

厚生労働行政推進調査事業費補助金（政策科学総合研究事業（政策科学推進研究事業））
「NDB のユーザビリティ向上を通じてクラウド上でのデータ二次利用を推進するための研究」
分担研究報告書（令和7年度）

HIC のユーザビリティテストと改善提案

研究代表者 明神 大也（浜松医科大学 健康社会医学講座）

研究分担者 村松 圭司（千葉大学医学部附属病院）

研究要旨

【目的】本研究は、厚生労働省の NDB（匿名医療保険等関連情報データベース）の利用に関するホームページに掲載されている「匿名医療保険等関連情報データベース（NDB）の利用に関するガイドライン」・「医療・介護データ等解析基盤（HIC）の利用に関するガイドライン」（以下「HIC/NDB ガイドライン」という。）、及び二次利用ポータルに掲載されている利用者マニュアルの記載に従って HIC のユーザビリティを確認し、HIC の利用上の課題を発見することを目的とする。

【方法と結果】通年パネルデータセット及び NDB-β を申請し、HIC での NDB データの利用を開始した。その過程で気づいた課題や改善提案を HIC 解析環境・二次利用ポータル・関連ドキュメントに分けて記載した。

【考察】今後は、複数の研究課題を立て、HIC の探索的利用環境において現行配置されているデータセットを扱い解析し、これらのデータセットの研究可能性を調査する。

A. 研究目的

【背景】

医療DXの推進に関する工程表（令和5年6月2日医療DX推進本部決定）、また規制改革実施計画（令和5年6月16日閣議決定）

「NDB や公的統計データの利活用の円滑化・迅速化」に従って、HIC を活用して、NDB データを迅速に提供する取組が進められてきた。しかし、プリセットデータがクラウド上にも関わらず CSV ファイルで提供される、解析環境内ではコマンドラインの操作が必要である等、これまでのオンプレミス環境における解析環境と比較しても、ユーザビリティに関して多くの課題が残っている。

こうした中で、規制改革推進に関する中間答申（令和6年12月25日）においては、公的 DB の仮名化情報について安全であるのみならず迅速かつ円滑に利用・解析を行うことができるクラウド環境について、HIC との関

係を整理しつつ実現することとされている。

【目的】

本研究は、「匿名医療保険等関連情報データベース（NDB）の利用に関するガイドライン」及び「医療・介護データ等解析基盤（HIC）の利用に関するガイドライン」（以下「HIC/NDB ガイドライン」という。）、及び二次利用ポータルに掲載されている利用者マニュアルの記載に従って HIC のユーザビリティを確認し、HIC の利用上の課題を発見することを目的とする。

B. 研究方法

通年パネルデータセット及び NDB-β を申請し、第 34 回匿名医療情報等の提供に関する専門委員会にて通年パネルデータセット及び NDB-β の HIC 上での利用承諾（いわゆる迅速提供）を得た。その後、Windows 環境で利用開始した。本研究ではその過程で気づいた課題や改

善提案を、HIC 解析環境と二次利用ポータル、関連ドキュメントの3つに分けてまとめた。

C. 研究結果

HIC/NDB ガイドライン及び利用者マニュアルの記載に従って、HIC にて通年パネルデータセット及びNDB-βを利用する際のユーザビリティの確認を行った。課題及び改善提案は、HIC 解析環境・二次利用ポータル・関連ドキュメントに分け、下記のとおり整理した。なお、課題及び改善提案は一覧化し、「別添_課題及び改善事項一覧」（以下、別紙）に整理した。

1. HIC 解析環境に関する課題及び改善提案

【課題 1-①】

HIC を利用した研究では、複数のユーザーが同一のサーバーで解析する必要がある一方で、仮想サーバーや Docker の利用が想定されておらず、ユーザー間の調整コストが高いほか、情報連携不足による誤操作のリスクが存在する。

また、Windows OS では Hyper-V が利用不可であり、利用者全員が同じ OS を利用する必要がある。

【改善提案 1-①】

仮想サーバーの構築や Docker の活用を可能とすることで、複数のユーザーが同一のサーバーで解析する際の調整コストやリスクが低減できるほか、利用者ごとに習熟した OS が利用可能となるなど、研究全体の生産性が向上することが想定される。

なお、上記の設計が難しい場合も、その旨をあらかじめ周知しておくことにより、利用者が当該前提のもと運用方針を検討できる。

【課題 1-②】

複数の利用者が同一のサーバーを利用する場合は、ある利用者の操作に起因したトラブルが利用者全体に波及し、データの取得や前処理、分析をやり直す必要性が生

じるなど、研究が停滞する可能性がある。

【改善提案 1-②】

ある利用者の操作に起因するトラブルが他の利用者に波及しうること、及びその防止策（案）を利用者マニュアルにより周知することで、トラブルの発生を未然に防げる可能性がある。

