

厚生労働科学研究補助金

(臨床研究等 ICT 基盤構築・人工知能研究事業) 総括研究報告書

電子カルテ情報をセマンティクス (意味・内容) の 標準化により分析可能なデータに変換するための研究

堀口 裕正 国立病院機構本部総合研究センター 診療情報分析部 副部長

岡田 千春 国立病院機構本部 企画役

狩野 芳伸 静岡大学情報学部行動情報学科 准教授

森田 瑞樹 岡山大学大学院医歯薬学総合研究科 准教授

奥村 貴史 国立保健医療科学院研究情報支援研究センター 特命上席主任研究官

平成 28-29 年度においては、我々が日本語における医療用自然言語処理の研究コミュニティを形成し研究に取り組んで来た標準化技術を実カルテへと適用することで、カルテからの情報抽出の自動化に向けた予備的な検証を行うことを計画した。

研究代表者堀口及び分担研究者岡田は、国立病院機構本部との調整を中心とした基盤構築を行った。まず、NDCA データの研究利用に向け、倫理審査申請に加えて、内部規定にて定められている内部委員会の調整を図った。また、閲覧・解析に特化した自然言語処理用の研究基盤の構築を行った。研究基盤の概念図は図に示したとおりで、セキュリティを維持しつつ、空間的制約をなるべく少なく研究が進められるようなものになっている。

研究分担奥村は、臨床的なニーズを自然言語処理における個別技術へと橋渡しする役割を担った。具体的には、臨床医側より退院サマリの自動生成に求められる要件定義を進めるとともに、先行研究の整理を行い今後の研究アプローチの策定を行った。さらに、今後の研究に役立てられる入院カルテ・退院サマリの高品質な個人情報を含まない模擬のデータセットを構築した。

研究分担狩野・森田は、上記の実データ・テスト用データ双方を活用し、自然言語処理の医療テキストへの適用を進めた。狩野は、時系列で蓄積してく入院カルテデータを対象として、既存の自然言語処理ツールによる処理性能と今後の改良に向けた課題抽出を図った。森田は、医師が要約した退院サマリデータを対象として、医師の記載する退院サマリの定量的・定性的な特徴の把握を図った。この知見は、今後、入院カルテの自動要約技術の研究に際した精度管理に役立てられる。

平成 30 年度については研究の最終年度として以下のテーマに取り組んだ。

研究代表者堀口及び分担研究者岡田は、他のため研究者の分析に資するため、国立病院機構本部との調整を中心としたセキュリティを維持しつつ、空間的制約をなるべく少なく研究が進められるような基盤構築・運用を行った。

また研究代表者堀口は、本研究の当初の目的である、汎用的にどの電子カルテからで

も利用できるシステムの構築という課題に対して、JAHIS の HL 7 CDA 規約に基づくデータを SS-MIX 拡張ストレージに保管するという現在日本における標準的なデータ交換フォーマットを起点として、特定の患者の医師記録を抽出し、そこからサマリを作成するフローを実現するモジュールの開発を行った。SS-MIX ストレージを活用したデータ保管システムはすでに国内の大手ベンダー 8 社においては NCDA 上で実現されており、横展開可能なシステムの構築という本研究に求められたミッションの 1 つは達成できたと考えている。

研究分担者奥村は、入院カルテの自動要約に向けた研究基盤の整備に取り組んできた。臨床医は、入院治療をしていた患者が退院する際、それまでに記載していた入院カルテから退院サマリを作成する。この退院サマリの作成を効率化することができれば、医師の診療負担を直接軽減することが出来ることに加えて、様々な副次的な効果が期待される。

今年度は、昨年度までの研究をさらに発展させ、4 つの課題に取り組んだ。まず、一連の研究には自由に研究利用できるカルテの存在が必要となる。そこで、ダミーカルテの収集を進め、100 件の整備を目指して活動を進めた。また、収集したダミーカルテを対象として、機械的な処理を可能とするためのアノテーション作業に取り組んだ。その際、初年度に行ったアノテーション、昨年度に試みたアノテーションを踏まえ、さらなる改善を図った。次に、このアノテーション作業と平行して、退院サマリの分析モデルであり生成モデルである「CASE モデル」の改善に取り組んだ。また、退院サマリの自動生成処理を精度管理するうえで必要となる「理想の退院サマリ」の確保に向けた検討を行った。

研究活動の結果、ダミーカルテは、目標を超える 108 件を集めることが出来た。また、これらのカルテを対象としたアノテーションを進めると共に、アノテーションガイドラインを高品質化することが出来た。さらに、このアノテーション済みカルテを用いて実現する退院サマリの分析に向けたモデルとして、修正 CASE モデルを提示することが出来た。理想のサマリに向けた検討では、高品質な退院サマリを低コストに実現するための作業仮説を整理すると共に、実証に向けた研究デザインを策定することが出来た。

研究分担者森田は、退院サマリの自動生成技術の実現を目指し、昨年度までの調査結果の整理およびそれを踏まえた退院サマリの自動評価のコンセプト検証を実施した。退院サマリを生成および評価するためには、どのような退院サマリを生成しなくてはならないかを示す「理想的な退院サマリ」の定義が必要となる。昨年度に理想的な退院サマリについて言及をした国内外の文献調査を実施し、今年度はその結果を整理した。退院サマリの記載に関して「記載すべき項目」および「定性的な要件事項」を抽出し、前者として 38 項目を得た。このうちの 14 項目はカルテの構造化データより抽出できるものの、残りの 24 項目はカルテの自由記載より抽出してサマリとして文章を作成や要約をする必要があるものと考えられた。また、後者として 9 項目を得た。このうちの 5 項目は医学的な知識・経験がないと採点が難しいものの、残りの 4 項目は形式的に判断をすることが可能と考えられた。この 4 項目のうち医学用語辞書を必要としない 3 項目について文章の特徴を用いて自動評価することを試み、人による評価との相関および点数の分布を踏まえて自動評価の可能性を考察した。退院サマリの自動評価手法の確立に向けた今後の課題が明らかとなった。

研究分担者狩野は、退院サマリの自動生成に向けたテキストの分析についての研究を

行った。入院患者の退院に際し、医師は入院中に記載したカルテ等の情報から退院サマリを作成する必要がある。この退院サマリを自動的に生成することが出来れば、臨床現場の負担を下げる事が出来ると共に、医療の質に貢献することが期待される。カルテの処理にあたっては、事前に匿名化が必要となる。匿名化作業を自動化するための匿名化ツールの実装と性能向上に取り組んだ。そのために、既存の正解付き模擬カルテデータに加え、別のダミーカルテデータセットに対し匿名化のためのアノテーション付与を行い、これらを用いてルールベースおよび機械学習による匿名化ツールの実装と性能検証を行った。

サマリ生成にあたっては、対象とするカルテやサマリのドメイン、すなわち診療科や疾患により、サマリ生成に必要な情報が異なると考えられる。サマリと対応する電子カルテの履歴データについてクラスタリングを行い、どのようなタイプのサマリやカルテがどう類似しているかの分析を行った。

A. 目的

本研究は、電子カルテに実装可能な医療用語の標準化を行うシステムの開発に向け、そのステップとして、臨床現場からのニーズが極めて大きい退院サマリの自動生成技術の実現を目標と定める。電子カルテの自動解析は技術的な難易度が高く、実用的な精度を実現するためには多額の研究開発投資が求められる。そこで、本研究提案では、医療現場に直接的なメリットが生じる研究課題に取り組むことによって、現場の協力と今後の追加的な研究開発投資を呼び込み、その過程を通じて実用性の高い電子カルテの自動解析技術を実現する戦略を採る。初年度、我々が今まで模擬カルテを用いて研究開発を進めてきた標準化技術を、国立病院機構の有する広域電子カルテ網(NCDA)上の実カルテへと適用し、技術的な課題を抽出する。2年目には、NCDAを用いて集積した電子カルテに加えて、退院サマリ情報を用いることで、電子カルテの自動要約技術の検討を行う。3年目においては、両技術の統合により、継続的な精度向上の体制を実現するとともに、研究成果を既存の各社電子カルテへと組み込む枠組みを構築する。本研究により、退院サマリの自動要約技術や紹介状の作成支援技術等、医療用の自然言語処理に関連する多彩な応用技術が実現する。これは、医療現場における負担軽減策として極めて効果が期待される。また、こうした応用の発展により、要素技術である電子カルテ上の記載からの自動情報抽出において、継続的な精度向上が実現する。この手法は、電子カルテにおける用語の標

準化技術単独に研究開発投資を行うことと比して、投資効率が極めて高いと考えられる。さらに、こうして医療用自然言語処理技術が発展することにより、大量の電子カルテからの効率的な情報抽出が実現する。これは健康医療政策に資する統計データの収集コストを劇的に低廉化し、今後、政策に求められる様々なエビデンスを継続的に生み出していく基盤となることが期待される。

B. 方法

本研究の今年度申請時に記載した研究計画および方法は以下のとおりであった。

【取り組む課題】

本研究では、電子カルテの記載より標準化した情報の抽出を行う技術の研究開発を行う。たとえば、カルテに「精神遅滞」と記載されている場合、この語を解釈することは機械にとって容易である。しかし、カルテの記載は単純ではない。「MR」と記載されていた場合、その語が精神遅滞(Mental Retardation)を指すのか、僧帽弁閉鎖不全症(Mitral Regurgitation)か、製薬会社の営業担当(Medical Representative)かを正確に判別することは機械にとって容易ではない。さらに、「学業成績が芳しくない」と記載されたとき、医学的には文脈より mild MR と解する柔軟性も求められる。カルテ記載は、往々にして略記法や省略が多用される。「腹部：圧痛、反跳痛なし」という

記載から、圧痛の有無を判別することは、高度な処理を要する。医療用自然言語処理はこのように難度の高い処理であり、現時点では情報抽出の精度に限界があり、研究開発体制におけるブレークスルーが求められている。

【研究体制】

研究代表者 堀口は、「大規模に電子カルテデータを入手できる体制」として、全国の国立病院 41 施設より年間 80 万患者の電子カルテ情報を自動収集する診療情報集積基盤 (NCDA) を構築し、運用している。この研究基盤を用いることにより、大学病院等のように患者の偏りが生じない多彩な施設からカルテ情報を収集することが出来る。

研究分担者 狩野・森田は、日本語カルテに記載された所見や病名を自動的に標準化するコンテスト (NTCIR MedNLP) を主催する、医療用自然言語処理における我が国の代表的な研究者である。このコンテストは、日本語カルテからの情報抽出精度を競う唯一の学術集会であり、森田は参加者としても首位の成績を収めて来た。狩野は、コンピュータによる大学入試問題の自動解答に向けた人工知能研究「ロボットは東大に入れるか」プロジェクトにおいて、社会科分野の科目担当リーダーを務めている。両名は、公募研究課題の求める「工学系人工知能研究者・自然言語処理研究者」の役割を果たす。

