

A commentary on the value of hospital data for covid-19 pandemic surveillance and planning

Douglas G. Manuel^{1,2,3,*}, Carl van Walraven^{1,3,4}, and Alan J. Forster^{1,4,5}

Submission History	
Submitted:	13/08/2020
Accepted:	13/01/2020
Published:	03/03/2021

¹Senior Scientist, Ottawa Hospital Research Institute, Civic Campus, 1053 Carling Ave., Ottawa ON, K1Y 4E9, Canada

²Distinguished Professor of Family Medicine, School of Epidemiology and Public Health, University of Ottawa, 600 Peter Morand Crescent, Ottawa ON, K1G 5Z3, Canada

³Senior Core Scientist, ICES

⁴Professor of Medicine, School of Epidemiology and Public Health, University of Ottawa, 600 Peter Morand Crescent, Ottawa ON, K1G 5Z3, Canada

⁵Adjunct Scientist, ICES

Abstract

Hospital data for covid-19 surveillance, planning and modelling are challenging to find worldwide in public aggregation portals.

Detailed covid-19 hospital data provides insights into covid-19's health burden including identifying which sociodemographic groups are at greatest risk of covid-19 morbidity and mortality. *Timely* hospital data is the best source of information for actionable forecasts and projection models of hospital capacity, including critical resources such as intensive care unit beds and ventilators that take time to plan or procure.

A challenge to generate timely and detailed hospital data is the reliance on separation or discharge abstracts and census counts. What are needed are well-maintained lists of patients hospitalized with covid-19.

From the standpoint of public health and health services researchers and practitioners, we describe the role of hospital data for studying covid-19, why admission data are hard to find, and how improved data infrastructure can meet surveillance and planning needs in the near future. Modern hospital electronic health records can create covid-19 patient lists and these decision support tools are increasingly used for research. These tools can generate patient lists that are transmitted and combined with public health data systems.

Keywords

covid-19; SARS-CoV-2; public health; hospital records; decision support techniques; medical records systems, computerized; forecasting

The value of hospital data for covid-19 pandemic surveillance and planning

"She found that not even the numbers of soldiers entering the hospitals, or leaving them – alive or dead – was known. From the first she kept meticulous records. The data she collected was the evidence that saved lives."

Reference to Florence Nightingale, Crimean War, to Scutari in Turkey, 1854 [1]

Detailed and timely publicly available hospital data for covid-19 are challenging to find worldwide. Even basic, timely and accessible hospital counts are difficult to find and use. Of thirteen covid-19 data aggregate portals cited by the Research Data Alliance, only one had hospital census counts and none had hospital admissions [2–5]. From the standpoint of public health and health services researchers and practitioners, we describe the role of hospital data for covid-19 surveillance and planning; why timely, integrated data are hard to find; and how improved data infrastructure can meet public health needs.

*Corresponding Author:

Email Address: dmanuel@ohri.ca (Douglas G. Manuel)

The main challenge for reports of covid-19 hospitalization is the reliance on administrative data, namely hospital separation or discharge abstracts and census counts. There are promising signs that clinical electronic health records are being organized for research, and open data standards for covid-19 are maturing [5, 6]. Jurisdictions around the world are beginning to link clinical lists of patients hospitalized for covid-19 with public health patient case data. In our setting in Ottawa, Canada, we found linked data provide infrastructure that is valued by a wide range of health planners and researchers.

What data are currently used for covid-19 surveillance and planning and why hospital data are needed

Two types of data are chiefly used worldwide for covid-19 surveillance: reports of newly confirmed covid-19 cases; and deaths classified as being related to covid-19. Hospital data have additional benefits to these two data types.

Benefit 1: A population-based covid-19 ascertainment

Detailed information about hospitalized patients with covid-19 provides insight regarding which populations are most

affected by covid-19 illness. Detailed data include predisposing comorbidities and sociodemographic characteristics, as well as measures of morbidity (length of stay, ICU admission, death, discharge to rehabilitation). Detailed information allows improved prevention and control strategies, including interventions to target people at highest risk of covid-19 morbidity or death.

Covid-19 case data can be misleading due to selective testing and they are susceptible to ascertainment bias (i.e. people tested for the disease do not represent the entire population). Widespread testing has been challenging in the earlier stages of the pandemic with low coverage and selective testing. However, the further expansion of testing will yield neither a consistent nor representative sample of covid-19 patients across time, disease severity, and social demographic groups. Serology testing has a complementary role to inform how many people have had covid-19. Still, serology testing does not describe current outbreak dynamics, and it is expensive.

In a universal health care system, most covid-19 patients who require intensive treatment are admitted to a hospital. This means that hospital patient data reflect a population perspective of severe covid-19 infection and transmission. Overall covid-19 infection and transmission can also be estimated by 'back calculating' cases using hospital infection rates [7].

