Posted on Authorea 25 Jul 2020 — The copyright holder is the author/funder. All rights reserved. No reuse without permission. — https://doi.org/10.22541/au.159569198.82432694 — This a preprint and has not been peer reviewed. Data may be preliminary.

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

July 25, 2020

Abstract

Hospital data for covid-19 surveillance, planning and modelling are challenging to find well into the pandemic. Data for new covid-19 admissions are particularly challenging to find. Accurate hospital data are the best source of information for short-term forecasts of health care capacity, and they improve long-term covid-19 models that are typically based on cases. 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 over the next year. Critical infrastructure includes a well-maintained list of patients hospitalized with covid-19 that is integrated with public health information systems.

Authors and Affiliations

Douglas G. Manuel, MD, MSc[1,3,4]

Carl van Walraven MD, MSc[2,3,4]

Alan J. Forster MD MSc[2,3,4]

1) Department of Family Medicine, University of Ottawa, Ontario, Canada

2) Department of Medicine, University of Ottawa, Ontario, Canada

3) Ottawa Hospital Research Institute, Ottawa, Ontario, Canada

4) ICES, Ontario, Canada

Data and Code: https://github.com/Big-Life-Lab/covid-19-hospital

Corresponding Author: Douglas G. Manuel

The authors have no conflicts of interest. Acknowledgements are at the end of the manuscript.

The Value of Hospital Data for covid-19 Pandemic Surveillance and Planning

Publicly available hospital data for covid-19 are challenging to find in Canada and elsewhere well into the pandemic. Of thirteen covid-19 data aggregate portals cited by the Research Data Alliance only one had hospital census counts and none had hospital admissions.(Group, 2020; Dong et al., 2020; Eurosurveillance Editorial Team, 2020; Atlantic, 2020) China and Belgium report admission data, but data for other G7 or OECD countries are absent from public health websites.(Bodart, 2020; Organization, 2020) The lack of new covid-19 hospital admission data as part of surveillance reporting is an indicator or long-standing poor data integration between hospitals and public health.

Posted on Authorea 25 Jul 2020 — The copyright holder is the author/funder. All rights reserved. No reuse without permission. — https://doi.org/10.22541/au.159569198.82432694 — This a preprint and has not been peer reviewed. Data may be preliminary.

Covid-19 will continue to affect our society, health, and hospitals for at least the next year. Short-term forecasts of hospital use are vital to planning for covid-19 surges and waves. Long-term covid-19 projections are required to assess the effectiveness of different control strategies related to physical distancing, healthcare planning, contract trancing or economic policies. Hospital data are the best source of information for short-term forecasts of health care capacity, and they improve long-term covid-19 models that are typically based on cases.

We describe the role of hospital data, why admission data are hard to find, and how improved data infrastructure can meet surveillance and planning needs over the next year. There are promising signs that hospital-based data are being organized for research, and open data standards for covid-19 are maturings.(Group, 2020; The University of Rochester Center for Leading Innovation & Collaboration, 2020) These new covid-19 data initiatives support the use of individual hospital patient data, as opposed to the current use of aggregate hospital census counts. Routinely-generated lists of hospitalized covid-19 patients should be shared and combined with public health patient case data. The combined and integrated data provide robust support covid-19 surveillance, modelling and planning.

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

Hospital data records a useful *population* perspective of the pandemic. 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 neither yield a consistent nor representative sample of covid-19 patients across time, disease severity, and social demographic groups. That stated, a high capacity for rapid testing is warranted because the diagnosis of confirmed cases is invaluable for early outbreak detection, outbreak detection in specific populations (e.g. long-term care homes), control using contact follow-up, and management of patients using isolation precautions. Serology testing will have 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 with logistical challenges to create a population perspective.

In a universal health care system, most covid-19 patients who require intensive treatment will likely be admitted to a hospital. This means that routinely-collected hospital patient data reflect a population the perspective of severe covid-19 infection that can be used to describe covid-19 transmission in both populations at high risk of complications and the overall community.