分担研究者 奥村は、我が国においては数少ない計算機科学分野において学位を

取得した医師であり、自然言語処理の医療応用分野で継続的な貢献を果たしてきた。本研究において奥村は、「病名・治療の標準化に詳しい臨床医学系研究者」として、基礎技術研究と応用研究とを橋渡しするコーディネータ役を担う。

分担研究者 岡田は、本研究の対象となる各国立病院カルテを記載する臨床医を取りまとめ、研究開発成果の臨床的有用性の担保に取り組む。

【H28 年度実施状況】

初年度においては、上述の NTCIR MedNLP コンテストを中心に研究開発を進めてきた標準化 (正規化) 技術を、NCDA 上の実カルテへと適用し、予備的な検証を行うこととしていた。実際には

1, 実データは利用場所が限られている等、研究利用に対して制限が多いことから、実際の診療を行っている医師の協力を得て、実際の診療経過に沿ったダミーデータを作成し、各種研究開発を加速する為の基盤を作成した。

2, 実際の分析を効率的に行うため、セキュリティを考慮した分析環境の整備/構築を行った。

3, 退院サマリと、電子カルテの医師記録についての関係について、現場医師から情報提供を受け予備的な検討を行った。

【H29 年度計画】

2 年目においては、NCDA を用いて集積した電子カルテに及び、退院サマリデータを対象とした電子カルテの自動要約技術の検証を行う。平成 28 年度に構築した分析環境基盤を活用して、まずは、医師記録や検査

データ等入院期間内に生成された情報と、退院サマリデータの差分の検証を行い、医師がどのような思考経路を経て情報処理を行い、文書を作成しているかについての検討を行う。その際、一般的な文章を対象とした自動要約技術の応用に加えて、医療文書を対象とする正規化技術を活用した疾患や病期毎のテンプレートを用いた情報抽出を試みる。この過程は、医療従事者による詳細な手本(教師データ)を用いることで精度向上が望める。その際、平成28年度に作成したダミーデータを活用し、本番データの分析環境と平行して各々の施設内での作業を並行して行っていくことで研究活動の効率化を図る。

【H30年度計画】

最終年度においては、それまでに培った知見を取りまとめるとともに、上記の過程を組織化し、継続的な精度向上の体制を実現する。また、研究成果の電子カルテへの組み込みもしくは SS-MIX ストレージを活用した1部門システム化を目指し、各種ベンダーのシステムへの退院サマリ自動作成支援システムの適用可能なモジュール化を試みる。

この研究計画を踏まえ我々が日本語における医療用自然言語処理の研究コミュニティを形成し研究に取り組んで来た標準化技術を実カルテへと適用することで、カルテからの情報抽出の自動化に向けた予備的な検証を行うことを計画した。

その計画の実現に向け、採択後研究者全員で組織する「総括・企画調整班」を作り、平成28-29年度においては月2回程度

のミーティングを行い、研究の方向性の整理及び各分担班の作成及び役割の決定、進捗の管理及び調整を行う方法で研究を遂行することとした。また、「総括・企画調整班」以外の分担班についてはそれぞれ責任研究者を決め、その裁量で研究を進める方法をとった。

平成30年度は、昨年度まで実施していた研究者全員で組織する「総括・企画調整班」の機能を縮小し、それぞれの分担班においてそれぞれ責任研究者を決め、その裁量で研究を進める方法を中心として運営を選択した。

(尚、総括・企画調整班以外の分担班については総括・企画調整班の活動結果から生まれたものであり、それぞれの班の目的・方法についてもC.結果セクションで記載することとする。)

C.結果

1. 総括・企画調整班

まず、本研究班はその応募要項の段階からデータ収集に掛かる部分については研究の中に組み込まないことを求められており、データの収集基盤の構築・運営については本研究のカバーする範囲ではない。しかしながら、本研究の前提となる国立病院機構が作成・維持運営するNCDAについて本報告書でその概要や意義について記述を行わないとするならば、本報告書の内容の理解に大きな妨げになると考えここに報告を行うこととする。

NCDAの概要については参考資料1にその概要を資料を添付した。また、実際の病院におけるSS-MIXデータ作成に掛かるシステム仕様についても参考資料2に示した。これらの仕様等のドキュメントについては

その改版履歴も含め、github 上で管理、公開している。

https://github.com/nhoHQ/SSMIX2_support_documents

次に、実際の本分担研究班の活動についての報告を行う。本研究班においては、まずは研究者が独立して研究活動を進めるのではなく、10回の研究班会議（うち8回はWeb 音声会議）を行い、1つの有機的な研究班として活動が行える環境で運営してきた。

各班会議での調整事項は以下の通りである。

第1回

- 研究管理面の話題
- 各人の状況 update
- 仮説構築作業
- 「退院サマリとは何か？」

第2回

- 作業仮説構築
- ツールドリブン/リソースドリブン/臨床ニーズ/病院管理ニーズからの整理
- 倫理審査に向けた論点整理
- 病院訪問に向けた調整

第3回

- 三重病院にて退院サマリの記載内容について臨床家とともにディスカッション

第4回

- 今後のスケジュール確認
- 三重病院訪問での成果確認
- 倫理審査に向けた調整
- 研究分担の整理
- 「良質な退院時サマリとは？」問題の整理

第5回

- ダミーカルテ作成の是非
- 継続申請書類の作成について
- 倫理審査の申請書確定について
- 医師アンケート企画について
- テストデータについて
- 解析のアプローチについて

第6回

- H29 継続申請について
- H28 倫理審査の状況報告
- 年度内達成目標の再確認
- 各分担研究状況報告
- テストデータについて

第7回

- 各分担研究状況報告
- 研究基盤の整理
- 倫理委員会・利活用審査委員会の報告

第8回

- 各分担研究状況報告
- 退院サマリの関する文献サーベイについて
- カルテ要約の要素技術についての議論

第9回

- 各分担研究状況報告
- 退院サマリの関する文献サーベイについて
- 報告書作成について

第10回

- 分析環境の整備について報告

第11回

- 各分担研究状況報告
- 報告書作成について
- 平成 29 年度計画の確認

第12回

- 各自状況報告
- 研究成果報告書
- 記録の保管について

第 13 回

- 各自状況報告
- 研究成果報告書
- 年間スケジュール

第 14 回

- CLEF シェアードタスク
- 荒牧班との共同研究班会議について
- 情報共有／相談
- 研究分担の整理
- 「良質な退院時サマリとは？」問題の整理
- ダミーカルテ作成の是非

第 15 回

- 匿名化について
- 荒牧班との共同研究班会議
- 各自状況報告

第 16 回

- 匿名化について
- 各自状況報告
- 年度内達成目標の再確認
- 各分担研究状況報告
- テストデータについて

第 17 回

- 各分担研究状況報告
- 研究基盤の整理

第 18 回

- 各分担研究状況報告
- 研究発表の場について
- スケジュール確認

第 19 回

- 各分担研究状況報告
- 報告書作成について
- 分析環境の整備について報告

第 20 回

- 各分担研究状況報告
- 報告書作成について

なお、NCDA データは国立病院機構が契約するデータセンター内で厳重に管理されている。研究に際しては、このデータベースから研究テーマごとに匿名化したサブセットを切り出し、国立病院機構本部内のオンサイト利用に限っている。以上により、データセットの利用対象と利用目的を厳しく制限することにより、患者個人情報の保護を行っている。それに対応する分析基盤の作成に関して、分担研究班を組織し、堀口・岡田が責任者として活動を行うこととした。(分担研究の結果は後述)。また、最終年度においてはこの分担班で「部門システム化を目指したモジュール開発」のタスクを担当することとした。

また、「退院サマリの自動生成に向けたアプローチの検討」というテーマの分担研究を奥村が、「退院サマリの自由記載文の特徴解析」というテーマの分担研究を森田が、「退院サマリの自動生成に向けたアプローチの検討」というテーマの分担研究を狩野が担当することとした。

2. SS-MIX2 分析用データセットの作成・開発班

研究代表者堀口及び分担研究者岡田は、国立病院機構本部との調整を中心とした基盤構築を行った。初年度 NDCA データの研究利用に向け、倫理審査申請に加えて、内部規定にて定められている内部委員会の調整を図りともに承認を得た。また、閲覧・解

析に特化した自然言語処理用の研究基盤の構築を行った。研究基盤の概念図は図1に示したとおりで、セキュリティを維持しつつ、空間的制約をなるべく少なく研究が進められるようなものになっている。

また、本研究で中心的に使われる医師記録等（経過記録・退院サマリ）については、SS-MIX2の標準仕様に含まれていないが、JAHISの提供している仕様を参考に、資料1で示した仕様でNCDA内に実装することとした。

汎用的にどの電子カルテからでも利用できるシステムの構築という課題に対して、JAHISのHL7CDA規約に基づくデータをSS-MIX 拡張ストレージに保管するという現在日本における標準的なデータ交換フォーマットを起点として、特定の患者の医師記録を抽出し、そこからサマリを作成するフローを実現するモジュールの開発を行った。SS-MIX ストレージを活用したデータ保管システムはすでに国内の大手ベンダー8社においてはNCDA上で実現されており、横展開可能なシステムの構築という本研究に求められたミッションの1つは達成できたと考えている。

3. 退院サマリの自動生成に向けたアプローチの検討班

入院カルテの自動要約に向けた研究基盤の整備に取り組んできた。臨床医は、入院治療をしていた患者が退院する際、それまでに記載していた入院カルテから退院サマリを作成する。この退院サマリの作成を効率化することができれば、医師の診療負担を直接軽減することが出来ることに加えて、

様々な副次的な効果が期待される。

今年度は、昨年度までの研究をさらに発展させ、4つの課題に取り組んだ。まず、一連の研究には自由に研究利用できるカルテの存在が必要となる。そこで、ダミーカルテの収集を進め、100件の整備を目指して活動を進めた。また、収集したダミーカルテを対象として、機械的な処理を可能とするためのアノテーション作業に取り組んだ。その際、初年度に行ったアノテーション、2年度に試みたアノテーションを踏まえ、3年目にさらなる改善を図った。次に、このアノテーション作業と平行して、退院サマリの分析モデルであり生成モデルである「CASEモデル」の改善に取り組んだ。また、退院サマリの自動生成処理を精度管理するうえで必要となる「理想の退院サマリ」の確保に向けた検討を行った。

