90-86.30）であった。

工程別の参照基準との一致率は以下のとおりである。すなわち、辞書照合工程 92.03%（33,945／36,883、95%CI 91.75-92.31）、形態素解析を活用した手付け工程 83.48%（49,594／59,408、95%CI 83.18-83.78）、ならびに手付け符号化工程 74.87%（11,749／15,693、95%CI 74.18-75.54）であった[注 2]。

[注 2]：分担研究 1 原典における一次工程の表記について、本総括報告書では「辞書照合」に統一する。

（2-2）形態素解析との比較

形態素解析を活用した手付け工程の対象における比較を行うため、当該工程の対象 59,408 語のうち、形態素解析側で第 1 位候補が出力された 59,375 語を対象に、両者の出力 ICD-10 コードと確定符号との関係を集計した。

両者の出力が 3 桁分類単位で一致した語は 41,754 語（70.32%）であった。確定符号との一致率は、PCMA が 83.49%（49,575／59,375）、形態素解析が 72.02%（42,764／59,375）であり、その差は＋11.47 ポイントであった（McNemar 検定 $\chi^2 = 3,165.39$、$p < 0.0001$）。この結果は、標準病名マスターを照合源とするアプローチが、形態素解析とは異なる確認手がかりを提供し得ることを示すものである。

（2-3）辞書照合で処理されなかった語の 3 層構造

辞書照合で処理されなかった語は、形態素解析を活用した手付け工程の対象 59,408 語と、手付け符号化工程の対象 15,693 語の合計 75,101 語であり、主指標対象に対して 67.06%を占める。本研究では、これらの語のうち、形態素解析側で第 1 位候補が確認できた 59,375 語と、形態素解析が候補を提示しなかった語に対する手付け符号化対象 15,693 語を合わせた 75,068 語を、形態素解析と PCMA の出力関係に基づき 3 層に層別化した。

第 1 層は、形態素解析と PCMA が一致した符号を提示した語であり、41,754 語（辞書照合で処理されなかった語 75,101 語のうち 55.60%）であった。第 1 層における参照基準との一致率は 93.03%（95%CI 92.78-93.27）であった。第 1 層は、文字列照合と言語処理という異なる照合原理が同一の ICD-10 コードに到達した語であり、用手的確認における判断負荷の軽減可能性を示唆する層である。

第 2 層は、形態素解析と PCMA がいずれも符号を提示するが、両者が一致しない語であり、17,621 語（同 23.46%）であった。第 2 層は、両者の候補を手がかりとして手付け符号化に資する層である。

第 3 層は、形態素解析が候補を提示しなかった語に対する手付け符号化対象であり、15,693 語（同 20.89%）であった。このうち、PCMA が候補を提示した語は 15,652 語、PCMA が候補を提示しなかった語は 41 語であった。第 3 層は、形態素解析が手がかりを与えなかった語に対して、標準病名マスター照合に基づく候補情報を付加できる可能性が確認された層である。

第 1 層 41,754 語のうち、参照基準と一致しなかった語は 2,910 語（6.97%）であった。その内訳は、アルゴリズム改善余地として算出された語 2,171 語（74.60%）、符号化方針差に起因する語

696 語（23.92%）、ICD 未設定の語 28 語（0.96%）、ならびに参照基準不明の語 15 語（0.52%）であった。

（2-4）PCMI 閾値別の参照基準との一致率

PCMI 閾値別の該当語数と参照基準との一致率を集計したところ、閾値が高くなるにつれて該当語数（カバー率）は単調に減少する一方、一致率は閾値 0.50 から 1.00 の範囲で 88.6%から 89.2%の狭い帯域に留まり、単調関係を示す傾向は観察されなかった。この結果は、PCMI が候補順位付けを目的として設計された指標であることを反映しており、PCMI 値そのものを点検対象選別の主指標として用いることは適切でないことを示している。辞書照合で処理されなかった語の層別化は、PCMI 値ではなく、形態素解析と PCMA の出力関係という離散的根拠を主とする。