The recent divergence in the trend of new covid-19 cases and hospital census in several countries is instructive. In Canada, several provinces and cities report hospital and intensive care census (**Fig. 1**). During March and early April, there was a rapid increase in hospitalization that coincided with a corresponding rapid increase in confirmed covid-19 cases. However, since then hospital census rates dropped more quickly than confirmed cases.

Posted on Authorea 25 Jul 2020 — The copyright holder is the author/funder. All rights reserved. No reuse without permission. — https://doi.org/10.22541/au.159569198.82432694 — This a preprint and has not been peer reviewed. Data may be preliminary.

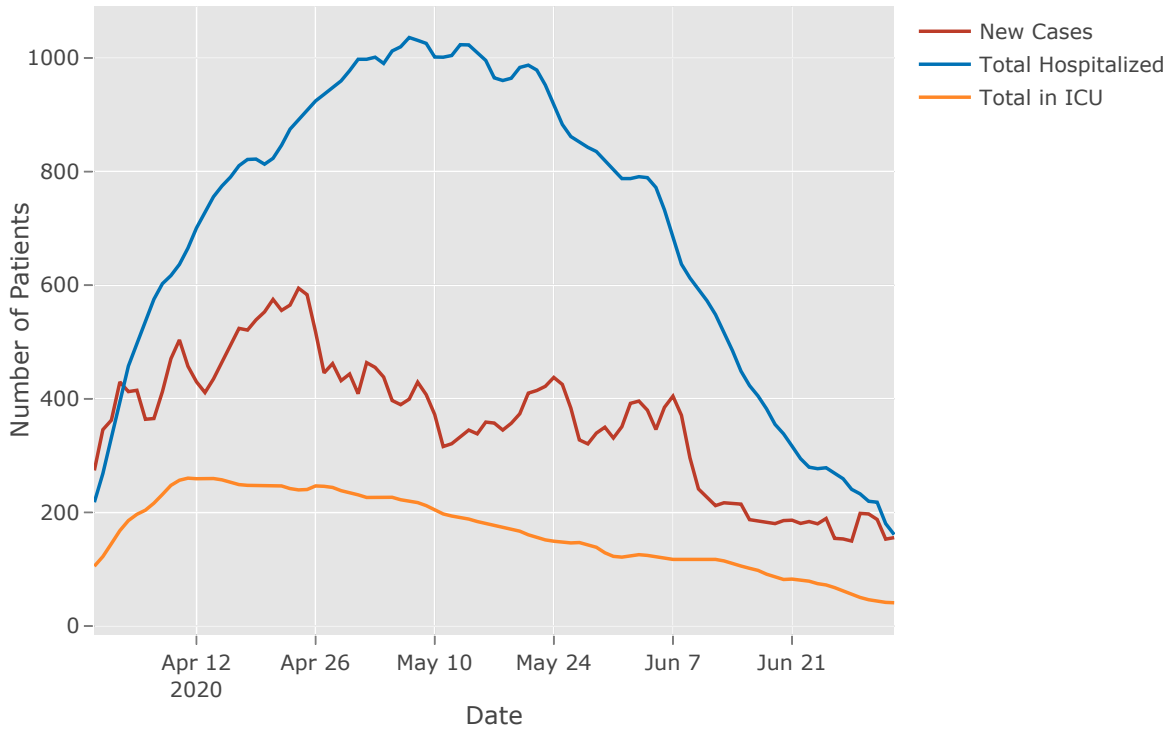


Figure 1: **Ontario covid-19 3-day moving average: new cases, total hospitalization patients, and total ICU cases.** The intensive care unit (ICU) census increased quickly in Ontario, Canada with a peak just a few weeks after the first covid-19 cases were confirmed. Following, ICU decreased steadily. Covid-19 cases continued to be diagnosed throughout the next two months before slowly decreasing.