研究活動の結果、ダミーカルテは、目標を超える108件を集めることが出来た。また、これらのカルテを対象としたアノテーションを進めると共に、アノテーションガイドラインを高品質化することが出来た。さらに、このアノテーション済みカルテを用いて実現する退院サマリの分析に向けたモデルとして、修正CASEモデルを提示することが出来た。理想のサマリに向けた検討では、高品質な退院サマリを低コストに実現するための作業仮説を整理すると共に、実証に向けた研究デザインを策定することが出来た。

4. 退院サマリの自由記載文の特徴解析班

退院サマリの自動生成技術の実現を目指し、昨年度までの調査結果の整理およびそれを踏まえた退院サマリの自動評価のコンセプト

ト検証を実施した。退院サマリを生成および評価するためには、どのような退院サマリを生成しなくてはならないかを示す「理想的な退院サマリ」の定義が必要となる。昨年度に理想的な退院サマリについて言及をした国内外の文献調査を実施し、今年度はその結果を整理した。退院サマリの記載に関して「記載すべき項目」および「定性的な要件事項」を抽出し、前者として38項目を得た。このうちの14項目はカルテの構造化データより抽出できるものの、残りの24項目はカルテの自由記載より抽出してサマリとして文章を作成や要約をする必要があるものと考えられた。また、後者として9項目を得た。このうちの5項目は医学的な知識・経験がないと採点が難しいものの、残りの4項目は形式的に判断をすることが可能と考えられた。この4項目のうち医学用語辞書を必要としない3項目について文章の特徴を用いて自動評価することを試み、人による評価との相関および点数の分布を踏まえて自動評価の可能性を考察した。退院サマリの自動評価手法の確立に向けた今後の課題が明らかとなった。

5. 退院サマリの自動生成に向けたアプローチの検討班

退院サマリの自動生成に向けたテキストの分析についての研究を行った。入院患者の退院に際し、医師は入院中に記載したカルテ等の情報から退院サマリを作成する必要がある。この退院サマリを自動的に生成することが出来れば、臨床現場の負担を下げることが出来ると共に、医療の質に貢献することが期待される。

カルテの処理にあたっては、事前に匿名化が必要となる。匿名化作業を自動化するための匿名化ツールの実装と性能向上に取り組んだ。そのために、既存の正解付き模擬カルテデータに加え、別のダミーカルテデータセットに対し匿名化のためのアノテーション付与を行い、これらを用いてルールベースおよび機械学習による匿名化ツールの実装と性能検証を行った。

サマリ生成にあたっては、対象とするカルテやサマリのドメイン、すなわち診療科や疾患により、サマリ生成に必要な情報が異なると考えられる。サマリと対応する電子カルテの履歴データについてクラスタリングを行い、どのようなタイプのサマリやカルテがどう類似しうるかの分析を行った。

退院サマリの自動生成にあたっては、**extractive** な処理を行うこととし、電子カルテ内の各文についてサマリに含めるべきか否かの判断を行った。判定には、文末表現に着目する手法、文ベクトルを生成して文間の類似度で決定する方法などいくつかの手法を行い、文一致率、単語一致率、**ROUGE** など異なる指標での評価を行った。結果、いずれの手法でもある程度の一致率を達成しうるということがわかった。今後、より実用的な性能の達成のためには、表記揺れの吸収、さらに大規模なデータの利用による学習性能の向上などの研究が考えられる。

D・E. 考察及び結論

医療用情報システムの研究開発においては、医療現場に直接の恩恵が及ばないゴー

ルが設定されることで、研究開発が現場のニーズから乖離するとともに、継続した開発投資に繋がらない悪循環が往々にして生じてきた。本研究提案は、医療現場における負担軽減策として期待が大きい退院サマリの自動要約技術の開発を目指す。これにより、紹介状の自動作成技術等、電子カルテの自動解析技術に関連する継続的な研究開発投資の実現が期待される。この研究開発サイクルを確立することにより、要素技術である電子カルテ上の記載から自動情報抽出における継続的な精度向上が期待される。

こうして確立する医療用自然言語処理技術は、大量の電子カルテからの効率的な情報抽出を実現し、健康医療政策に資する統計データの収集コストを劇的に低廉化することが期待される。とりわけ、様々な傷病や治療に関して、既存のDPCやレセプトには表れてこない深遠な実態を明らかとし、医療の質向上・均てん化・各種医療技術の臨床開発に必要なエビデンスを生み出すことが期待される。

また、国立病院機構の有する広域電子カルテ網は、各病院が独自に調達した電子カルテベンダー主要8社を網羅している。本研究によって、大口顧客としての交渉力を背景としたこれら主要ベンダーへの研究開発成果の技術移転が期待される。これらは、医療の情報化を進める厚生労働行政にとって、新たな政策手段の実現をもたらす

加えて、本研究で作成した入院カルテ・退院サマリの高品質な個人情報を含まない模擬のデータセットは、研究成果として公表することにしており、電子カルテの現物を持たない情報系研究者が、容易に本領域

の発展に貢献できる環境を提供出来る点も大きな成果になるものと考えている。

今年度の研究においては、昨年度に引き続きサマリ作成についての各種技術の研究を着実に進めたとともに、本研究の当初の目的である、汎用的にどの電子カルテからでも利用できるシステムの構築という課題に対して、JAHISのHL7CDA規約に基づくデータをSS-MIX拡張ストレージに保管するという現在日本における標準的なデータ交換フォーマットを起点として、特定の患者の医師記録を抽出し、そこからサマリを作成するフローを実現するモジュールの開発を行った。SS-MIXストレージを活用したデータ保管システムはすでに国内の大手ベンダー8社においてはNCDA上で実現されており、横展開可能なシステムの構築という本研究に求められたミッションの1つは達成できたと考えている

F. 研究発表

1. 論文発表

古崎晃司,堀口裕正,奥村貴史,津本周作:
OS-27 人工知能の医療応用 人工知能
33(6):843-848 2018

堀口裕正: 国立病院機構のデータベースを用いた臨床研究 *Progress in Medicine*
38(2):17-20 2018

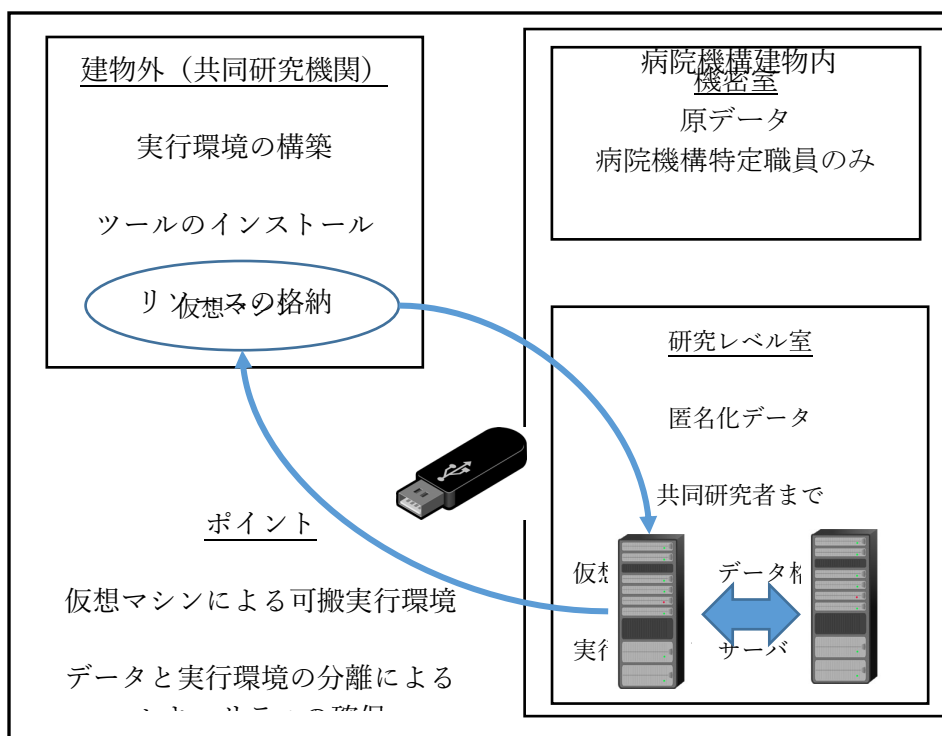
堀口裕正: NCDAの現況と成果および今後の展望 月刊新医療 2018年2月号

2. 学会発表

森田瑞樹, 奥村貴史, 狩野芳伸, 堀口裕正.
退院サマリの自由記載は何を書くこと
が望ましいのか: 文献レビュー, 第38回
医療情報学連合大会, 2018年11月22~25
日, 福岡.

[Kajiyama 18] Kajiyama, K. Horiguchi, H.
Okumura, T. Morita, M. Kano, Y.,
2018. De-identifying Free Text of
Japanese Dummy Electronic Health
Records. The Ninth International
Workshop on Health Text Mining and
Information Analysis(LOUHI2018) (p.
65).

図 1



資料 1 NCD A における医師記録等の仕様書

趣旨

本事業では、各社の **SS-MIX2** モジュールの拡張ストレージへの出力機能を利用し、以下の情報を出力することを求めている。その際、**SS-MIX2** 拡張ストレージ構成の説明と構築ガイドライン Ver.1.2d (以下、ガイドライン) に記載している仕様に対応していること。また、トランザクションストレージ、インデックスデータベースも同時に生成すること。

- 経過記録
- 退院時サマリー
- 診療情報提供書

以下に仕様を示す。

ドキュメントデータ 物理構造

```
|-- 拡張ストレージ ルートフォルダ
  |-- 患者 ID 先頭 3 文字
    |-- 患者 ID 4～6 文字
      |-- 患者 ID
        |-- 診療日
          |-- データ種別
            |-- コンテンツフォルダ
              |-- 主文書ファイル
```

診療日

特に指定しない。

データ種別

ガイドライン P4 (4) 「データ種別フォルダ」について に則ること。

```
[ローカル文書コード]^ローカル文書名称^[ローカルコード体系コード]^標準文書コード^
標準文書名称^標準コード体系コード
```

以下のように標準コードに対しローカルコードが複数あることは許容される。

```
L12345^入院診療録^99ZZZ^11506-3^経過記録^LN
```

```
L12346^外来診療録^99ZZZ^11506-3^経過記録^LN
```

コンテンツフォルダ

ガイドライン Ver.1.2d P5 (5) 「コンテンツフォルダ」について に則ること。

```
患者 ID_診療日_データ種別コード_特定キー_発生日時_診療科コード_コンディションフ
ラグ
```