The diverging trend in new covid-19 cases and hospital census during spring 2020 in many countries is instructive. In our setting, there was rapid increase in hospitalization that coincided with a corresponding rapid increase in confirmed covid-19 cases (Figure 1). Hospital census then dropped more quickly than confirmed cases. In the setting of differing case and hospitalization trends, both short and long-term projection models for hospitalization are improved with transmissibility estimates (effective reproduction number, R_t) based on hospitalized patients.

Using hospital data for covid-19 surveillance and planning has limitations. The time between symptom onset and hospitalization for covid-19 is about seven days, which will result in reporting delays for newly acquired disease. Hospitalization occurs in only a fraction of confirmed cases, particularly for younger people. Therefore, settings with a small number of hospitalizations should assess whether there is sufficient statistical power to estimate overall population infection rates and outbreak dynamics. There can be a change in the virulence of covid-19 that affects the risk of severe disease, which in turn affects how hospital cases of covid-19 reflect total population counts [8]. However, the examination of covid-19 virulence is difficult to examine without hospital data, further demonstrating the value of hospital data to inform the covid-19 pandemic.

Benefit 2: Measuring and forecasting the health burden of covid-19

Timely information about hospitalized patients with covid-19 is used to generate short and long-term projections for health care capacity. Most short and long-term projection models are based on growth measures of confirmed covid-19 cases, but they can also be generated using hospital-based

growth and transmission measures. Short-term forecast for hospital capacity is best estimated using current hospital use and growth measures based on *recent* hospital use. More sophisticated epidemic models can provide longer-range projections, but with more uncertainty in their estimates.

There are two main uses of projections. The first use is the counterfactual assessment for reducing new hospitalizations using control measures such as physical distancing or contact tracing (Figure 2). The second use is extended forecasts for ICU capacity, personal protective equipment (PPE) and other critical resources that take time to plan or procure [9].

Challenges for detailed and timely covid-19 hospital data

Hospital separation or discharge data were originally developed decades ago for administrative purposes, but they are now widely used for surveillance and research, amongst other purposes. Separation summaries include detailed information about covid-19 patient care. However, separation summaries are not timely. Separation data is generated after discharge from hospital. With covid-19, length of stay varies and can be long (weeks or months) resulting in delayed admission counts. Additional reporting delays occur because separation summaries are often generated by trained records staff and reviewed by the attending clinical team.

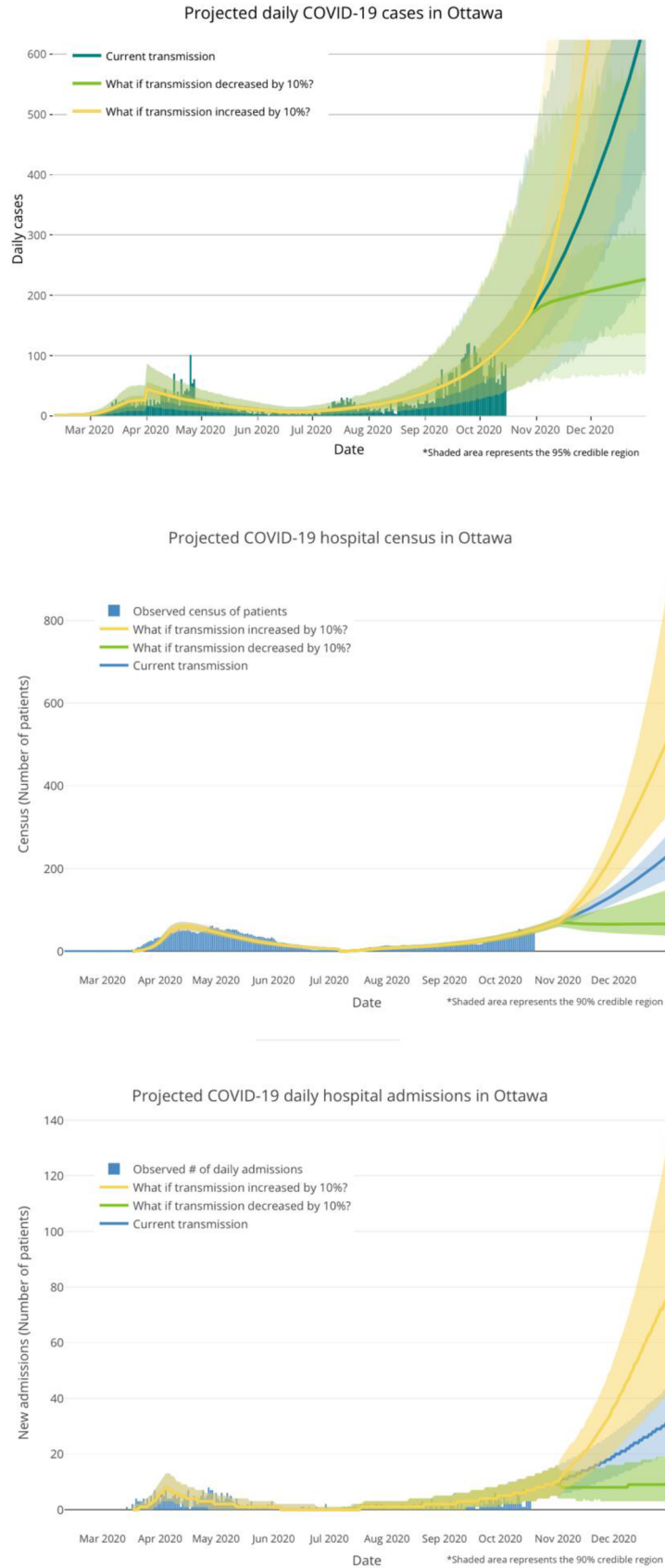
Hospital census is measured by simply counting how many patients have a confirmed covid-19 diagnosis (usually measured at midnight). Census data is timely but lacks detail. There are additional challenges such as counting new admissions. A new admission is straightforward to count when a patient arrives with a covid-19 diagnosis. But what if covid-19 is diagnosed after their admission day? What if a patient is transferred to another hospital? What if a covid-19 diagnosis was made outside the hospital setting, or a covid-19 patient is readmitted? The measure of new admissions for covid-19 is more difficult to generate compared to hospital census because people are counted if they have covid-19 either at the time of admission or they are retrospectively classified as a new admission when covid-19 is confirmed later during their hospital stay.