There are two likely reasons why rates of hospitalization and confirmed cases have not mirrored each other in Canada and other countries. The first explanation is an increased case detection of less severe covid-19 infection as testing capacity increased. In this situation, population transmission of covid-19 overtime is more consistently measured using hospital data compared to reported cases. The second explanation is a reduction in covid-19 transmission in people at high risk of complications compared to the general population. Reduced transmission in high-risk groups would happen if at-risk people more effectively adhered to physical distancing compared to low-risk people. In this situation, short and long-term models are improved with the use of different should transmissibility estimates (effective reproduction number, R_t) for high-risk groups and the general population. Both these explanations – changing case detect or changing transmissibility – argue for covid-19 forecasts and modelling of control measures that consider transmissibility based on hospital data.

Using hospital data to estimate population covid-19 incidence or past infection 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, the use of hospital data to estimate overall population infection rates and outbreak dynamics is statistically robust only for larger communities.

Posted on Authorea 25 Jul 2020 — The copyright holder is the author/funder. All rights reserved. No reuse without permission. — https://doi.org/10.22541/au.159569198.82432694 — This a preprint and has not been peer reviewed. Data may be preliminary.

Benefit 2: Measuring the health burden of covid-19

The second benefit of hospital data is the accurate and meaningful measurement of health care burden. A key consideration for covid-19 planning is health care capacity for patients with severe disease. Hospital data are clearly well-suited for this purpose.

The 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. Most long-term projection models are based on growth measures based on confirmed covid-19 cases, but they can also be generated using hospital-based growth and transmission measures. The inclusion of hospital-based transmission measures improves long-range models for two uses. The first use is the counterfactual assessment for reducing new hospitalizations using control measures such as physical distancing or contact tracing (**Fig. 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.(Shoukat et al., 2020)

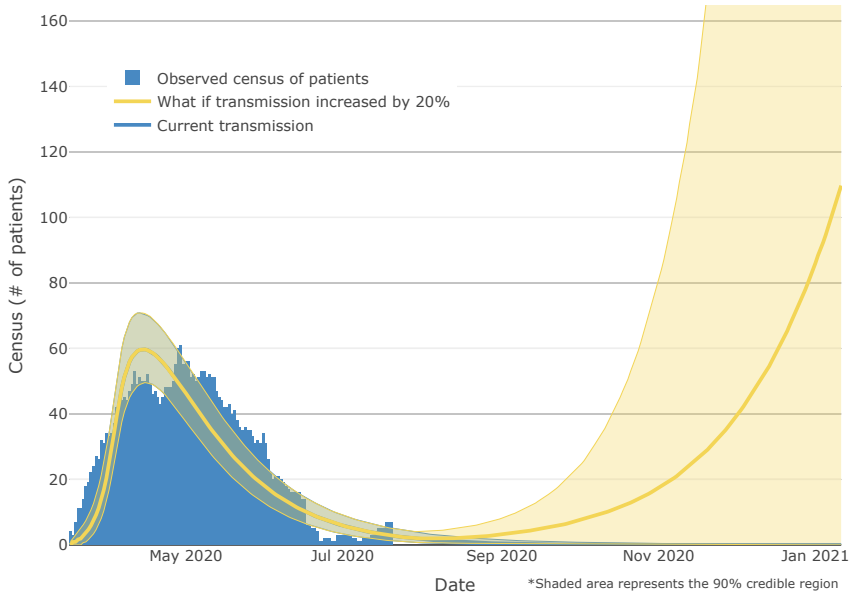


Figure 2: **Projected hospitalization for Ottawa.** Projection based on hospital census. Accessed from https:6130ttawa.ca on 2020-07-24

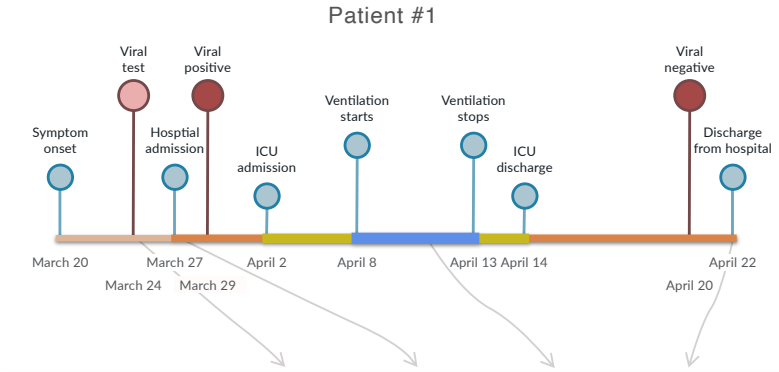
Issue: census and incident counts are not the same