いずれの文書も削除は想定していないが、電子カルテシステムによっては修正はあり得ると考える。その場合、ガイドライン P6 ④修正が発生する場合 に則り改版すること。

主文書ファイル

XML CDA R2 で出力すること。XML ファイル以外に画像ファイルや CSS ファイル等を出
力してもかまわない。

HEADER 部

いずれの文書も JAHIS 診療文書構造化記述規約 共通編 Ver.1.0 に則ること。

P27 6.3.11.検査・診療等行為 "documentationOf/ServiceEvent" によると、documentationOf
の制約・多重度は 0..1 となっているが、経過記録、退院時サマリについてはこれを 1..1 と
読み替えること。

経過記録は serviceEvent classCode(サービスイベントクラスコード)を ENC(診察)とし、
effectiveTime(実施日)は low value、high value とともに記録タイミングを出力すること。

退院時サマリは serviceEvent classCode(サービスイベントクラスコード)を ACCM(入院、
滞在)とし、effectiveTime(実施日)は low value に入院タイミング、high value に退院タイミ
ングを出力すること。

タイミングの粒度は日以上であれば良い。

BODY 部

診療情報提供書は、日本 HL7 協会 患者診療情報提供書 規格 Ver.1.00 に則ること。

診療情報提供書以外は、XML の文法に則ること。

参考資料

1. NCDA データベースの説明資料



国立病院機構 診療情報集積基盤について

～電子カルテデータの標準化から利活用へ～

NCD
NHO Clinical Data Archives

Contents

1. NCDAとは
2. NCDAの経緯
3. NCDAの現状
4. SS-MIX2変換プログラムの構成
5. 保有するデータ種別
6. 構築・運用費用について
7. NCDAの利活用について
8. NCDAにおける個人情報の取扱い
9. NCDAを活用した研究例
10. MIAについて
11. 本事業の今後

1. NCDAとは



国立病院機構のデータベースの特徴

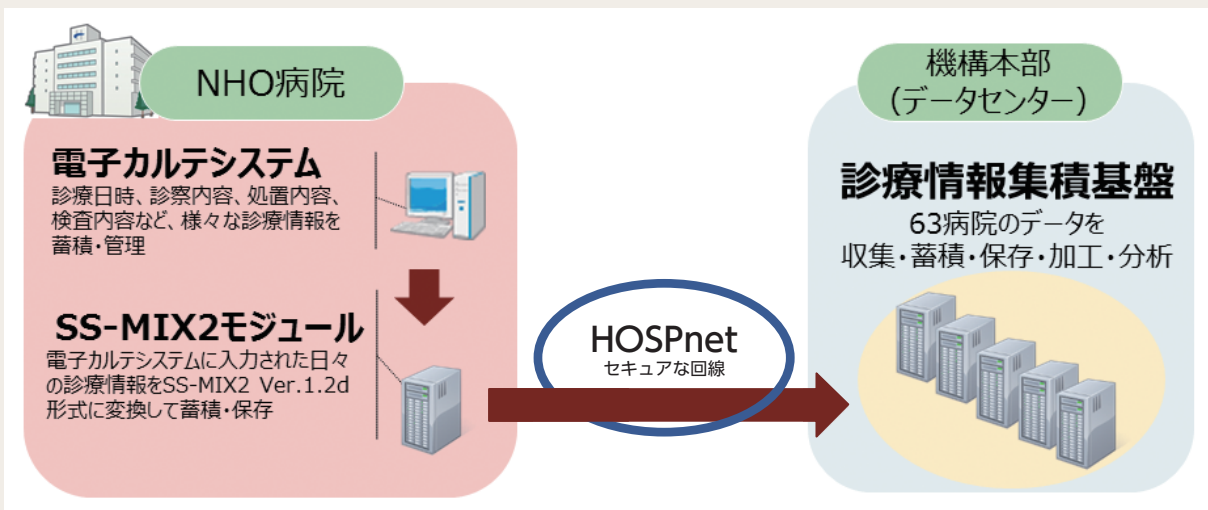
国立病院機構では、厚生労働省より平成26年度地域診療情報連携推進費補助金をうけて『電子カルテデータ標準化等のためのIT基盤構築事業』を実施しています。当事業では、SS-MIX2規格を利用し、データの標準化に最大限取り組んだ2次利用可能なデータベースとして国立病院機構診療情報集積基盤（NCDA：NHO Clinical Data Archives、平成30年度末で63病院が参加）を構築すると共に、SS-MIX2規格を用いて電子カルテデータの標準化を進め、その工程を示したドキュメント（手順書）を作成・公開しています。

NCDAの特徴は、以下のとおり3つ挙げられます。

- ①平成28年1月以降のSS-MIX2規格（標準化ストレージ機能）に含まれる全データ種別（病名・入退院・転棟・外来来院・食事・処方・投薬・検査）のデータを保有。
- ②入院患者のバイタルサインデータ（血圧・体温・心拍数）を保有。
- ③国立病院機構では単一の法人格によって複数の病院が運営されていることから、利活用におけるデータ提供の際には匿名化が行われているが、収集・データベース化の段階では匿名化せずに患者番号や生年月日・住所といった個人識別子を含んだ状態で運用。

例えばNCDAとDPCデータやレセプトデータとのデータ結合や病院における追加調査の実施等において技術的な制約がない点が特徴としてあげられます。加えて、IT基盤構築事業の結果、標準化されて各病院で同じマスターを利用し、数値が数値として認識でき、数値の単位も標準化されたデータベースとなっている点も大きな特徴です。

本パンフレットは、平成30年度末現在のNCDAの特徴と現況について説明を行っています。この資料が皆様のNCDAへの理解と利活用に関する興味を持っていただくことの一助になることを期待しております。



2. NCDAの経緯

年表

平成26年	SS-MIX2データの収集に向けて、パイロットスタディを4病院で開始
平成26年3月	パイロットスタディ参加病院を12病院まで拡大
平成26年4月	国立病院機構第3期中期計画開始（「電子カルテ情報の収集・分析について具体的な検討を進め、臨床研究等のIT基盤の充実を図る」ことを明記） 電子カルテ情報の収集・分析について具体的な検討を開始
平成27年2月	平成26年度地域診療情報連携推進費補助金「電子カルテデータ標準化等のためのIT基盤構築事業」開始
平成28年3月	IT基盤（NCDA）構築事業完了 6ベンダ41病院でNCDAを運用開始

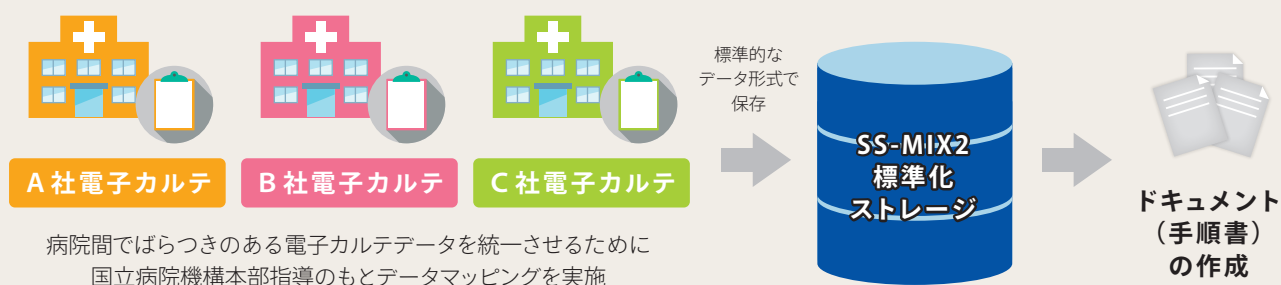
電子カルテデータ標準化等のためのIT基盤構築事業 平成26年度半ば～平成27年度末

電子カルテについては、ベンダごとで開発が行われ、各病院が使いやすいようにカスタマイズされるなど、電子カルテデータの形式が標準化されないまま普及したことから、電子カルテ上で使用されている病名や医薬品等のコードがベンダや病院で異なり、標準化の課題となっている。

本事業は、上記のような問題を解消するため、厚生労働省から平成26年度地域診療情報連携推進費補助金として事業を付託されたものであり、各病院の電子カルテデータを厚生労働省の定める標準コードに紐付けするデータマッピングを行い、SS-MIX2規格（標準化ストレージ機能）を用いて電子カルテデータの標準化を実施するとともに、その工程を示したドキュメント（手順書）を作成・公開することを目的としている。

事業内容

主要なベンダや多くの疾病領域について対応可能な精度の高いドキュメントを作成するために、41病院で電子カルテデータ標準化事業を行う。





平成28年11月	国立病院機構診療情報データベース利活用規程制定 平成28年度地域診療情報連携推進費補助金「電子カルテによる『災害診療記録』電子フォーマット自動出力実証事業」開始
平成29年8月	災害診療記録事業に7ベンダ55病院が参加
平成30年3月	災害診療記録事業完了 NCDAの運用は8ベンダ58病院まで拡大
平成31年3月	NCDAの運用は8ベンダ63病院まで拡大

電子カルテによる『災害診療記録』電子フォーマット自動出力実証事業 平成28年度半ば～平成29年度末

大規模災害時において、災害対策本部（都道府県）が被災地の医療概況を把握し、適確な医療支援活動を展開するうえで、極めて重要な情報は「疾病別症例数」等の集計情報であるが、それを迅速に集計する手法の確立が課題となっている。

この課題に対し、東日本大震災を契機に「災害時の診療録のあり方に関する合同委員会*」が設置され、災害時の標準的記録フォームといえる「災害診療記録」が作成され、平成28年熊本地震で初めて運用開始された。本事業はこの電子フォーマットの電子カルテへの実装と収集に係る実証を行うことを目的とする。

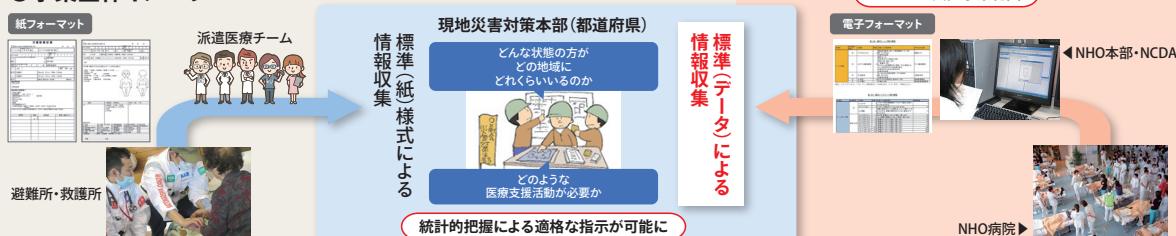
平成29年度末時点で、NCDA参加病院のうち災害拠点病院を中心に55病院で本モジュールを導入済みである。