また、利用者による環境の巻き戻しを可能とすることにより、トラブル発生時の研究への影響を最小限にとどめられると想定される。

なお、利用者マニュアルの事前準備編（2026/2/27 に取得）には、最新のバックアップイメージでのリストア（再構築）作業は「通常運用中に発生した、解析基盤の動作不具合時のみ対応」することとされており、また、「利用者様の操作に起因する、トラブル発生時の環境の巻き戻しを目的としたリストア依頼などにつきましては、対応致しかねます」と記載されている。

【課題 1-③】

HIC を利用した端末においては、利用端末に機微な情報が保存されない一方で、特別抽出等の利用形態と同様に、申請した利用場所以外での利用が不可能である。

【改善提案 1-③】

クラウド環境での分析であることを活かし、例えば HIC 環境での利用に限り、窃視対策を条件に利用場所の制限を緩和することにより、利用者がより柔軟に研究を推進可能となることが想定される。

【課題 1-④】

解析環境上にコミュニケーションツールが存在せず、複数の組織での共同研究を推進する際にコミュニケーションコストが発生する。

【改善提案 1-④】

解析環境にコミュニケーションツールを導入することで、解析環境内で、解析環境上の情報を基に相談でき、コミュニケー

セッションコストが低減すると想定される。

2. 二次利用ポータルに関する課題及び改善事項

【課題 2-①】

HIC を利用して NDB の研究を行う場合、複数の研究者の関与が想定される一方で、二次利用ポータルの公表前確認の申請や、最新の申請書の確認は申請の担当者のみが実施可能であるため、担当者を除く利用者は担当者を通して依頼する必要があり、コミュニケーションコストが発生する。

【改善事項 2-①】

二次利用ポータルにおいて、担当者のアカウントに加え、利用者のアカウントを研究と紐づけて管理し、利用者による公表前確認や最新の申請書の確認を実施可能とすることで、コミュニケーションコストを削減できると想定される。

また、二次利用ポータルログイン時にログインパスワードを誤った場合、「ユーザ認証に失敗しました」の画面ではなく「セッションが既に終了しています」の画面となったため、cookie などを削除して等で何度かログイン試していると最終的にアカウントロックになった。パスワードを間違っても必ずしもセッション終了の画面が出るわけではなく再現ができない（規則性の判断がつかない）ため、正式な仕様漏れやバグとも言えないが本報告書には記載した。

3. 関連ドキュメントに関する課題及び改善事項

【課題 3-①】

利用者マニュアルでは、Amazon Redshift をはじめとする専門的な知識が必要とされる箇所の情報量が相対的に少ないため、特に AWS 等の習熟度が低い研究者において、分析に用いるデータの前処理に時間を要すことや、意図しない操作によりトラブル

が発生する可能性がある。

【改善事項 3-①】

利用者マニュアルに記載されている操作の前提条件の記載を充実化させることや、Research Question を基に、Amazon Redshift からデータを抽出し、解析結果を S3 に格納するまでの一連の流れをユースケースとして記載することにより、NDB-β を用いて分析する研究者の障壁を削減できる可能性がある。なお、前者の具体的な改善案は別紙に記載している。

【課題 3-②】

HIC/NDB ガイドライン及び利用者マニュアルに関して、ドキュメント間での記載の不統一や、URL のリンク切れ等が存在する。

【改善提案 3-②】

ドキュメントを更新する。なお、具体的な改善案は別紙に記載している。

D. 考察

今年度は HIC 解析環境、二次利用ポータル、関連ドキュメントの 3 つの観点で課題及び改善事項の検討を行った。HIC 解析環境や二次利用ポータルに関連する課題には、複数人が利用者となり、研究を推進するという特性から生まれる課題が多く見つかった。

複数人が利用者となることを前提に、HIC 解析環境や二次利用ポータル、ドキュメントの改善を行うことにより、利用者のユーザビリティが向上すると考えられる。また、ドキュメントに関しては、HIC 上での操作に想定以上の時間を要することや、意図しないトラブルが発生することを避けるため、HIC を操作する際に必要な情報を充実化することが重要であると考えられる。

今後は、複数の研究課題を立て、HIC の探索的利用環境において現行配置されているデータセット (NDB-β) を扱い解析し、NDB-β の研究可能性を調査する。

E. 結論

本研究では、HIC/NDB ガイドライン及び利用者マニュアルの記載に従って、HIC のユーザビリティを確認し、HIC 解析環境・二次利用ポータル・関連ドキュメントに改善余地のある課題を発見した。これらの課題の改善によりユーザビリティ上の向上が期待される。今後は、複数の研究課題を立て、HIC の探索的利用環境において現行配置されているデータセット（通年パネルデータセット及びNDB-β）を扱い解析し、これらの研究可能性を調査する。