What is needed: A line listing of covid-19 patients

A line listing of covid-19 patients addresses at least three concerns including a lack of detail clinical data in aggregate hospital counts, a lack of timeliness of discharge abstracts, and the challenge of ascertaining covid-19 status throughout hospitalization.

A line listing includes individual key covid-19 patient information that is updated daily until discharge or illness resolution (Figure 2). Preferably, a patient's listing continues if a patient is transferred to another hospital or is readmitted. A patient's record is updated to confirmed covid-19 if their test comes back positive on any day of their admission. A line listing of covid-19 cases is long-standing and fundamental data approach used by public health epidemiologists for outbreak management. Epidemic curves are routinely reported on the date of illness onset. Covid-19 surveillance similarly benefits

Figure 1: Ottawa covid-19 historic and projected cases, hospital census and hospital admission



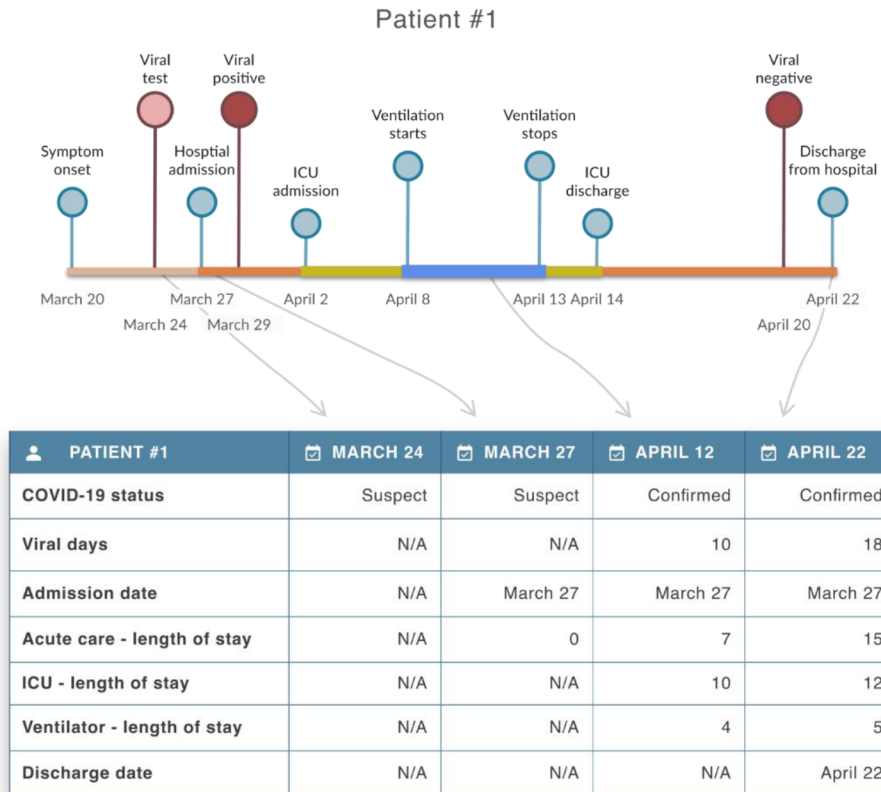
The trends in covid-19 incident cases and hospitalization admissions diverged in May 2020, with differing trends and projections since that time.