*日本医師会、日本病院会、日本集団災害医学会、日本救急医学会などが参加

事業内容

NCDAの標準化機能を活かして、様々なベンダの電子カルテから災害診療記録用の電子フォーマットの出力が可能となるよう対応モジュールをバージョンアップし、災害時に必要な診療情報の自動抽出化等の開発及び検証を行い、その結果を導入手順書として公開することを通じて、災害発生時の適確な医療支援活動の展開に役立つもの。

●事業全体イメージ



3. NCDAの現状

NCDA参加病院の状況 (平成31年3月現在)

NCDA参加病院

63病院

うち災害診療記録事業参加
60病院

北海道

(北海道) 北海道がん、北海道医療、函館、旭川医療、帯広
(青森) 弘前

東北

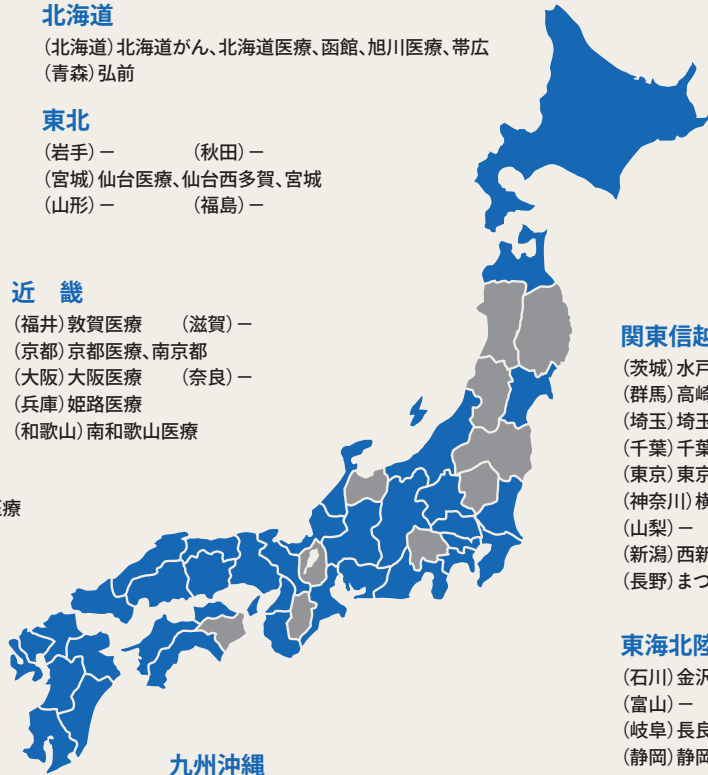
(岩手) - (秋田) -
(宮城) 仙台医療、仙台西多賀、宮城
(山形) - (福島) -

近畿

(福井) 敦賀医療 (滋賀) -
(京都) 京都医療、南京都
(大阪) 大阪医療 (奈良) -
(兵庫) 姫路医療
(和歌山) 南和歌山医療

中国

(鳥取) 米子医療 (島根) 松江医療
(岡山) 岡山医療
(広島) 呉医療、広島西医療
(山口) 山口宇部医療、岩国医療



関東信越

(茨城) 水戸医療 (栃木) -
(群馬) 高崎総合医療、渋川医療
(埼玉) 埼玉、東埼玉
(千葉) 千葉医療
(東京) 東京医療、災害医療、東京、村山医療
(神奈川) 横浜医療、箱根、相模原
(山梨) -
(新潟) 西新潟中央
(長野) まつもと医療、信州上田医療

東海北陸

(石川) 金沢医療、医王
(富山) -
(岐阜) 長良医療
(静岡) 静岡てんかん、天竜、静岡医療
(愛知) 名古屋医療、東名古屋
(三重) 三重、三重中央医療

四国

(徳島) - (香川) 高松医療
(高知) 高知 (愛媛) 四国がん

九州沖縄

(福岡) 小倉医療、九州がん、九州医療、福岡東医療
(佐賀) 肥前精神医療、嬉野医療
(長崎) 長崎医療 (熊本) 熊本医療
(大分) 別府医療 (宮崎) 都城医療
(鹿児島) 鹿児島医療、指宿医療
(沖縄) -

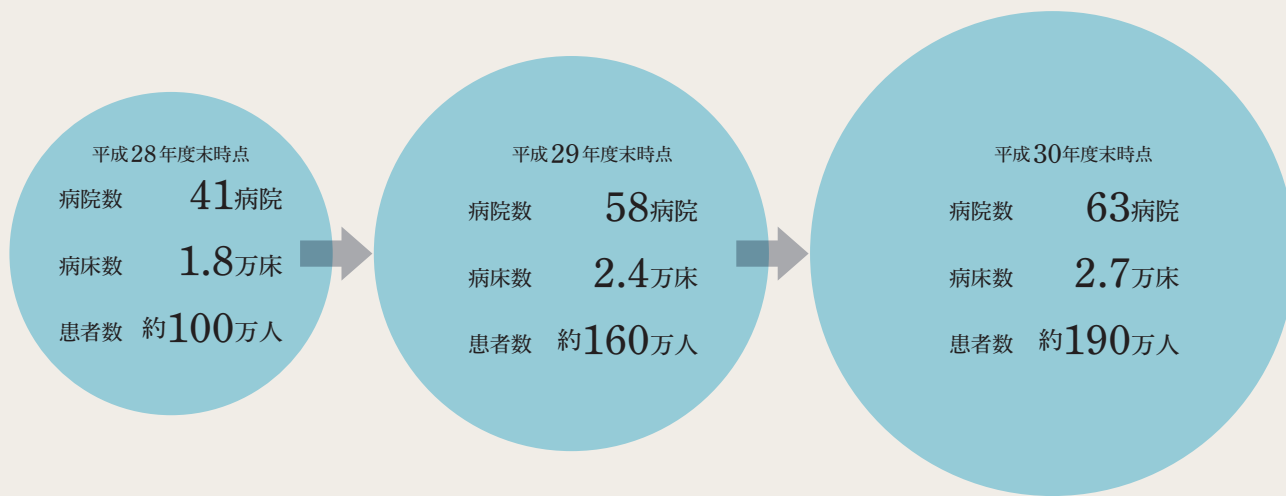
ベンダ・病院種別分布

	500床以上	350～499床	349床以下	複合(その他)	障害病床中心	精神科病床中心	総計
富士通	9病院 金沢医療、名古屋医療、 京都医療、大阪医療、 呉医療、岩国医療、 九州医療、長崎医療、 熊本医療	7病院 千葉医療、横浜医療、 相模原、三重中央医療、 姫路医療、小倉医療、 別府医療	2病院 南和歌山医療、 指宿医療*	6病院 渋川医療、東京、 村山医療、長良医療、 山口宇部医療、 福岡東医療	9病院 宮城、東埼玉、箱根、 医王、東名古屋、三重、 南京都、松江医療、 広島西医療		33
日本電気		3病院 北海道がん、埼玉、 災害医療	1病院 信州上田医療	6病院 北海道医療、旭川医療、 帯広、まつもと医療、 静岡医療、高知	1病院 仙台西多賀		11
ソフトウェア・サービス	1病院 岡山医療	6病院 水戸医療、高崎総合医療、 四国がん、九州がん、 嬉野医療、鹿児島医療	1病院 米子医療		1病院 高松医療*		9
亀田医療情報				1病院 敦賀医療	1病院 西新潟中央		2
SBS					2病院 静岡てんかん、天竜		2
日本IBM	2病院 仙台医療、東京医療						2
CSI			3病院 函館、弘前、都城医療				3
ナイス						1病院 肥前精神医療*	1
総計	12	16	7	13	14	1	63