F. 健康危険情報 なし

G. 研究発表

1. 論文発表
なし
2. 学会発表
なし

H. 知的財産権の出願・登録状況

1. 特許取得 なし
2. 実用新案登録 なし
3. その他 なし

(参考)

厚生労働省 HP 【NDB】匿名医療保険等関連情報データベースの利用に関するホームページ
https://www.mhlw.go.jp/stf/seisakunit-suite/bunya/kenkou_iryuu/iryuhoken/rese-puto/index.html, 2026/2/27 アクセス
二次利用ポータル
<https://www.hic.mhlw.go.jp/imart/top>, 2026/2/27 アクセス

No	カテゴリ	課題	改善に向けた対応方針	改善事項	本紙上のNo.
1	解析環境	医療・介護データ等解析基盤（Healthcare Intelligence Cloud、以下HICという。）を利用した研究では、複数のユーザーが同一のサーバーで解析する必要がある一方で、仮想サーバーやDockerの利用が想定されておらず、ユーザー間の調整コストが高いほか、情報連携不足による誤操作のリスクが存在する。また、Windows OSではHyper-Vが利用不可であり、利用者全員が同じOSを利用する必要がある。 なお、解析環境へのミドルウェアのインストールを試みる場合、2次利用ポータルへのアップロード機能（ファイル容量は500MBまで）を活用するか、保守事業者へ依頼する必要がある。	・システム改修 ・利用者への周知	仮想サーバーの構築やDockerの活用を可能とすることで、複数のユーザーが同一のサーバーで解析する際の調整コストやリスクが低減できるほか、研究者ごとに習熟したOSが利用可能となるなど、研究全体の生産性が向上することが想定される。 なお、上記の設計が難しい場合も、その旨をあらかじめ周知しておくことにより、利用者が当該前提のもと運用方針を検討できる。	1-①
2	解析環境	複数の利用者が同一のサーバーを利用する場合は、ある利用者の操作に起因したトラブルが利用者全体に波及し、データの取得や前処理、分析をやり直す必要性が生じるなど、研究が停滞する可能性がある。	・利用者への周知 ・制限緩和	ある利用者の操作に起因するトラブルが他の利用者に波及しうること、及びその防止策（案）を利用者マニュアル（二次利用ポータルに掲載されている。）により周知することで、トラブルの発生を未然に防げる可能性がある。 また、利用者による環境の巻き戻しを可能とすることにより、トラブル発生時の研究への影響を最小限にとどめられると想定される。 なお、利用者マニュアルの「事前準備編」（2026/2/27に取得）には、最新のバックアップイメージでのリストア（再構築）作業は「通常運用中に発生した、解析基盤の動作不具合時のみ対応」することとされており、また、「利用者様の操作に起因する、トラブル発生時の環境の巻き戻しを目的としたリストア依頼などにつきましては、対応致しかねます」と記載されている。	1-②
3	解析環境	HICを利用した端末においては、利用端末に機微な情報が保存されない一方で、特別抽出等の利用形態と同様に、申請した利用場所以外での利用が不可能である。	・制限緩和	クラウド環境での分析であることを活かし、例えばHIC環境での利用に限り、窃視対策を条件に利用場所の制限を緩和することにより、利用者がより柔軟に研究を推進可能となることが想定される。	1-③
4	解析環境	解析環境上にコミュニケーションツールが存在せず、複数の組織での共同研究を推進する際にコミュニケーションコストが発生する。	・ツールの導入	解析環境にコミュニケーションツールを導入することで、解析環境内で、解析環境上の情報を基に相談でき、コミュニケーションコストが低減すると想定される。	1-④
5	2次利用ポータル	HICを利用してNDBの研究を行う場合、複数の研究者の関与が想定される一方で、二次利用ポータルの公表前確認の申請や、最新の申請書の確認は申請の担当者のみが実施可能であるため、担当者を除く利用者は担当者を通して依頼する必要があり、コミュニケーションコストが発生する。	・システム改修	二次利用ポータルにおいて、担当者のアカウントに加え、利用者のアカウントを研究と紐づけて管理し、利用者による公表前確認や最新の申請書の確認を実施可能とすることで、コミュニケーションコストを削減できると想定される。	2-①