（2-5）標準病名マスターと患者調査の符号化方針差

PCMA が選択した索引用語と傷病名が完全に一致しながら確定符号と一致しなかった語は 1,184 語確認された。その内訳は、索引一致・参照基準相違 746 語、ならびに左右差不問・参照基準相違 438 語であった[注 3]。これらは文字列照合上の誤りではなく、標準病名マスターにおける ICD-10 付与方針と、患者調査における符号化方針との差に起因するものであり、参考指標対象から除外して扱った。

[注 3]：分担研究 1 原典の表 2 における左右関連の区分について、本総括報告書では「左右差不問・参照基準相違」に統一する。

（3）AI 類似度判定の試行から得られた要件

阿南研究は、自動符号付けアルゴリズムと傷病符号付けチェックの要件整理を目的として、慢性 C 型肝炎、慢性 B 型肝炎およびアルコール性肝炎の 3 疾患を対象に、Gemini および MeCab を用いた類似度判定の試行を行った。

MeCab による形態素解析の結果、「慢性 C 型肝炎」は「慢性／C 型／肝炎」、「慢性 B 型肝炎」は「慢性／B 型／肝炎」、ならびに「アルコール性肝炎」は「アルコール性／肝炎」に分割された。Gemini によるコサイン類似度の計算結果は、慢性 C 型肝炎と慢性 B 型肝炎の間で 0.853、慢性 C 型肝炎とアルコール性肝炎の間で 0.236、ならびに慢性 B 型肝炎とアルコール性肝炎の間で 0.211 であった。慢性 C 型肝炎、ならびに「慢性」を削除した C 型肝炎は、いずれも標準病名マスターに対応する ICD-10 コード B18.2 と整合する結果が得られた。

形態素解析単独では、用語および記述の類似性は把握可能であるが、ICD 分類条件（定義および意味）に基づく判断には限界がある。この限界は、原因菌、ウイルス種、病勢、病態あるいは原因に関する識別を必要とする傷病名において顕在化する。AI あるいは形態素解析は、傷病符号付けに対する補助的手がかりを提供し得る一方で、ICD 分類上の意味判断には限界があり、特に公的統計実務に求められる判定根拠の明示性および再現性の観点から、十分な配慮を要する。

実装上の課題として、(1) 使用する AI ツールの違いによる出力差が生じる可能性、(2) 形態素解析後の単語数が増加した場合の類似度低下の可能性、(3) 同一 ICD に対する複数の標準的傷病名（例えば、短い表現と長い表現）の準備の必要性、ならびに (4) 大量データ処理における時間的負担、の 4 点が確認された。

（4）符号化方針差事例に対する専門家検証

荒井・石割研究は、患者調査の確定符号と標準病名マスター側 ICD-10 コードとが相違する事例（符号化方針相違事例）に対し、診療情報管理士

4名（3組）が標準病名マスターを参照しない条件下で盲検下符号化を行う予備的検証を実施した[注4]。

[注4]：分担研究6では、原報告に従い、対象母集団を「符号化方針相違事例1,184件（索引一致・参照基準相違746件、左右の指定差異による438件）」として扱う。分担研究1（PCMA研究）の文脈における「符号化方針差1,184語」とは、評価単位および抽出条件が異なるため、各分担研究の文脈で扱う。

（4-1）多数決一致先

対象124症例の事例単位における多数決一致先は、標準病名マスター側ICD-10コードが80件（64.5%、95%CI 55.8-72.4）、患者調査の確定符号が34件（27.4%、同20.3-35.9）、多数決不成立が10件（8.1%、同4.4-14.2）であった。

この分布は、本予備標本における選択分布の記述にとどめ、両方針の優劣を示すものではない。なお、評価用ワークシートには参考ICD-Aおよび参考ICD-Bの双方が併記されており、評価者の付与結果の約90%が4桁レベルで参考ICD-AまたはBのいずれかに完全一致していたことから、本検証は実質的に2候補併記下での盲検選択評価としての性格を有する。

（4-2）評価者間一致度

独立評価者対であるA-BペアにおけるCohenのκ係数は、3桁レベルで0.577（中等度の一致）、章レベルで0.778（かなりの一致）であった。A-CおよびB-CペアのCohenのκ係数（3桁レベル0.661-0.694、章レベル0.829-0.837）、ならびに3組間のFleissのκ係数（3桁レベル0.647、章レベル0.819）は、A-Bペアより高い値を示したが、評価者群Cが2名共同付与であることに起因する一致度の偏りの可能性に留意し、参考値として扱う。

