

令和 5 年度厚生労働科学研究費補助金
地域医療基盤開発推進研究事業

「公的に標準化された医療情報を活用した感染症流行状況と一般診療状況を把握するための分析手法の開発
および評価方法に関する研究」
分担研究報告書

DPCデータを用いた感染症流行状況・診療状況の把握手法の開発

研究分担者	新城 大輔	東京医科歯科大学大学院 医療政策情報学分野 准教授
研究分担者	伏見 清秀	東京医科歯科大学大学院 医療政策情報学分野 教授
研究分担者	高嶋 隆太	東京理科大学創域理工学部経営システム工学科 教授
研究代表者	佐藤 大介	藤田医科大学大学院 病院経営学・管理学 教授

研究要旨：

○研究目的

新興感染症や既知の感染症・病態の流行時において、従来は速やかにその情報を把握する方法がなくアナログな方法に依存しており、Covid-19 等の感染時に大きな混乱をきたした。感染発症を契機に行う情報収集も非常に重要であるが、適切な初動対応や日々のモニタリングのために既存のデータベースを活用する方法が魅力的である。本研究では、本邦における公的に標準化された医療情報のひとつである DPC データベースにより当該手法論の開発可能性を検討することを目的とした。

○研究方法

2020 年度～2021 年度の DPC データを利用し、Covid-19 を対象として手法論開発の検証を行った。Covid-19 病名が付与された患者を抽出し、SQL Server にてデータ処理を行い、Tableau Desktop にて集計・可視化を実施した。

○研究結果

入院契機病名が Covid-19 である患者から、日本語病名で”疑い”と記載されているあるいは日本語病名から Covid-19 病名判断が困難であった症例を除外し、259407 症例が対象となった。対象患者の基礎情報の創出（年齢階級別症例数、都道府県別症例数）、時系列・年齢階級別患者数、時系列・呼吸器関連医療資源情報の創出等について実現可能性を確認した。一方、即時性の観点からは、入院期間、データ提出までの期間という2つのボトルネックがあることが改めて確認された。

○結論

即時性については改善の余地があるものの、DPCデータをもちいた感染症流行状況・診療状況の把握は可能であり、その礎となる方法論を開発した。

A. 研究目的

Covid-19 は我々の生活・生き方にも大きな影響を与えたが、今後同様の事例が生じることに備え

て適切な対策を講ずる必要がある。本邦では Covid-19 後に法改正等を実施し、新興感染症等の感染拡大時における医療提供体制の確保に関す

る事項が医療計画に位置付けられることになり、その一步を踏み出していると考えられる。

しかし、具体的な方法論について、従来のアナログ的な手法を用いるには限界がある。従来の方法は、発症患者数が少なければアナログ的な手法でも対応可能かもしれないが、患者数に比例して人的資源の不足や診療・保健所業務への影響が大きくなり、業務破綻が懸念される。いわゆるレジストリデータのようにデータを登録する形式と弱点として、医療現場・保健所における情報入力の手間が膨大であることが挙げられる。そのため、利用可能な既存データベースや IT を活用する方法論が効果的と考えられる。

本邦にある既存データベースのうち、急性期医療が必要となる重要な感染症の情報を得るのに最も有益と考えられるデータベースは、本邦の急性期医療のデータベースである DPC およびレセプト情報である。

本研究では、本邦における公的に標準化された医療情報のひとつである DPC データベースを用いて、新興感染症等の感染拡大時における医療提供体制の確保に資する情報を創出する手法論の開発可能性を検討することを目的とした。

B. 研究方法

データソース

厚生労働科学研究費・DPC研究班が保有するDPCデータベースから、2020年4月～2022年3月分のデータを用いた後向きコホート研究である。

組み入れ基準・除外基準

入院契機病名としてコロナ病名(U07.1 コロナウイルス感染症2019, ウイルスが同定されたもの)を付与された入院患者を対象とした。データ期間を考慮し、ウイルス同定されていないCovid-19(U07.2)については対象外とした。除外基準は以下の通りである。