There is a subtle but important distinction between census and admission counts that explains why there are few if any hospital admission reports for covid-19 worldwide. Hospital census is a measure that is easily generated by simply counting how many patients have a confirmed covid-19 diagnosis (usually measured at mid-night). Counting a new admission is straightforward 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? New admissions covid-19 is a measure that is more difficult to generate because people are counted if have covid-19 either at the time of admission or retrospectively classified as a new admission if or when covid-19 is confirmed later during their hospital stay. A line listing of covid-19 patients is required to address these concerns.

Posted on Authorea 25 Jul 2020 — The copyright holder is the author/funder. All rights reserved. No reuse without permission. — https://doi.org/10.22541/au.159569198.82432694 — This a preprint and has not been peer reviewed. Data may be preliminary.

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

A line listing includes individual key covid-19 patient information that is updated daily until discharge or illness resolution (**Fig. 3**). 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. Epidemiologists routinely report epidemic curves of cases based on the date of illness onset. Covid-19 surveillance similarly benefits from reporting hospitalization based on the date of illness or admission. However, hospitals do not routinely report this information, and the creation of a line listing of hospital patients is often not well supported in electronic medical record systems.



PATIENT #1	MARCH 24	MARCH 27	APRIL 12	APRIL 22
COVID-19 status	Suspect	Suspect	Confirmed	Confirmed
Viral days	N/A	N/A	10	18
Admission date	N/A	March 27	March 27	March 27
Acute care - length of stay	N/A	0	7	15
ICU - length of stay	N/A	N/A	10	12
Ventilator - length of stay	N/A	N/A	4	5
Discharge date	N/A	N/A	N/A	April 22

Figure 3: **Line Listing for a single patient**

Many hospitals have overcome the challenges to generate covid-19 patient data to support research. However, there are additional challenges to use 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. Instead, public health staff update their public health information line list with patient hospital through a laborious process of phone calls or searches in hospital information systems to which they've been granted access.

Opportunity and Caution

There are signs that Canada and other jurisdictions are responding by adapting existing information systems and developing new ones to connect covid-19 patient information from different health care sectors. In our

Posted on Authorea 25 Jul 2020 — The copyright holder is the author/funder. All rights reserved. No reuse without permission. — https://doi.org/10.22541/au.159569198.82432694 — This a preprint and has not been peer reviewed. Data may be preliminary.

local jurisdiction, all hospitals worked together to create a standard line listing that is transmitted to the regional public health department where is it merged with their information systems. The public health information system serves as the central point for covid-19 information and forms the basis of contract 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.

References

Recommendations and Guidelines on data sharing. (2020). [Report]. https://doi.org/https://doi.org/10.15497/rda00052

An interactive web-based dashboard to track COVID-19 in real time. (2020). [Journal Article]. *The Lancet Infectious Diseases*.

Latest assessment on COVID-19 from the European Centre for Disease Prevention and Control (ECDC). (2020). [Journal Article]. *Eurosurveillance*, *25*(8).

The COVID Tracing Project. (2020). [Online Database]. https://covidtracking.com/contact

Covid-19 Hospitalizations in Belgium. (2020). [Online Database]. https://rpubs.com/JMBodart/Covid19-hosp-be

Report of the WHO-China Joint Mission on Coronovirus Disease 2019 (COVID-19). (2020). [Report]. https://www.who.int/publications-detail/report-of-the-who-china-joint-mission-on-coronavirus-disease-2019-(covid-19)

Using the ACT network to gain insight into COVID-19. (2020). [Online Database]. https://clic-ctsa.org/events/using-act-network-gain-insight-covid-19

Projecting demand for critical care beds during COVID-19 outbreaks in Canada. (2020). [Journal Article]. *Canadian Medical Association Journal*, cmaj.200457. https://doi.org/10.1503/cmaj.200457