延べ観測（最大370観測）における標準病名マスター側ICD-10コードへの一致は246件（66.5%）、患者調査の確定符号への一致は108件（29.2%）であった。観測の独立性が成立しないため、事例クラスタを考慮したノンパラメトリック・ブートストラップ法による95%信頼区間（標準病名マスター側ICD-10コード：60.2-72.2%、患者調査の確定符号：23.5-35.0%）を併記した。

（4-3）探索的所見

事後的に同定された記述統計として、以下の所見が観察された。

第一に、術後状態に関する記述13件において、患者調査の確定符号は原疾患章（第Ⅱ章、第Ⅸ章、第Ⅺ章等）または第ⅩⅧ章に符号化されている一方、標準病名マスター側ICD-10コードは多くの場合、第ⅩⅪ章Z90-Z99に該当するコードに符号化されていた。具体例として、「冠動脈バイパス術後」におけるI25.1とZ95.1、「早期胃癌術後」におけるC16.9とZ08.0、ならびに「前立腺癌術後状態」におけるC61とZ98.8が観察された。

第二に、入力語「コロナウイルス感染症」が1件含まれ、患者調査の確定符号としてはU07.1（COVID-19、ウイルスが特定されたもの）が、標準病名マスター側ICD-10コードとしてはB34.2（コロナウイルス感染症、部位不明）が、それぞれ符号化されていた。本事例においては、3組すべてが標準病名マスター側ICD-10コードを選択した。

これらの探索的所見は、両方針の優劣を示すものではなく、両方針はそれぞれの目的（基幹統計としての疾病別医療需要・受療動向の把握、ならび

に電子カルテ等の臨床用途における手術後の継続的状態の表現）に照らしいずれも妥当な、運用目的の違いに由来する構造的差異として位置づけられる。

D．考察

（1）研究計画との整合性

本研究班は、令和 7 年度研究計画書に示された 3 つの整理項目を、5 課題の分担研究によって相補的に網羅した。すなわち、手付け符号化が必要となった傷病名の事例分類と発生要因の分析は桜澤研究、來島研究および森藤研究が担い、患者調査用辞書ファイルの要件整理は來島研究、森藤研究および荒井・西本研究が担い、自動符号付けアルゴリズムと傷病符号付けチェックの要件整理は荒井・西本研究、阿南研究、森藤研究および荒井・石割研究が担った。

研究計画書の【期待される効果】に示された 5 項目――傷病符号付けの適正化および効率化に向けた基盤構築、辞書ファイル等の更新の方向性提示、傷病符号付けチェックプログラム精緻化、迅速な統計調査公表への基盤形成、ならびに次期統計処理システム導入準備――に対しても、本研究班の成果は具体的に応答している。

（2）患者調査の意義を踏まえた研究の位置づけ

患者調査は、統計法に基づく基幹統計であり、医療施設を利用する患者の傷病状況を明らかにする我が国唯一の網羅的な疾病統計である。基幹統計には、継続性、信頼性、説明可能性および再現性が一貫して求められる。過去調査の蓄積、符号化担当者による判断、ならびに分類運用上の経験は、いずれも基幹統計としての品質を支える要素であり、本研究はこれらを尊重する立場をとった。

このため、本研究は、患者調査の既存プロセスを変更、代替あるいは置換するものではなく、当該プロセスに対する補助的検証手段の試作、ならびに改善に資する基礎的知見の提示として位置づけられる。本研究の成果は、患者調査の確定符号を否定するものではなく、当該符号の確認可能性および判定根拠の明示性を高めるための資料として活用されることを意図したものである。

（3）本研究が示す最適化

本研究における最適化は、単なる一致率の向上を意味するものではない。それは、符号候補の適合度、候補順位の妥当性、判定根拠の明示性、再現性、ならびに一貫性の向上に資するものとして定義される。