* 災害診療記録事業には参加していない病院

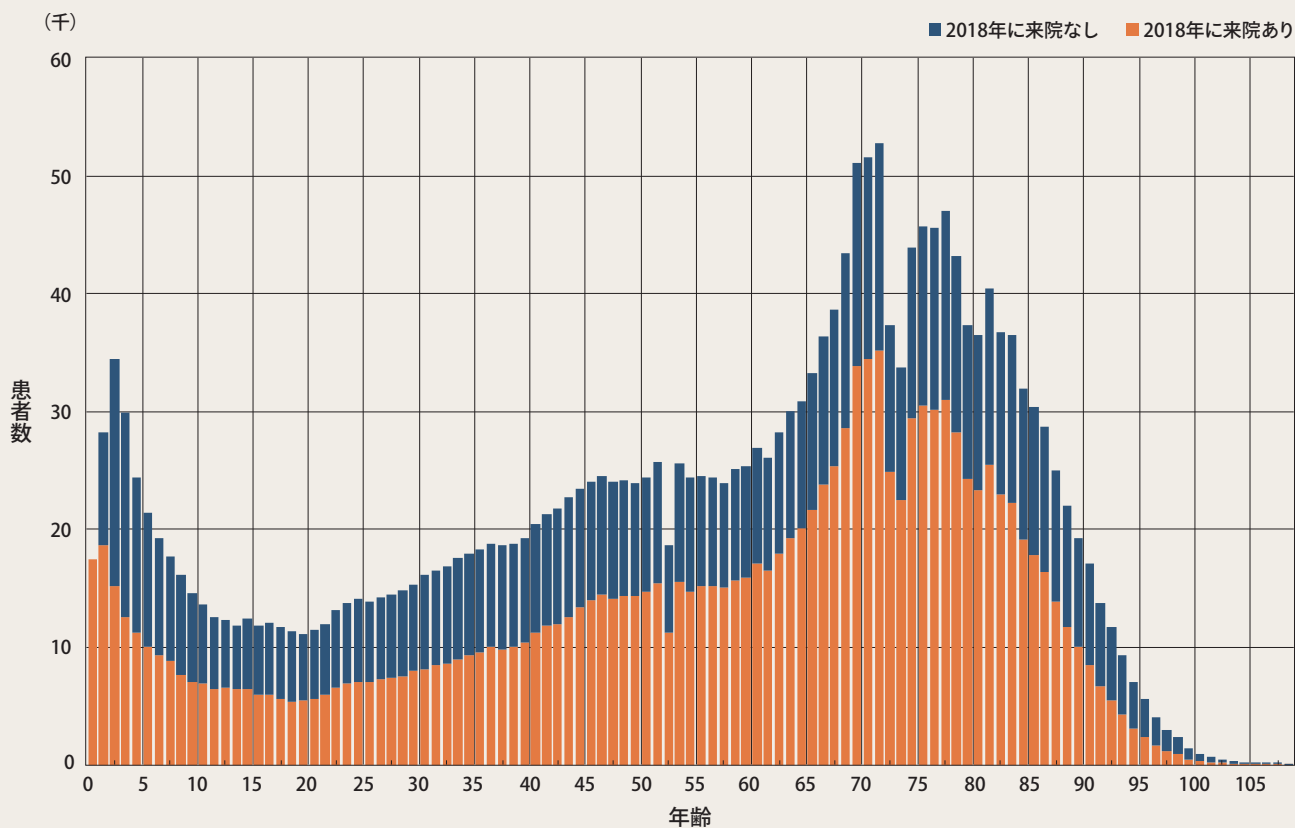


病院数・病床数・患者数



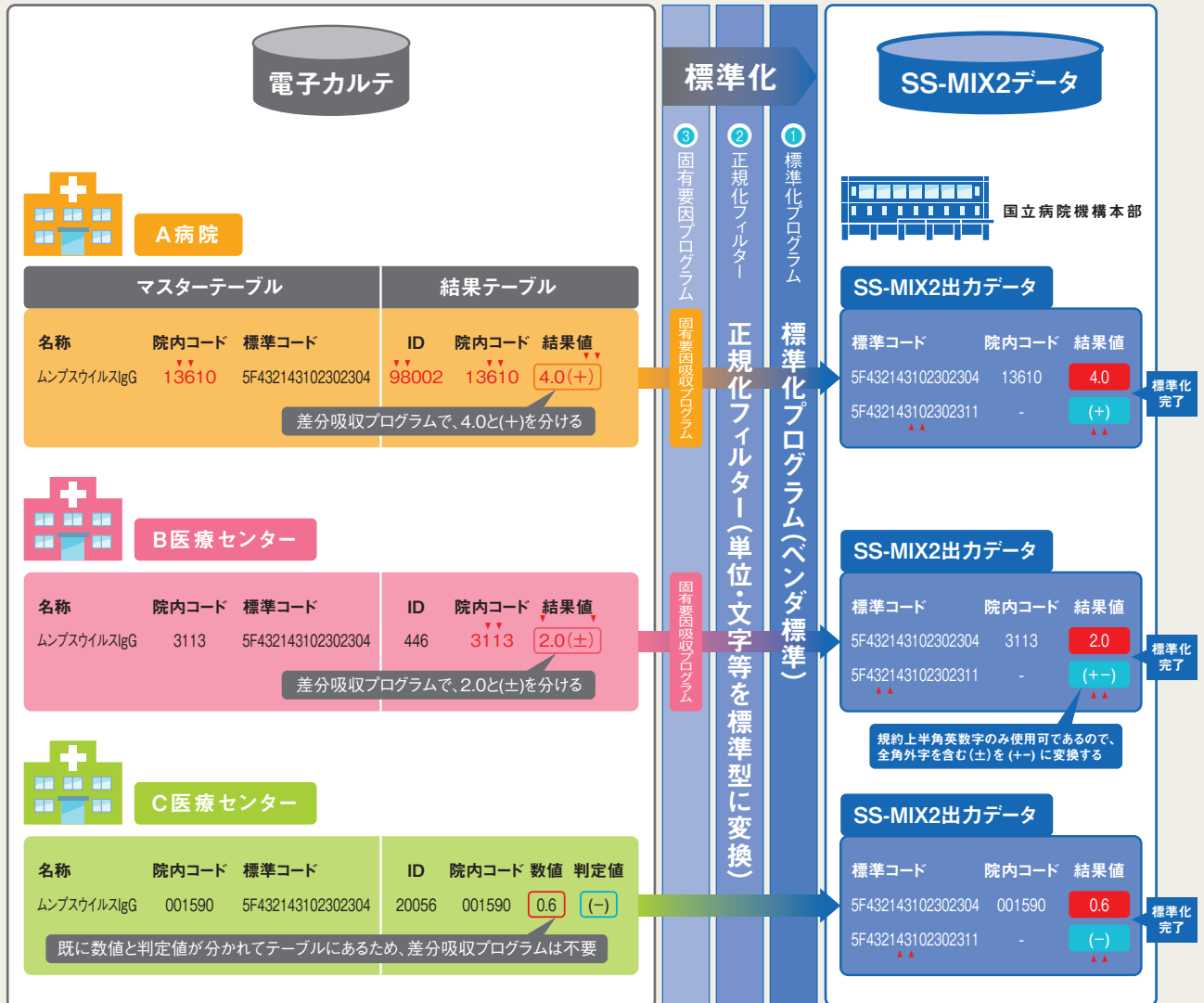
NCDAに集積している患者数（年齢）の状況

NCDAデータ保有患者数（平成30年12月現在）



4. SS-MIX2変換プログラムの構成

SS-MIX2変換プログラムの構成



結果値の表記が、病院独自[例・数値+判定値が一体化]となっているため、差分吸収プログラムで別々に表示する。
※判定値とは、基準値をもとに規定(例えば、ムンプスウイルスIgGでは、2.0未満は(-)、2.0~3.9は(±)、4.0以上は(+))

- ①②の標準化プログラムは他の医療機関でも使用可能な汎用的なもの。
 - ③の固有要因プログラムについては病院固有のもの。
- 各ベンダーが構築する標準化プログラムに固有要因プログラム③の機能が含まれていると、病院独自の仕様となり汎用化できないため、普及促進を図る手順書としての品質は不可。
- 今回の事業では、複数病院で標準化プログラムを運用して、それが汎用的なものであること(病院固有の変換機能が入っていないこと)を確認する。
- ※①②③のプログラムの著作権はベンダーにあるため国立病院機構はコード等中身を見ることは不可。
国立病院機構は出力結果により汎用性を確認する。

5. 保有するデータ種別



データ種別	データ内容	使用している標準コード
ADT-00 ADT-01 ADT-12 ADT-21 ADT-22 ADT-31 ADT-32 ADT-41 ADT-42 ADT-51 ADT-52 ADT-61	患者基本情報 担当医 外来診察受付 入院 外出泊 転科・転棟(転室・転床) 退院 アレルギー情報	
PPR-01	病名(歴)情報	ICD-10
OMP-01 OMP-02 OMP-11 OMP-12	処方・注射	HOT9
OML-01 OML-11	検体検査	JLAC-10
OMG-01 OMG-02 OMG-03 OMG-11 OMG-12 OMG-13	放射線・内視鏡・生理検査	
L-OBSERVATIONS^OBSERVATIONS^99ZL01	バイタル検査結果	
^(ローカル名称) ^^11506-3^経過記録^LN	診療録(外来/入院含む)	
^(ローカル名称) ^^34108-1^外来診療録^LN	診療録(外来)(入院・外来が別の場合)	
^(ローカル名称) ^^34112-3^入院診療録^LN	診療録(入院)(入院・外来が別の場合)	
^(ローカル名称) ^^18842-5^退院時サマリー^LN	退院時サマリー	
^(ローカル名称) ^^57133-1^紹介状^LN	診療情報提供書	
L-JSPEED^災害時JSPEED記録^99ZL01^74465-6^ 災害時JSPEED記録^LN	災害時JSPEED記録	

6. 構築・運用費用について

各病院における構築費用については、ベンダ側からの当初提示額から1～2割の低減を行い、1病院あたり平均約700万円となりました。また、本事業を安定的に継続していくため、運用コストの低廉化について検討した結果、1病院あたり年間約20万円での運用が可能となりました。

各病院での電子カルテ更新の際に発生するコストについては、今後も交渉を行い、さらなる低廉化が図れるよう努力していきます。

7. NCDAの利活用について

平成28年11月に「国立病院機構診療情報データベース利活用規程」を制定し、研究における利活用については本規程を遵守するとともに、倫理規程等の研究に関連する法令やルールを遵守することを定めました。利活用の際は、匿名化を原則としています。

また、国立病院機構本部内に「診療情報データベース利活用審査委員会」を設置し、データ利活用及び成果公表の適切性について審議を行っています。

利活用申請件数（データソースがNCDAのもの）

	平成28年度	平成29年度	平成30年度
業務	1	0	1
研究	2	0	4

8. NCDAにおける個人情報の取扱い


NCDAにおいては、以下の方針で個人情報を取り扱っています。

●患者同意

各病院において掲示している個人情報の利活用目的の範囲内で実施することを原則とします。そのために、ポスターを院内に掲示するとともに、患者からの利用不可の申出に対応できる体制を整備しています。

●法令対応

個人情報保護法、独立行政法人における個人情報保護に関する規程、ガイドラインおよび研究の倫理指針等に適切に対応します。医療分野の研究開発に資するための匿名加工医療情報に関する法律にも適切に対応していきます。



「国立病院機構 診療情報集積基盤 (NCDA)」
運用開始のお知らせ

平成28年1月1日より国立病院機構は、「国立病院機構 診療情報集積基盤 (略称NCDA)」の運用を開始し、全国の患者様の診療情報分析を本格的に実施していきます。

※診療情報は、患者様の年齢、性別、病名、薬の処方、検査結果などです。
※分析結果は、医学と医療の発展やより良い患者サービスの提供に活かします。
※利用に当たっては、個人情報は匿名化するなど厳密に管理します。
※自身の病歴が研究に利用されることを希望しない方は、病院にて申し出ください。

独立行政法人 国立病院機構
問い合わせ先 独立行政法人 国立病院機構 情報システム部
電話：05-5712-3105 FAX：05-5712-5682 e-mail: 700.ncda@nho.go.jp

9. NCD Aを活用した研究例



【研究名】

「腹部悪性腫瘍手術患者における術前血糖管理と術後創部感染症の関連に関する研究」

（研究目的）

腹部悪性腫瘍患者において術前の血糖管理状態と術後の創部感染症の関連を検証し適切な術前血糖管理目標を特定する

【研究名】

「電子カルテ情報をセマンティクス（意味・内容）の標準化により分析可能なデータに変換するための研究」

（研究目的）

電子カルテに実装可能な医療用語の標準化を行うシステムの開発に向け、そのステップとして、臨床現場からのニーズが極めて大きい退院サマリの自動生成技術を実現する。

10. MIAについて

～電子レセプトデータとDPCデータの利活用～

国立病院機構では、NCDAとは別に、診療情報データベース（MIA:Medical Information Analysis databank）を運営しています。

MIAでは、国立病院機構の全病院から診療情報の利活用を目的としてDPCデータ及びレセプトデータを収集しており、これらの既存データを二次活用することで、疫学研究の推進に役立てています。