Figures from <https://613covid.ca> for the City of Ottawa (area code for Ottawa is 613).

The website was created in collaboration between Ottawa Public Health (the regional public health department) and the five local hospitals that accept covid-19 patients. The website started in April and updates projections daily with hospital and public health line list data.

Accessed: 2020-10-23.

Figure 2: Line Listing for a single patient



from timely and regular reporting hospitalization based on the date of illness or admission.

Patient cohorts are increasingly used in covid-19 research. The creation of research cohort and patient line lists are similar; both have detailed individual-level patient data to describe the course of illness and patient characteristics. Once generated, patient data can be linked to additional data to increase the information regarding patient characteristics and long-term outcomes. The difference between a public health line list and research cohort is the timeliness and update period. A line list is current, ongoing and updated throughout a patient’s hospital stay. A research cohort is usually historic and either not updated or updated periodically. Both patient line list and research cohorts can be linked to public health covid-19 case and outbreak management data.

The creation of a line list of hospital patients is straightforward in most electronic medical records systems (EMRs). In practical terms, most hospitals use specially trained analytic staff to create line listing of confirmed, suspect or probable covid-19 patients using definitions that are based on EMR covid-19 test values and infectious disease ‘flags’. Ideally, both public health and hospitals use the same covid-19 case definitions.

Many hospitals have overcome the challenges to generate covid-19 patient data to support research. However, there are additional challenges when using hospital patient data for covid-19 surveillance, monitoring and planning. Chief among the challenges is a lack of infrastructure to connect hospital information systems with public health information systems. Lacking linkage to hospital line lists, public health staff update

their public health information line list through a laborious process of phone calls or searches in hospital information systems to which they’ve been granted access. Low and middle-income counties have a wider range of challenges including a lack of EMRs and analytic staff.

Opportunity

There are signs that countries are responding by adapting existing information systems and developing new ones to connect covid-19 patient information from different health care sectors. Locally, in Ottawa, Canada, hospital decision support analysts and public health epidemiologists worked together to create a standard line listing that is transmitted to the regional public health department where it is merged with their information systems. The public health information system serves as the central point for covid-19 information and forms the basis of contact tracing, surveillance and modelling. The availability of robust and integrated hospital data improves surveillance and planning, including modelling future hospital capacity in the event of surges.

There is a tremendous effort to organize covid-19 hospital data for research. These initiatives can inform how hospitals and public health can work together to generate integrated information systems that are needed to control and plan for covid-19 in the coming year or more.

Acknowledgements

Website and analyses: Warsame Yusuf, Rostyslav Vyuha, Yulric Sequeira, David Schramm and others on Big Life

Lab team for doubling time figure, analyses and website at 613covid.ca.

The Ottawa Hospital: Deanna Rothwell and the Decision Support team for covid-19 line list data.

Ottawa Public Health: Amira Ali, Jacqueline Willmore, Dara Spatz for Ottawa Hospital data.

Funds from the Canadian Institute for Health Research grant FRN 162222 were redirected to support the analysis infrastructure for the doubling time analyses and projections.

Conflicts of interest

The authors declares that they have no conflicts of interest.

References

1. Magnello E. Great moments in statistics: Florence Nightingale. *Significance*. 2013;10(6):21–8.
2. Dong E, Du H, Gardner L. An interactive web-based dashboard to track COVID-19 in real time. *The Lancet infectious diseases*. 2020.
3. Eurosurveillance Editorial Team. Latest assessment on COVID-19 from the European Centre for Disease Prevention and Control (ECDC). *Eurosurveillance*. 2020;25(8).
4. The COVID Tracing Project [Internet]. The Atlantic. 2020. Available from: <https://covidtracking.com/contact>
5. RDA COVID-19 Working Group. Recommendations and Guidelines on data sharing. 2020.
6. Center for Leading Innovation and Collaboration. Using the ACT network to gain insight into COVID-19 Bethesda: The University of Rochester Center for Leading Innovation and Collaboration (CLIC); 2020 [Available from: <https://clic-ctsa.org/events/using-act-network-gain-insight-covid-19>
7. Bhatia S, Cori A, Parag KV, Ainslie KEC, Baguelin M, Bhatt S, et al. Short-term forecasts of COVID-19 deaths in multiple countries 2020 2020-04-20.
8. Gandhi M, Rutherford GW. Facial Masking for Covid-19 — Potential for “Variolation” as We Await a Vaccine. *New England Journal of Medicine*. 2020.
9. Shoukat A, Wells CR, Langley JM, Singer BH, Galvani AP, Moghadas SM. Projecting demand for critical care beds during COVID-19 outbreaks in Canada. *Canadian Medical Association Journal*. 2020:cmaj. 200457.