PCMA および PCMI は、患者調査の確定符号とは独立した照合ロジックに基づき、文字列照合（PCMI による候補順位付け）と概念点検（ICD-10 内容例示表[4]に基づく医学概念整合性確認）から構成される判定根拠を、明示的かつ再現的に提供する。PCMA の主指標対象における第 1 位候補の参照基準との一致率は 85.09％を示し、辞書照合工程における同一致率は 92.03％であった。これらの結果は、PCMA が、患者調査の確定符号と一定程度整合する形で機能し得ることを示している。

桜澤研究による ICD 大分類別の発生要因分析は、辞書ファイル機能改良の重点領域（U コード、Z コード、皮膚、呼吸器、循環器等）を定量的に提示することにより、辞書ファイル更新の優先順位判断に寄与する。

來島研究による ICD1 桁および 3 桁分類別の分析は、悪性新生物に関する表記、側性表現、ならびに新型コロナウイルス感染症関連表記の整理が、辞書ファイルの網羅性向上に資する具体的方向性を示した。

森藤研究によるクラスタリング分析は、既存の表記揺れ補正処理、目印コード抽出条件、ならびにアンマッチ基礎表における参考情報の充実といった、既存プロセスの論理的改善方向性を、AI や新規技術に依存しない形で提示した。

荒井・石割研究による専門家盲検下符号化は、患者調査と標準病名マスター側 ICD-10 コードとの符号化方針差を、専門家判断という第三の視点から検討する基盤を提供した。

阿南研究は、AI や形態素解析が提供し得る用語類似度の手がかりと、ICD 分類上の意味判断との差異を試行的に示し、公的統計における説明可能性および再現性の重要性を補強する位置づけとして機能した。

（4）本研究が示す効率化

本研究における効率化は、所要時間の短縮を意味するものではない。それは、用手的確認における確認対象の重点化と、符号候補の選択に伴う判断負荷の軽減可能性として定義される。

PCMA の 3 層構造は、用手的確認における確認対象の重点化に資する構造を具体的に示した。

第 1 層（41,754 語、参照基準との一致率 93.03%）は、文字列照合と言語処理という異なる照合原理が同一の ICD-10 コードに到達した語であり、用手的確認における判断負荷の軽減可能性が相対的に高い層として位置づけられる。

第 2 層（17,621 語）は、形態素解析と PCMA の双方が候補を提示するが両者が一致しない語であり、両者の候補を併存する手がかりとして、手付け符号化に資する層である。

第 3 層（15,693 語）は、形態素解析が候補を提示しなかった語であり、このうち 15,652 語に対しては PCMA が候補を提示できた。すなわち、形態素解析が手がかりを与えなかった語の大部分に対して、標準病名マスター照合に基づく候補情報の付加が可能であることが示された。PCMA が候補を提示できなかった 41 語は、標準病名マスターを照合源とする本手法の限界として整理する。

桜澤研究、來島研究および森藤研究の成果は、辞書ファイル機能改良およびチェック機能整備の重点領域を明らかにすることにより、用手的確認対象の重点化に資する基礎情報を提供する。

なお、本研究では作業時間の実測は行っておらず、効率化の効果は実測値ではなく、用手的確認の対象重点化および判断負荷の軽減可能性として解釈される。

（5）辞書ファイル更新方針への示唆

桜澤研究の結果は、辞書ファイル機能改良の優先領域として、U コード、Z コード、呼吸器系の疾患、皮膚および皮下組織の疾患、ならびに循環器系の疾患を提示する。

來島研究の結果は、悪性新生物（C コード）に関する表記整理――上皮内癌・非浸潤癌の判別、複数病名による別疾患への分類への対応、ならびに境界悪性病変の取扱いを含む――が、辞書ファイル更新における重点課題となることを示す。また、側性表現の前処理あるいは正規化が辞書照合に資する可能性が示される一方、片側性および両側性の違いが分類に影響する疾患（変形性膝関節症、鼠径ヘルニア等）における慎重な対応の必要性が併せて示された。新型コロナウイルス感染症関連表記の標準化および同義語統合についても、辞書ファイル更新における対応の方向性として整理される。