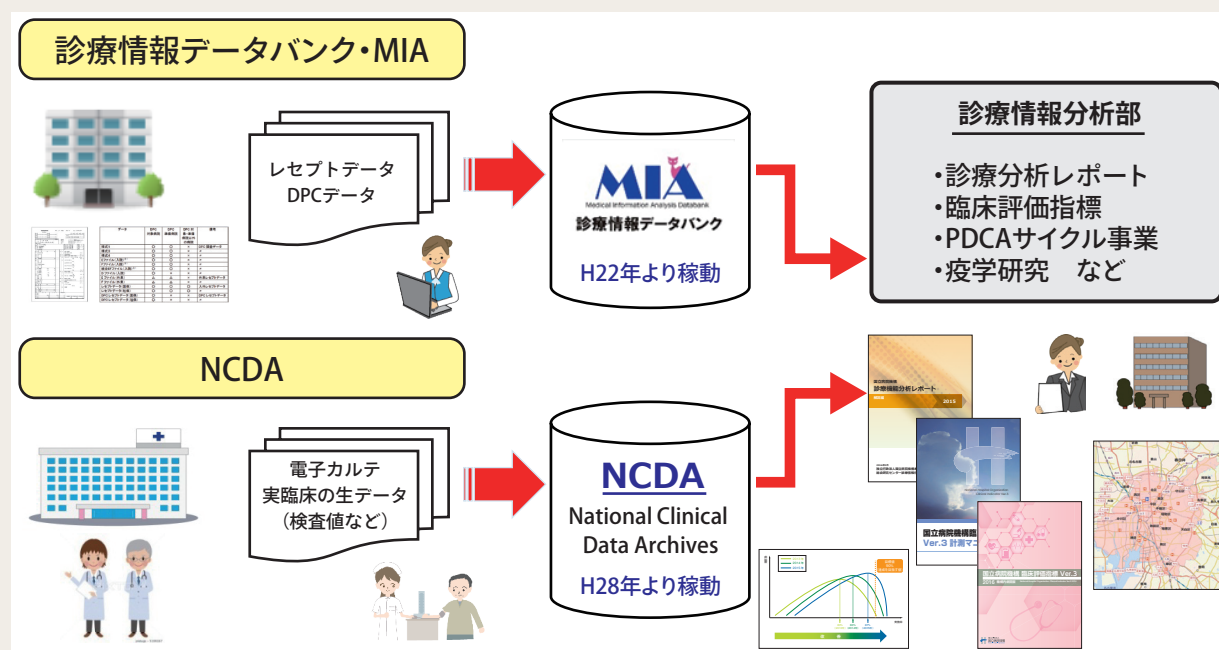
MIAの特徴

NCDAの特徴として挙げた、「③国立病院機構では単一の法人格によって複数の病院が運営されていることから、利活用におけるデータ提供の際には匿名化が行われているが、番号や生年月日・住所といった個人識別子を含んだ状態で運用。」について、MIAも同様の運用を行っています。そのため、NCDAとのデータ結合や病院における追加調査の実施等において技術的な制約がありません。

病院種別の変遷とMIAに蓄積されているデータ

	H22	H23	H24	H25	H26	H27	H28	H29
電子レセプト	144	144	144	143	143	143	143	143
DPC参加病院	45	49	53	52	54	54	64	64
DPC準備病院*	9	4	4	4	15	28	25	43
電子レセプトのみ	89	90	86	86	73	60	54	36

*データ提出加算病院を含む





MIA利活用の経緯

平成22年 4月	電子レセプト、DPCデータの収集開始
平成22年10月	臨床評価指標の開発 診療情報分析レポートの編纂
平成26年 4月	国立病院機構第3期中期計画開始（「電子カルテ情報の収集・分析について具体的な検討を進め、臨床研究等のIT基盤の充実を図る」ことを明記） 電子カルテ情報の収集・分析について具体的な検討を開始
平成28年 4月	医療の質改善事業（PDCA事業）開始に伴い、臨床評価指標の集計サイクルを四半期ごととする
平成28年11月	国立病院機構診療情報データベース利活用規程制定
平成29年 4月	NCDAとのデータ相互利活用の開始
平成30年 3月	データの品質を管理するためのバリデーション研究開始

MIAを活用した研究例

【競争的資金の獲得による研究例】

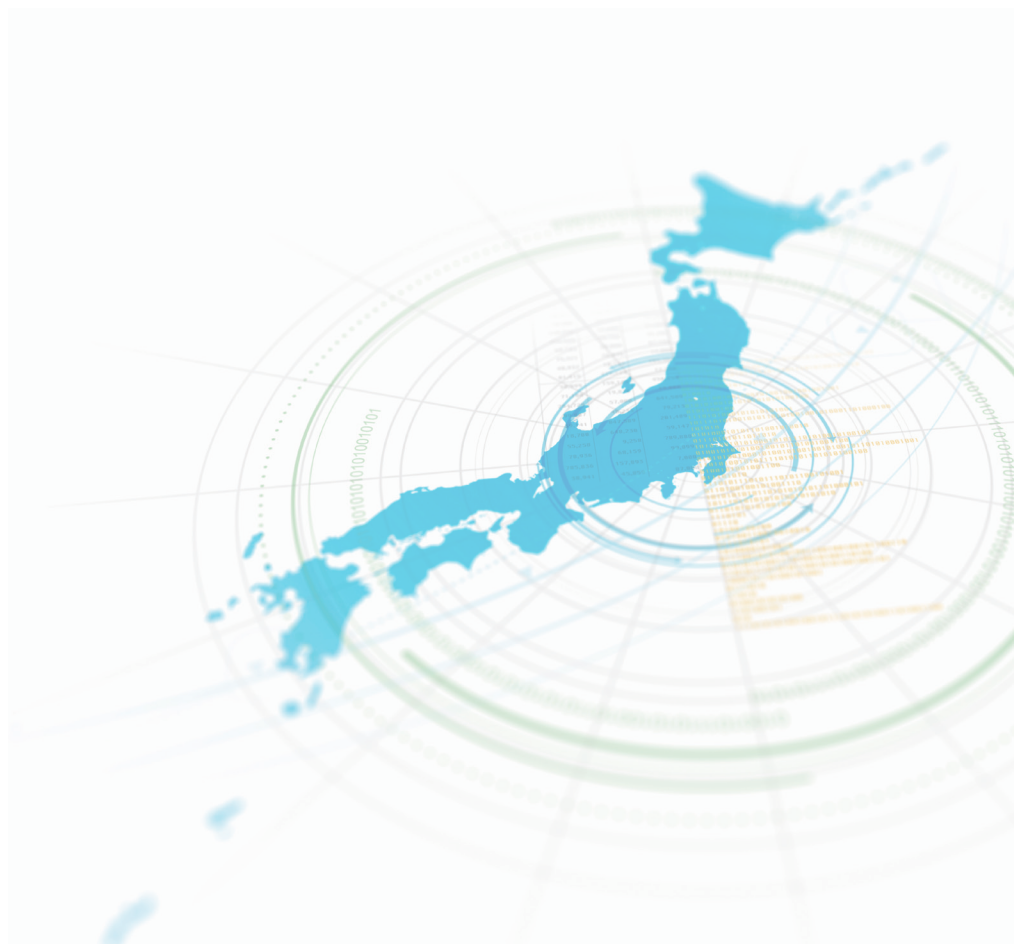
- ・ アレルギー疾患対策に必要とされる疫学調査と疫学データベース作成に関する研究
- ・ 診療情報データベースにおける重症敗血症および播種性血管内凝固症候群の特性に関する研究
- ・ 医薬品使用パターンによる重症度リスク補正を用いたアウトカム評価手法の開発
- ・ B型肝炎ウイルス再活性化に関与するウイルス・宿主要因の解明に基づく予防対策法の確立を目指す研究
- ・ 肝硬変患者の予後を含めた実態を把握するための研究
- ・ 診療情報データベースに記録された情報についての妥当性検証研究
- ・ 新薬へのスイッチの実態が後発医薬品推進政策へ及ぼす影響を評価する研究
- ・ 医療情報データベースを用いた治療効果検証手法の開発：カルテ調査との比較を通して
- ・ 手術患者の術後日常生活活動(ADL)悪化に対する鎮痛薬・鎮静薬・向精神病薬の影響

11. 本事業の今後

これまでの取り組みにより、SS-MIX2規格を用いて電子カルテベンダ毎に異なるデータを標準形式に変換して集積するIT基盤（NCDA）を構築し、その導入手順等の工程を標準作業手順書として作成、公開を行いました。

引き続きNCDAを運用するとともに、今後は、対応ベンダや実施病院の拡大、更には集積されたデータから新たな臨床評価指標の作成・揭示モニタリング、臨床疫学研究の推進、診療機能分析レポートの作成、薬剤副作用調査、被験者データベースによる治験の推進などの利活用を積極的に進めます。

公開した手順書は、他の医療機関・病院グループにおいても、より簡便に変換作業ができるようにすることを旨とするものであり、問い合わせに対しても助言等対応を行い、我が国の医療情報の標準化の普及推進に継続的に取り組んでいきます。





国立病院機構本部
情報システム統括部

〒152-8621 東京都目黒区東が丘2-5-21
TEL 03-5712-5130

2. NCDA システム仕様書

SS-MIX2 を用いた診療情報データベース構築の為の SS-MIX2 モジュール技術仕様書

1. システム要件

国立病院機構の各病院にて「国立病院機構診療情報分析基盤(NCDA)」に参加する為に調達する SS-MIX 2 モジュールの機能は以下の通りである。但し、本体の電子カルテシステム等の仕様上、作成が不可能であるものについては作成を要しない。その場合、何が不可能かを導入標準作業手順書に記載すること。

1.1 SS-MIX2 Ver.1.2d 機能

SS-MIX2 Ver.1.2d に準拠することとして、以下の機能を有すること。

- 日本医療情報学会発行の「SS-MIX2 標準化ストレージ構成の説明と構築ガイドライン Ver.1.2d」、「SS-MIX2 拡張ストレージ構成の説明と構築ガイドライン Ver.1.2d」、「SS-MIX2 標準化ストレージ仕様書 Ver.1.2d」、「標準化ストレージ仕様書別紙：コード表 Ver.1.2d」、「SS-MIX2 拡張ストレージ構成の説明と構築ガイドライン Ver.1.2d 別紙：標準文書コード表」に記載している仕様に対応していること。(尚、当初 Ver.1.2c 準拠としていたが、標準ストレージ部分では Ver.1.2c からの変更点について影響がないため Ver.1.2d 準拠ということとした。)
- 標準化ストレージ、拡張ストレージ、トランザクションストレージ、インデックスデータベースの4つのファイルを生成すること。
- 標準化ストレージにはデータ種別として 36 種のデータを出力すること。