- ・日本語病名で”疑い”の記載がある、もしくは、日本語病名からCovid-19病名の判断が困難で

あった症例

地理情報の付与

原則として、患者住所に基づく地理情報を用いた。DPCの様式1にある患者郵便番号情報に、「郵便番号・市区町村コード対応表」を紐づけて、医療施設調査結果として公表している「医療施設調査令和3年医療施設（動態）調査 二次医療圏・市区町村編 第1表」により市区町村コードと二次医療圏の紐づけを行った。今回の分析では利用していないが、所属医療機関の地理情報に基づく集計も技術的に可能である。また、患者住所ベース・二次医療圏または市区町村レベルの集計についても同様に実現可能であるが、いくつかの課題が散見されたため今回の分析には含めていない。

今回の分析では、郵便番号情報がマスタにヒットしない症例も対象として含めることとした。(地理情報が付与できない症例となる)。

マスタの作成

呼吸器関連資源情報、重症度が高い患者を受け入れるICU等の病床利用情報、Covid-19の治療等で利用される薬剤資源利用情報の抽出に必要なマスタを作成した。

統計解析

Microsoft SQL Server 2019にてデータ処理を行い、Tableau Desktop (2022.4)にて集計・可視化を実施した。

研究倫理に関わる事項

本研究はヘルシンキ宣言に則って行われた。DPCはデータベース収載時点で匿名化されているため、対象者からの同意取得は行っていない。東京医科歯科大学の倫理審査で承認されている(承認番号M2000-788)。

C. 研究結果

DPC データベースから、入院契機病名が

Covid-19 である 26898 症例が抽出され、除外条件に該当した 1491 名を除き、**259407 名**が解析対象となった。

患者概要を把握するため、年齢階級・性別別の集計結果(Figure1)および都道府県別入院患者数(Figure2)を算出した。Covid-19 患者が最も多い年齢カテゴリについて、男性では 45-64 歳、女性では 75 歳以上となっているなど、年齢階級別に患者割合が異なることが確認された。都道府県別入院患者数の集計について、地理情報が欠落している/マスタにヒットしない症例が 25 症例あったものの、ほとんどの症例では問題なく利用可能であった。

Figure1 年齢階級・性別別の入院患者数

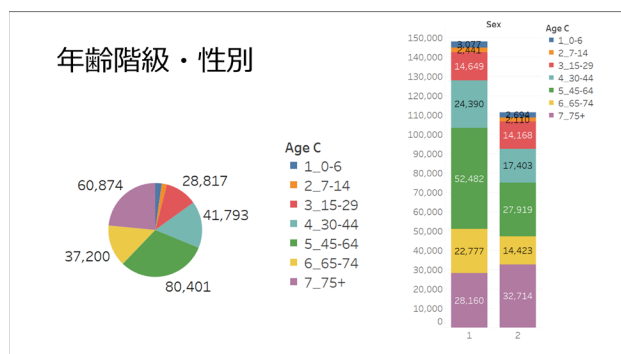


Figure2 都道府県別入院患者数

都道府県別入院患者数(1)



属性別・時系列患者数状況把握のために、毎月の年齢階級別患者数(Figure3)、ICU 利用日数別患者数(Figure4)、呼吸器関連資源利用別患者数(Figure5)を得た。横軸の時間軸については、入院年月を基準としている。時系列別患者数の傾向も既報と同様であり、全国調査と代表性に

おける一定以上の親和性が確認された。年齢階級別だけでなく、ICU 系の病床利用日数や呼吸器関連資源利用等の重症度に近い情報と紐づけて分析することで、入院年月ベースでどの程度いるのかを確認する仕組みが構築可能であることが確認された。