森藤研究の結果は、既存の表記揺れ補正処理における「Z 付きコード」の基本コードへの置換ルール、感染症等における原因菌差への対応、ならび

にアンマッチ基礎表における過去修正事例の自動提示等、ルールベースによる辞書ファイル機能改良の具体的方向性を示した。

なお、本研究は、標準病名マスター側 ICD-10 コードに対する是非を論じるものではない。標準病名マスターは、医療情報システムにおける臨床用途を主目的に整備された標準的病名資源であり、その付与方針は当該目的に照らして妥当なものである。本研究の成果は、患者調査の傷病名辞書の更新に資する基礎的知見の提供にとどまる。

（6）傷病符号付けチェックプログラムへの示唆

PCMA の第 1 位候補と確定符号との一致・不一致は、傷病符号付けチェックプログラムにおける確認対象の絞り込みに資する可能性を提供する。特に、PCMA の 3 層構造に基づく重点化は、形態素解析が手がかりを提供しなかった語に対するチェックの優先順位付けに資する。

森藤研究が提示した目印コード抽出条件の精緻化、ならびにアンマッチ基礎表における過去修正事例の自動提示は、チェックプログラム精緻化に直接資する実装方向性である。

荒井・石割研究が示した符号化方針差事例の構造的特徴――術後状態における原疾患章と Z90-Z99 の選択、コロナウイルス感染症事例における U07.1 と B34.2 の選択など――は、チェックプログラムにおける符号化方針差の事前識別および評価対象としての位置づけに資する。

阿南研究が示した AI 類似度判定の可能性と限界は、傷病符号付けチェックにおける AI 活用の位置づけ――補助的手がかりとしての活用と、ICD 分類上の意味判断における人間判断あるいは追加チェック機構との組み合わせ――を明確化する基礎的知見を提供する。

（7）標準病名マスターを照合源とする意義と限界

標準病名マスターは、厚生労働省標準規格 HS005 に位置付けられた、医療情報システムにおける標準的病名資源である。また、レセプト電算処理システム用傷病名マスターと共通化が図られ、電子カルテおよび電子レセプトに広く用いられている。これらの事実は、標準病名マスターを照合源として用いることが、患者調査の傷病名辞書とは独立した外部資源に基づく第二の判定根拠の提供を可能にすることを示している。

一方で、標準病名マスター側 ICD-10 コードは、修飾語の追加、語順の差、一語違い、多部位記載、ICD 未付与、ならびに符号化方針差により、そのまま患者調査の確定符号として適用できない場合がある。本研究において確認された 1,184 語の符号化方針差は、両体系における符号化方針差を検討する必要性を示すものである。

このため、本研究の方法論的立場は、標準病名マスター側 ICD-10 コードを絶対的な参照基準として扱うものではない。標準病名マスターを照合源とする意義は、独立した外部資源に基づく候補提示にあり、その限界は、患者調査の確定符号との関係を構造的に整理することの必要性にある。

（8）AI・機械学習を主要方法としない意義

公的統計には、再現性、説明可能性および判定根拠の明示性が一貫して求められる。生成 AI および教師あり機械学習は、学習データ、推論過程、出力の安定性、ならびに判定根拠の説明可能性の各観点において、本研究の段階では、公的統計実務への適用に際して検討すべき課題が残されている。

このため、本研究班は、AI および機械学習を主要方法としては採用せず、ルールベース照合を中核に据えた。本研究班における AI に関する分担研

究（阿南研究）は、AI や形態素解析が提供し得る用語類似度の手がかりと、ICD 分類上の意味判断との差異を明らかにする位置づけで実施された。また、森藤研究におけるクラスタリング分析は、自動学習モデルの構築を目的とせず、実務上の既知の不一致傾向を論理的に整理する目的で実施された。

これらの設計上の選択は、本研究の方法論的限界ではなく、公的統計実務に必要な説明可能性および再現性を重視した方針として位置づけられる。

（9）研究の限界

本研究の限界として、以下の事項を挙げる。

第一に、本研究は主として令和 5 年患者調査データを対象としており、過年度患者調査データによる外的妥当性検証は今後の課題である。

第二に、PCMA の第 1 位候補一致率は、患者調査の確定符号との一致率であり、独立した医学的・分類学的な絶対解との一致率ではない。患者調査の確定符号は、本研究における参照基準であり、医学的・分類学的な絶対解として扱うものではない。

