Apr 30, 2024

NDB_covid19_medical_resource_usage

DOI
dx.doi.org/10.17504/protocols.io.eq2lyj34qlx9/v1

Keita Fukuyama[1]

[1]Kyoto university

K Keita Fukuyama
Kyoto univiersity

DOI: https://dx.doi.org/10.17504/protocols.io.eq2lyj34qlx9/v1

Protocol Citation: Keita Fukuyama 2024. NDB_covid19_medical_resource_usage. **protocols.io https://dx.doi.org/10.17504/protocols.io.eq2lyj34qlx9/v1**

Protocol status: Working
We use this protocol and it's working

Created: July 05, 2023

Last Modified: April 30, 2024

Protocol Integer ID: 84495

Keywords: ndb-covid19-medical-resource-usage, threatening global health care, pandemic, medical resource shortage, medical care provision, future pandemic, global health care, coronavirus disease, fewer patient, number of patient, ill patient, medical resource, kind of medical resource, healthcare provider, resources for all patient, patient, healthcare, medical insurance claim records in japan, hospital admission, respiratory support, health, working population, disease

Abstract

Purpose: The coronavirus disease 2019 (COVID-19) pandemic exhibited several different waves threatening global health care. During this pandemic, medical resources were depleted. However, the kind of medical resources provided to each wave was not clarified. This study aimed to examine the characteristics of medical care provision at COVID-19 peaks in preparation for the next pandemic.

Methods: Using medical insurance claim records in Japan, we examined the presence or absence of COVID-19 infection and the use of medical resources for all patients monthly by age group.

Results: The wave around August 2021 with the Delta strain had the strongest impact on the working population in terms of hospital admission and respiratory support. For healthcare providers, this peak had the highest frequency of severely ill patients. In the subsequent wave, although the number of patients with COVID-19 remained high, they were predominantly older adults, with relatively fewer patients receiving intensive care.

Conclusions: In future pandemics, we should refer to the wave around August 2021 as a situation of medical resource shortage resulting from the COVID-19 pandemic.

Key words: COVID-19, NDB, medical resource

Troubleshooting

Approval to use NDB data.

1 **Approval to NDB usage**
There are two ways to conduct analysis using Japanese receipt data in NDB: either On-site Research Center, where the analysis is performed in an analysis environment, or by having the data set for the extracted analysis provided to the researcher.
In this study, On-site analysis is used.

Regarding the use of the system, an application for use must be submitted to the following office.
Please refer to the manual on the following page for application details.

https://www.mhlw.go.jp/stf/seisakunitsuite/bunya/kenkou_iryou/iryouhoken/reseputo/index.html
(Japanese)

2 **Ethical approval**
To use NDB data, users must have approval for the study from your institution's ethics committee.
Please check with your institution's Ethics Review Committee for more information.

3 **Access to onsite**
Currently, on-site analysis environments exist at only two locations: the University of Tokyo, and Kyoto University. Please consider access to the analysis environment when using an onsite location.

Analyiss Environment Construction

4 **OS**
On-site analysis environment allows selection of either Windows or Ubuntu as the OS.
Ubuntu is used in this study.

5 **PostgreSQL setting**
Onsite DB settings are left at default.
The data folder needs to be changed to a data drive and the shared_buffers and work_mem settings need to be changed.
The access speed to the data drive is set to 125 MB/s.
The DB access speed is not faster by any means, though, the low cache will eat up disk access by writing back.

The use of GUI support represented by pgAdmin is not allowed. Please be able to build, manage, and operate the DB only with CLI.

6 **R and Rstudio**

R and Rstudio settings must also be modified.
Renv should be changed according to the process, as it is not set up to load huge objects.

7 **Utitilies**
If you're doing edge-of-the-edge analysis while monitoring memory, storage capacity, and disk access, you need utilities to monitor these.
We used nmon in our study.

Materials and methods

8 From here onwards, the contents is the same as in the manuscript.

9 **Ethics statement**
The Kyoto University Graduate School of Medicine Ethics Committee approved this study (http://www.ec.med.kyoto-u.ac.jp/) as R3620. All study procedures involving human participants conformed to the principles of the 1964 Declaration of Helsinki and any subsequent amendments or equivalent ethical standards.
Acquisition of informed consent for participation was waived in accordance with Article 16, Paragraph 2 of the Act on Securing Medical Care for the Elderly in Japan.