Figure3 時系列入院患者数

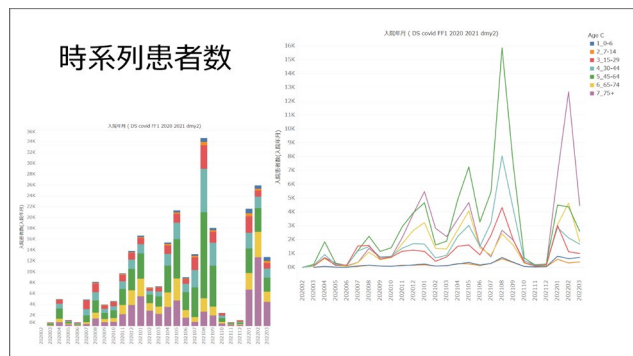


Figure4 時系列入院患者数・ICU 利用日数別

時系列・ICU利用日数別

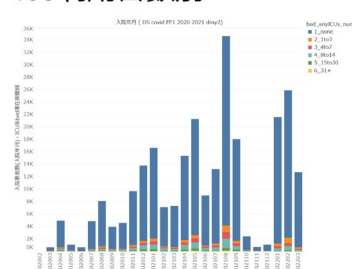
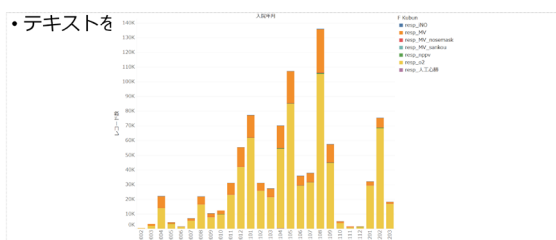


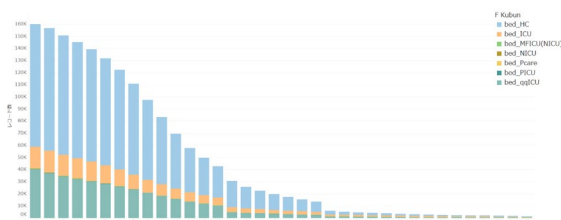
Figure5 時系列入院患者数・呼吸器関連資源利用別

時系列・呼吸器関連資源利用(F)



入院経過中の医療資源利用状況を把握するために、ICU 系 bed 資源利用状況(Figure6)、呼吸器関連資源利用状況(Figure7)、指定薬剤資源利用状

Figure6 入院後経過・ICUbed 系資源利用状況

[illegible]

一方、医薬品とは異なり医療機器については一部の製品を除いてどの製品(メーカー、品番)を利用しているかをレセプト・DPC から把握することが困難である。例えば、通常の人工呼吸を提供した場合、どの製品を利用しても原則同一点数でレセプト・DPC から

は製品判別が困難であるため、①より柔軟な設定変更が可能な製品により患者容態に合わせたコントロールを提供することとアウトカムの関連性が不明である、②機器不足時に暫定的に導入したケースと通常運用のケースの間のアウトカムの関連が不明である、等の制約が生じる。これらの制約を取り除くためには、レセプト上に医療機器詳細情報を記載する制度が必要と考える。本研究とは直接的な関係性は高くないが、当該制度を導入することにより、管理する医療機器製品が多さとME体制の評価や、保険医療材料制度のチャレンジ申請への活用なども可能ではないかと推察する。当該制度の実現可能性の検証等は別途必要であるものの、代表的な医療機器については、個別製品を識別可能な情報をレセプトに記載することも考慮しても良いであろう。

・即時性について

第二に、即時性に一定の制約がある。まず、DPCデータは、退院しないと作成されない。入院期間が長期間にわたる新興・再興感染症(疾病)の場合、DPCでの状況把握にタイムラグが生じることとなる。また、現行の診療報酬請求の仕組みでは最速でも毎月データを確定・提出することになるため、退院からデータ請求までの間隔が長いと把握がその分遅くなる。現在、内閣府SIPのプロジェクトにより「日々のレセプトデータによる医療動態解析基盤の開発※」が進められている。もし日々提出レセプトが実現するのであれば、日々提出レセプトを利用して把握する情報と、DPCから取得可能な情報(死亡の正確性、病名精度の高さ、一部疾患の重症度情報等)を加えて把握すべきものを分けて考え、制度設計すべきと考える。

※ 内閣府、第3期戦略的イノベーション創造プログラム(SIP)統合型ヘルスケアシステムの構築、テーマD1「医療機関・ベンダー・システムの垣根を超えた医療データ基盤構築による組織横断的な医療情報収集の実現」