No	カテゴリー	課題	改善に向けた対応方針	改善事項	本紙上のNo.
6	ドキュメント	利用者マニュアルでは、Amazon Redshiftを始めとする、専門的な知識が必要とされる箇所の情報量が相対的に少ないため、特にAWS等の習熟度が低い研究者において、分析に用いるデータの前処理に時間を要すことや、意図しない操作によりトラブルが発生する可能性がある。	・ドキュメントの見直し	利用者マニュアルに記載されている操作の前提条件の記載を充実化させることや、Research Questionを基に、Redshiftからデータを抽出し、解析結果をS3に格納するまでの一連の流れをユースケースとして記載することにより、NDB-βを用いて分析する研究者の障壁を削減できる可能性がある。 なお、前者の具体的な改善案は下記の通り。 ・利用者マニュアルの「事前準備編」から「トライアルデータセット・通年パネルデータセット操作編」に記載されている内容を一覧として整理し、特定の操作を実施する際にどのマニュアルを参照すればよいか分かるようにする。 ・利用者マニュアルの「事前準備編」の「2.1.3.ログイン（解析サーバ（EC2））」を実施するに当たり、事前にEC2を起動する必要がある旨、及びEC2の起動方法は利用者マニュアルの「基本操作編」に記載されていることを記載する。 ・Redshiftでのスキャン実施等、高額な利用料が発生しうるケース及びその対策の記載を充実させる。 ・NDB利用者マニュアルの別紙「別紙_コマンド一覧.xlsx」の情報をHICのPowerShellにコピーする方法を記載する。 ・通年パネルデータセットのダウンロードの考え方が利用者マニュアルごとに異なるため、どのような時にどの考え方でダウンロードするのが望ましいのかを記載する（利用者マニュアルの「S3～EC2操作編」を基に、「別紙_コマンド一覧」のコマンドNo.10を実行すれば、指定した年度の通年パネルデータセットがダウンロードされる一方で、利用者マニュアルの「トライアルデータセット・通年パネルデータセット操作編」のp.10やp.11に記載されている操作を行うと、S3上に保管されている通年パネルデータセットのうち、一部のddl及びcsvがダウンロードされる）。	3-①
7	ドキュメント	「匿名医療保険等関連情報データベース（NDB）の利用に関するガイドライン」・「医療・介護データ等解析基盤（HIC）の利用に関するガイドライン」、及び利用者マニュアルに関して、ドキュメント間での記載の不統一や、URLのリンク切れ等が存在する。	・ドキュメントの見直し	ドキュメントを更新する。 ・「解析環境利用に係る通知書」のQRコードを手打ちせずとも表示できるようにする（表示するURLを押下すると404エラーが表示され、URLを全て手打ちする必要があるため）。 ・一般的なWindows製品において二要素認証を実現するための具体的な対策例をHIC/NDBガイドラインに記載する（HIC利用端末には二要素認証が求められるが、一般的なWindows製品に搭載されている認証方法（Windows Hello）だと顔認証した時点でログインできてしまい、二要素認証とならないため）。 ・利用者マニュアルに記載されているURLのリンク切れを解消する（「S3～EC2操作編」の1.3章に記載の「匿名医療保険等関連情報データベース（NDB）の利用に関するガイドライン」のリンクが切れているため）。 ・利用者マニュアルの別紙「別紙_コマンド一覧.xlsx」に記載されているコマンドを直接ペーストできるようにする（セキュリティ対策によりマクロ付Excelを扱わない方針としている可能性はあるものの、「別紙_コマンド一覧.xlsx」に記載されているコマンドは関数により作成されており、直接ペーストできないため）。 ・利用者マニュアルの「トライアルデータセット・通年パネルデータセット操作編」と「別紙_論理テーブル仕様書_通年パネルデータセット（NDB）（2022年度）」の「v1」版と「v1なし」版との違いを記載する。 ・ドキュメントの更新時は、関連するドキュメントを横並びで更新する（記載が統一されていないことにより、誤った理解を促す可能性があるため。例えば、2025/11/1に第3.1版に更新された「匿名医療保険等関連情報データベース（NDB）の利用に関するガイドライン」では、「なお、オンサイト環境におけるリモート用全量 NDB、若しくは HIC におけるリモート用全量 NDB 又は NDB-β を利用する場合、利用開始時点から遡って最大で 10 年分のデータが参照可能であるため、この範囲内の期間を抽出対象期間として記載すること。」と記載されている一方で、利用者マニュアルのRedshift操作編では、「Redshiftによるレポート情報取得時の性能劣化防止の観点から利用可能範囲を10年分と規定しております。ただし、ご利用期間中にRedshiftにてデータの再抽出を行うことを想定し、利用開始月から起算して10年(120か月)前までの診療月のデータは、常に参照可能となるよう設定しております。※なお、仕様上最大10年分のレポート情報が利用可能ですが、申請いただいた抽出期間を超えてデータ参照を行うことはできませんのでご注意ください。※やむを得ない理由により利用期間延長の変更申請を行われた場合は、この限りではありません。」と記載されている）。	3-②