第三に、効率化について、作業時間の実測は行っていないため、所要時間の短縮効果は本研究では示していない。

第四に、AI 類似度判定の試行は、慢性肝疾患領域の 3 疾患を対象とした限定的な範囲のものであり、一般化にはさらなる検証を要する。

第五に、専門家による盲検下符号化は、124 症例を対象とした予備的検証であり、本格的な独立参照基準の構築には、検証規模および評価設計の拡張を要する。

第六に、標準病名マスターを照合源とする方法には意義がある一方、標準病名マスター側 ICD-10 コードを患者調査の確定符号として単純に扱うことはできない。1,184 語の符号化方針差の存在は、両体系の関係を構造的に整理する必要性を示している。

（１０）今後の課題

今後の課題として、以下の事項を挙げる。

第一に、過年度患者調査データによる外的妥当性検証を進める。

第二に、診療情報管理士複数名による独立判定と評価者間一致度に基づく、独立参照基準の構築を進める。

第三に、ICD4 桁レベル等での要因分析の深掘り、ならびに入院・外来別、病院・診療所別、地域別の層別解析を進める。

第四に、本研究で明らかとなった重点領域に基づき、辞書ファイル更新方針の具体化を進める。

第五に、PCMA および森藤研究が提示した既存プロセスの論理的改善策に基づき、傷病符号付けチェックプログラムの実装要件の整理を進める。

第六に、ICD-11 への移行を見据えて、変換テーブルの整備、ならびに判定根拠の明示性を確保するための基盤整備を進める。

第七に、令和 8 年 1 月から導入予定の次期統計処理システムへの段階的適用可能性の検討を進める。

Ｅ．結論

本研究は、患者調査における傷病名符号化プロセスを前提として、令和 5 年患者調査により得られた傷病名を試料に、5 課題の分担研究により、傷病符号付けプロセスの最適化および効率化に資す

る基礎的知見を得た。

本研究を通じて、手付け符号化を要する語の発生要因（ICD 大分類別の傾向、ICD1 桁および 3 桁分類別の分布、悪性新生物および側性表現に関する表記上の課題、ならびに新型コロナウイルス感染症関連表記の揺らぎ）、辞書判定コードと手付けコードの不一致パターン（K-means 法による 8 クラスタへの整理と、クラスタ別の代表的な不一致パターン）、標準病名マスターを照合源とする並行照合手段（PCMA および PCMI）の有用性と限界（主指標対象における参照基準との一致率 85.09%、3 層構造による確認対象の重点化、ならびに符号化方針差 1,184 語の存在）、ならびに患者調査と標準病名マスターの符号化方針差（術後状態およびコロナウイルス感染症関連事例における方針差の予備的観察、ならびに専門家盲検下符号化における多数決一致先の分布）が明らかとなった。

本研究の成果は、患者調査における傷病符号付けの適正化、辞書ファイル更新、傷病符号付けチェックプログラムの精緻化、ならびに次期統計処理システム更新方針の検討に資するものである。

本研究は、患者調査の既存プロセスを置換するものではなく、基幹統計としての継続性と信頼性を尊重しながら、符号化判断の再現性、説明可能性、ならびに確認可能性を高めるための基礎資料を提示したものである。

Ⅰ．参考文献

1）厚生労働科学研究費補助金 政策科学総合研究「患者調査の効率的な実施手法の確立に資する研究（21AB1001）」令和 3 年度報告書.

2）厚生労働科学研究費補助金 政策科学総合研究「患者調査の効率的な実施手法の確立に資する研究（21AB1001）」令和 4 年度報告書.

Ｆ．健康危険情報

該当なし。

Ｇ．研究発表

1．論文発表

学術誌投稿準備中（荒井・西本研究）。

その他の分担研究については、本年度の論文発表は該当しない。

2．学会発表

阿南研究は、2026 年度日本医療マネジメント学会学術総会において発表予定である。

その他の分担研究については、本年度の学会発表は該当しない。

Ｈ．知的財産権の出願・登録状況（予定を含む。）

1．特許取得

該当なし。

2．実用新案登録

該当なし。

3．その他

該当なし。