・情報公表等のルールについて

第三に、公表可能な情報粒度についてルールを

明確にする必要がある。今回の集計でも、複数の軸や集計単位を細かくしていくと、患者N数が匿名医療保険等関連情報データベースでのマスキング基準となる10や5未満となることがある。今回報告した図を例にすると、横軸の時間軸を月別→週別の集計に置き換えや、地理的単位を都道府県→二次医療圏/市区町村等に置き換えることにより、より細かい情報が把握でき政策的な価値も高くなるが、その分単位当たりの患者数が少なくなるため、10や5未満となる情報が多くなる。現時点でどの法令に紐づけるかという点については把握していないが、適切かつ迅速な初動とモニタリングを可能とするために最適なルール設定が必要と考える。

・地理的情報/医療機関情報について

今回の分析では患者所在ベースでの集計を行ったが、医療機関所在地の情報でも同様に把握可能とする必要があると考えられる。しかし、医療機関情報の取得・整理において、医療機関番号は地方厚生局にて発番され、当該情報は公表情報であるものの原則PDFベースのみの公開となっているため、過去情報も含めるとデータ収集・クリーニング等の手間が膨大である。全国の所属医療機関情報を利活用しやすいルール整備が必要になると考えられる。

また、市区町村の統廃合と二次医療圏の変更がしばしば生じることから、現行のルールでは「都道府県-二次医療圏-市区町村-医療機関」の管理コストが膨大になる。公的医療DBの連結解析の環境も整備されつつあることから、その周辺領域として適切な管理する仕組みを構築することが肝要であると考えられる。

・分析対象集団について

細かい点となるが、転院している患者の場合、複数のDPCデータを連結して解析が必要になる。DPCの様式1を連結した解析について、個々の研究者の努力でなされているものの、一般的な方法論は開発されていない。また、DPC対象外ではあるがCovid-19の診療に携わっている医療機関で治療された症例の情報が欠落してしまう等の制約もある。

・モデルについて

本分析に用いた DPC データを基に、確率モデルや機械学習等の学習アルゴリズムを用いた予測モデルを構築し、感染症データ収集の頻度変化に伴う感染症流行状況の把握の精度に及ぼす影響について経営工学、オペレーションズ・リサーチ分野の工学的手法を用いて評価、分析を行うデータが整った。本分析に用いたデータは、COVID-19 が未知のウイルス性による呼吸器系の新興感染症であり、我が国の医療体制や治療指針の経緯および緊急事態宣言等の国の諸政策ならびに国民の行動を踏まえた回帰的データである。これらの前提に基づくデータを予測モデルのパラメータに投入する際、各パラメータの値を取る確率を、上記データから機械学習を用いて算出することで、仮定の値ではなく過去の状況や実績に基づいた説明可能性の高い値に基づく予測モデルを作成することが可能となる。

このような政策科学の手法を取り入れた、感染症流行状況を把握する方法の提案を進めていく。

最後に、本研究の限界を述べる。まず、本邦で標準化された医療情報には、残念ながら検体検査結果や感染症検査結果等の検査結果情報は含まれていない。そのため、Covid-19 の診断は各施設で付与された病名情報に基づいているが、その正確性を確認した validation study は実施されていない。次に、レセプト/DPC に掲載されない情報について把握することは困難であることから、保険適用外や本邦未承認の診療行為等を中心として underestimate の可能性が否定できない。

E. 結論

本研究で開発したDPCデータベースに基づく新興感染症等の感染拡大時における医療提供体制の確保に資する情報を創出として感染症流行状況・診療状況の把握手法論は、情報不足可能性や即時性等に一定の制約があるものの、更に検討を深める価値があることが確認された。実施可能な政策として実現させるためには、日々提供レセプト等の情報を精査しつつ、標準化された医療情報の効果的な活用が期待される。

F. 健康危険情報

特になし

G. 研究発表

該当なし

H. 知的財産権の出願・登録状況

1. 特許取得

特になし

2. 実用新案登録

特になし

3. その他

特になし