10 **Dataset preparations**
In Japan, all citizens and long-term residents can avail a universal health insurance through the National Health Insurance System. After being anonymized, claims for medicines and services available in almost all national health insurances are collected and stored in the NDB. Researchers can directly analyze the data stored in the NDB in the Onsite Research Center, a specified analysis environment with the prescribed application and permission.
This retrospective observational study was conducted using NDB data recorded from January 2020 [when the novel coronavirus (SARS-CoV-2) was first detected in Japan] to December 2022. We extracted the medical records of all citizens during such period, imported them into the PostgreSQL database existing in the Onsite Research Center, and then analyzed them.

11 **Definition of patients**
We used ID1n, a hash value generated from the insurer's ID, beneficiary's ID, birth date, and sex, as a patient identifier among patient IDs in NDB. NDB has a rule that IDs such as insured person's ID cannot be directly used. In NDB, the ID1n changes even for a single patient, mainly because of a change of insurer; therefore, following up a single patient over a long term and investigating the utilization of medical resources can be difficult. Hence, we aggregated the data on a person-month basis.
We defined the code corresponding to acute-phase COVID-19 infection from the disease name master used for claiming claims (Supplemental Table 2). Patients assigned with a

COVID-19 disease code without a suspicion flag were defined as patients with COVID-19 and then compared with those without a COVID-19 disease name in each month. Patients were stratified by age as follows: young (<14 years), adult (15–64 years), and older (>64 years).
To verify the validity of the abovementioned definitions, we compared the number of patients defined in this study with the number of newly infected patients in the open data provided by HER-SYS, information acquisition, and management system for new COVID-19 cases provided by the Ministry of Health, Labor and Welfare (https://covid19.mhlw.go.jp/extensions/public/index.html).

12 **Medical resource usage**

We examined the number of patients with or without COVID-19 admitted to general wards, ICUs, and psychiatric wards. We extracted claim codes corresponding to the use of oxygen administration, artificial respiration, and ECMO as respiratory support and counted the number of patients who received these medical treatments. Furthermore, the use of sedatives, corticosteroids, and anticoagulants, which were feared to be depleted during the pandemic, was assessed. We then counted the total amount of drugs used in vials, ampoules, tablets, capsules, and more. Regarding the implementation status of tests for diagnosing COVID-19, we classified the codes according to whether the detection target was nucleic acid or antigen, and whether the test was outsourced or not. Subsequently, we counted the number of tests performed.
To validate the total medical expenses invested in patients with COVID-19, we stratified and validated the total insurance claims and public subsidies by age and COVID-19 status. Dental claims were excluded from this tabulation. Infectious disease experts determined the codes used for tabulation. We calculated the sum of claims for insurance claims and public subsidies, stratified by patients' age and presence of COVID-19, followed by verification.

13 **Relationship with other pandemic respiratory viral infections**

To ascertain the control status of other respiratory viral infections under COVID-19 control, we visualized the trends of the number of records of influenza virus and respiratory syncytial virus (RSV) infections, along with the trends of COVID-19 infections. Supplemental Table 2 lists these codes.

14 **Data Availability**

The datasets used during our study are not publicly available because of the rule that the NDB can only be used if decided by the Ministry of Health, Labor and Welfare in Japan. Thus, the data and analysis environment are available only if the NDB use is approved (https://www.mhlw.go.jp/stf/seisakunitsuite/bunya/kenkou_iryou/iryouhoken/reseputo/index.html).
Codes, queries, and output contents of processing are published on GitHub (https://github.com/fk506cni/ndb_covid_infection) within the permitted files. Additionally, the procedures up to the use of the analysis environment are summarized in the protocols.io platform.

15 **Analysis Environment**

The analysis was conducted in a prebuilt environment using R (version 4.1.2)/RStudio (version 2021.09.01 Build 372) on Ubuntu 20.05.5 LTS at a closed network. The packages and their versions used in each stage of extraction and aggregation are listed in the html text on GitHub. Furthermore, the local database was built using PostgreSQL (version 14.6), and the NDB dataset was extracted from Redshift (version 1.0.50468) on Amazon Web Service.