(表 1-1 標準化ストレージ格納データ)

No	データ種別	種別名称	HL7 メッセージ型
1	ADT-00	患者基本情報の更新	ADT^A08
2	ADT-00	患者基本情報の削除	ADT^A23

No	データ種別	種別名称	HL7 メッセージ型
3	ADT-01	担当医の変更	ADT^A54
4	ADT-01	担当医の取消	ADT^A55
5	ADT-12	外来診察の受付	ADT^A04
6	ADT-21	入院予定	ADT^A14
7	ADT-21	入院予定の取消	ADT^A27
8	ADT-22	入院実施	ADT^A01
9	ADT-22	入院実施の取消	ADT^A11
10	ADT-31	外出泊実施	ADT^A21
11	ADT-31	外出泊実施の取消	ADT^A52
12	ADT-32	外出泊帰院実施	ADT^A22
13	ADT-32	外出泊帰院実施の取消	ADT^A53
14	ADT-41	転科・転棟(転室・転床)予定	ADT^A15
15	ADT-41	転科・転棟(転室・転床)予定の取消	ADT^A26
16	ADT-42	転科・転棟(転室・転床)実施	ADT^A02
17	ADT-42	転科・転棟(転室・転床)実施の取消	ADT^A12

No	データ種別	種別名称	HL7 メッセージ型
18	ADT-51	退院予定	ADT^A16
19	ADT-51	退院予定の取消	ADT^A25
20	ADT-52	退院実施	ADT^A03
21	ADT-52	退院実施の取消	ADT^A13
22	ADT-61	アレルギー情報の登録／更新	ADT^A60
23	PPR-01	病名（歴）情報の登録／更新	PPR^ZD1
24	OMD	食事オーダー	OMD^O03
25	OMP-01	処方オーダー	RDE^O11
26	OMP-11	処方実施通知	RAS^O17
27	OMP-02	注射オーダー	RDE^O11
28	OMP-12	注射実施通知	RAS^O17
29	OML-01	検体検査オーダー	OML^O33
30	OML-11	検体検査結果通知	OUL^R22
31	OMG-01	放射線検査オーダー	OMG^O19
32	OMG-11	放射線検査の実施通知	OMI^Z23

No	データ種別	種別名称	HL7 メッセージ型
33	OMG-02	内視鏡検査オーダー	OMG^O19
34	OMG-12	内視鏡検査の実施通知	OMI^Z23
35	OMG-03	生理検査オーダー	OMG^O19
36	OMG-13	生理検査結果通知	ORU^R01

「SS-MIX2 標準化ストレージ構成の説明と構築ガイドライン Ver.1.2d p11」

1.2 拡張ストレージへの出力機能

現在の SS-MIX2 モジュールでオプションとして既に導入している拡張ストレージへの出力機能は、そのまま提供すること。また、1.3.0 で規定する出力を行うこと。

1.3 NHO 対応としての設定

1.3.0 拡張ストレージへの出力機能

各社の SS-MIX2 モジュールの拡張ストレージへの出力機能を利用し、以下の情報を出力すること。その際、日本医療情報学会発行の「SS-MIX2 拡張ストレージ構成の説明と構築ガイドライン Ver.1.2d」に記載している仕様に対応していること。また、トランザクションストレージ、インデックスデータベースも同時に生成すること。

No	データ種別	種別名称	HL7 メッセージ型
1	L-OBSERVATIONS^OBSERVATION S^99ZL01	バイタル検査結果	HL7 V2.5 ORU^R30
2	^ (ローカル名称) ^^11506-3^経過記	診療録 (外来/入院含)	HL7 CDA R2

No	データ種別	種別名称	HL7 メッセージ型
	録^LN	む)	
2.1	^ (ローカル名称) ^^34108-1^外来診療録^LN	診療録 (外来) (入院・外来が別の場合)	HL7 CDA R2
2.2	^ (ローカル名称) ^^34112-3^入院診療録^LN	診療録 (入院) (入院・外来が別の場合)	HL7 CDA R2
3	^ (ローカル名称) ^^18842-5^退院時サマリー^LN	退院時サマリー	HL7 CDA R2
4	^ (ローカル名称) ^^57133-1^紹介状^LN	診療情報提供書	HL7 CDA R2

1.3.1 バイタル検査結果通知の出力

(1) バイタル検査結果通知のデータを、別紙の形式で拡張ストレージに出力する。尚、「診療日」に出力する日付は **OBX-14** トランザクション日時 (測定した日) とする。

(2) ファイル作成の単位は、データの格納構造として日付の下にあるため、最大でも一日分が1ファイルにまとまっている形とする。一日の中で測定のたびに作成するのでも良い。一日1ファイルなら、特定キーは測定日を出力する。一日に複数回のデータを出力する場合は、特定キーに測定日の時間まで (YYYYMMDDHH) 出力すること。

1.3.2 バイタルデータの項目及び形式等

(1) バイタルデータとして取得する項目は、「拡張期血圧、収縮期血圧、脈拍数、呼吸数、体温」の5項目とする。

(2) **OBX-3** 検査項目に出力するコードは **JLAC10** コードとする。バイタルデータを参考に適切な **JLAC10** を選択すること。

(3) 上記以外の項目を **SS-MIX2** に出力することは問題ないが、今回の対応では扱わない。但し、今後の検討で仕様として扱うことになる場合は、**JLAC10** コードを基準とした標準コードを必須とすることを想定している。この今後想定される検査項目は別表として提供する。

1.3.3 標準コード変換機能

SS-MIX2 データの出力に際しては、コードのマッピング表などに従って、院内のローカルコードを厚労省が定める標準コードに変換する機能を有すること。またマッピング表については、容易にその内容を変更できるマスターメンテナンスプログラム等の機能を有すること。

JLAC10 コード、**JANIS** コード、**HOT** コードについては、機構病院が **NCDA** 事業に参加する場合においては機構から提供する。

1.3.4 標準化ストレージにおける文字コードについて

メッセージの文字コードについては、「標準化ストレージガイドライン」で示されているとおり、1バイト系文字は **ISO IR-6 (ASCII)**、2バイト系文字は **ISO IR87 (JIS X 0208 第一水準、第二水準)** とする。ただし現実には上記以外の文字コードが電子カルテシステムに登録されている可能性があるため、以下のように対応することとする。

- 1 半角カナ文字 → 全角カナ文字に置き換えて **SS-MIX2** に出力する。
- 2 外字 → ■で置き換えて **SS-MIX2** に出力する。
- 3 環境依存文字については変換表を機構より提供するのでそれにより変換して **SS-MIX2** に出力する。

1.3.5 単位の文字表記の統一

SS-MIX2 データの出力に際して、臨床検査データの **OBX** セグメントの **6** フィールド目の単位の文字表記を統一すること。

【単位の文字表記の統一ルール例】 **ASCII** コードで表記すること

- ・かける → . (ドット)
- ・乗 → * (アスタリスク)
- ・ μ → u (小文字ユー)
- ・語尾に名称 → () で

- ・℃ → cel
- ・‰ → permil
- ・個 → pcs

【上記ルールの適用例】

- ・ mL → mL (ASCII コード)
- ・ $X10^2/\mu\text{l}$ → $.10^2/\text{uL}$ (かける、乗、 μ)
- ・ /HPF → /(hpf) (語尾に名称)

1.3.6 単位変換機能

SS-MIX2 データの出力に際して臨床検査データの単位に関しては、JLAC10 コードごとに、機構が定める単位に変換を行った上で SS-MIX2 データを生成すること。尚、JLAC10 コード別の単位表は別途機構から提供する。単位表は「SS-MIX2 標準化ストレージ仕様書 Ver.1.2」にも別表として添付する。

【単位変換例】

JLAC10 コード	数値	単位	→	JLAC10 コード	数値	単位
1A02500000127201	10.5	mg/l	→	1A02500000127201	1.05	mg/dL

1.3.7 計測値等の表記方法について

(1) 定性値・検出限界以下・検出限界以上の表記

- OBX (検体検査結果) セグメントの5フィールド目 (検査値) に検査結果を記述する場合、現在そのデータ形式は OBX-2 フィールドの説明にあるように NM 型、ST 型、CWE 型のうちいずれかの形式で記述することとなっている。
- 今回の仕様では、定性値・検出限界以下・検出限界以上のデータについては、SN 型の表現方法を用いて SN 型の”^”を” ” (スペース) に置き換える。
- この件の説明は、「SS-MIX2 標準化ストレージ仕様書 Ver.1.2」 P104 表 3-77 検査結果セグメント (OBX) 定義 の OBX-2 の項目説明にも記述する。

(2) 複数の要素が一つの値で表現されている場合の表記

複数の要素が組み合わせられ一つの結果値として表記されている場合は、それぞれの要素に分離して表記すること。例えば定量値とクラス値が組み合わせられた結果値については、定量値とクラス値に分離する。

【定量値とクラス値の分離の例】

定量値とクラス値が組み合わせられた例

検査名称	院内コード	結果値	
ムンプス Virus IgG	001591	2.3(±)	
↓			
定量値とクラス値を分離した例			
SS-MIX2 標準コード	院内コード	結果値	備考
5F432143102302304	001591	2.3	
5F432143102302311	001591	+-	(半角スペース2つプラスマイナス)

1.3.8 トランザクションストレージのデータ保持期間

トランザクションストレージのデータ保持期間は、現在の標準化ストレージ及び拡張ストレージを作っているデータの再現に必要な分だけ保持しておくこと。

1.3.9 ST 型の長さ

- RXE-23(与薬速度)は ST 型で長さが 6 であるが、正負の記号と小数点を考慮し (例: +266.865)、本事業では 8 桁まで許容するものとする。

- **CX**型は先頭成分が**ST**型で長さが**15**であるが、**IN1-10**(被保険者グループ雇用者 ID)に長い名称の保険者が出力される場合などを考慮し、本事業では**CX**型の先頭成分は**30**桁まで許容するものとする。
- **XAD**型は第**8**成分(その他地理表示)が**ST**型で長さが**50**であるが、全角**50**文字(**100**バイト)と解釈しているシステムがあり半角文字で**100**文字登録出来るため、本事業では**XAD**型の第**8**成分は**100**桁まで許容するものとする。

1.3.10 トランザクションストレージのファイル切り替え機能

SS-MIX2の仕様上、トランザクションストレージはカレントの日付が変わった時点、もしくは記録中のトランザクションデータファイルのファイルサイズが一定量を超えた時点で、新たなファイルを作成して記録先を切り替えるものとなっているが、同一日付内において一定時刻（例えば**17:00**）を経過した時点で記録先を切り替える機能を追加する。

