

Article details: 2020-0246	
Title	Descriptive epidemiology of patients hospitalized with laboratory-confirmed COVID-19 in a network of Canadian acute-care hospitals, March 1 to August 31, 2020
Authors	Robyn Mitchell MHSc, Kelly Baekyung Choi MSc, Linda Pelude MSc, Wallis Rudnick PhD, Nisha Thampi MD, Geoffrey Taylor MD; for the CNISP COVID-19 Working Group*
Reviewer 1	Alon Vaisman
Institution	Infection Control, University Health Network, Toronto, Ont.
General comments (author response in bold)	<p>This paper describes a sample of COVID-19 cases from hospitalized patients across Canada. It is well-written and the rationale is very clear. The Discussion section is succinct, easy to follow, and intelligently written. Although the findings are not novel in general, they represent an important mile marker for Canadian cases of COVID-19. The main area of improvement is the Results section and interpretation of values.</p> <p>Introduction</p> <p>1) Page 3, line 57: How was this definition chosen? Why not 5 days? Each province may have differing definitions on this, so it would be helpful for a reference or justification for selecting 7 days.  <b>Thank you for this suggestion. We added the following information regarding how the definition of healthcare acquisition was determined (line 62-64). Seven days was chosen as the cut-off to attribute acquisition to the hospital based on an estimate of a median incubation of 4 days (IQR 2-7 days) for hospitalized patients with COVID-19.</b>  <a href="https://www.cdc.gov/coronavirus/2019-ncov/hcp/clinical-guidance-management-patients.html">https://www.cdc.gov/coronavirus/2019-ncov/hcp/clinical-guidance-management-patients.html</a> (accessed March 2020). We applied a conservative cut-off using the upper quartile of 7 days.</p> <p>2) Page 3, line 60: Change 'epi' to 'epidemiological' link  <b>Thank you for this suggestion. We have updated line 62: 'epi' has been changed to 'epidemiological'</b></p> <p>Results</p> <p>3) Line 76 – “The proportion of males in the 50–64 year age group was significantly higher compared to the proportion in other age groups (p=0.0008) (Table 2).” – I don't think this is accurate. Although not explicitly stated, I assume a Chi Square/Fisher exact test was performed here based on what's written in the Methods section – a comparison of proportions? The null hypothesis for the test is that sex and age are independent. If the p value indicates statistical significance, the null hypothesis is rejected, and you've demonstrated the two variables are not independent. In other words, the p values are non-directional (two-tailed). It does not necessarily tell you that a specific group has a characteristic at more or less frequency. Therefore, this phrasing should be changed. Alternatively, a different statistical test should be performed whereby you compare 50-64 year olds to all other patients. In this Chi test, if 50-64 year olds have a significantly higher proportion of males, then the statement is accurate (although I doubt this what is being conveyed in the Table).  <b>Thank you for this suggestion. This sentence was removed and replaced with line 81-82 “The proportion of patients hospitalized with COVID-19 significantly differed by age group (Table 2).”</b></p>

4) Line 79 – why is the denominator 15 here and not 20?

**Thank you for this note. Data for this variable was not available for all cases. Therefore, a sentence (line 74-75) was added to the Methods section to indicate that missing and incomplete data were excluded from the analyses and that denominators may vary.**

5) Line 87 – are 53 and 71 years median or means? Please clarify

**Thank you for this note. The results have been updated (line 90) to reflect that these are median ages.**

6) Line 90 – isn't the word 'hospitalized' redundant, as all the patients in this study are hospitalized patients? The way it's phrased here suggests there are some non-hospitalized patients in this study.

**Thank you for this suggestion. We removed "among all hospitalized cases" and replaced with "among all patients" (line 94).**

7) Line 93 – the total number of patients mentioned here –  $313 + 686 = 999$ . Why is this value lower than the total of 1 030? Likewise, Line 95 – why is the denominator 867 and not 1 030? This appears to be an issue in multiple areas in the paper. I'm assuming it's because of missing data? For clarification, it would be helpful in the Methods section to describe how missing data was dealt with in the analysis.

**Thank you for this suggestion. Data for some variables were not available for all cases. Therefore, a sentence (line 74-75) was added to the methods to indicate that missing and incomplete data were excluded from analyses and that denominators may vary.**

8) Line 96 – what statistical test proves this? Again, a Fisher exact/Chi square test would only tell you that age and co-morbidity are not independent, but does not tell you which direction this association is. These tests do not know there are ranks to the dependent variable (age) such that you could say that the proportions rise with age. The p value here is non-directional (two-tailed). Was there a different test done here?

**Thank you for this note. The direction of association was removed and the sentence (line 99-101) was reworded to "The majority of patients (85.8%, 1,602/1,876) had at least one underlying medical condition and the presence of medical conditions significantly differed by age ( $p < 0.0001$ )"**

9) Line 102 – The meaning of 'significant differences' may be interpreted differently between readers, so this should be spelled out here more clearly, or at least the key findings highlighted here.

**Thank you for this suggestion. We removed the sentence "Significant differences in symptoms were identified between age groups" and replaced with the following key finding based on updated data (line 107-110) "Younger patients (<40 years) were less likely to report symptoms compared to patients 40 years or older (84.8% vs. 94.4%,  $p < 0.0001$ ). Common reasons for admission among asymptomatic younger patients included trauma, mental health, and labour or pregnancy related complications."**

10) Line 110 - The interpretation of the p-value variables of age and ventilation has

the same issue as age and males (first comment above)

**Thank you for this note. The direction of association was removed and the sentence (line 116-117) was reworded to “The proportion of patients who received mechanical ventilation and/or ICU admission significantly differed by age group (Table 3).”**

Discussion

11) Line 154 – While true, this paper did not look at public health measures or how they relate to the epi curves at all, so this statement should be removed.

**Thank you for this suggestion. The authors agree and we have removed the following sentence “COVID-19 hospitalization rates indicate that, with a few exceptions, public health measures were able to mitigate a surge in patients during the first pandemic wave”.**

12) Line 161 – What direction would the selection bias go as a result of an oversampling of academic centres? More young patients? Less males? More co-morbidities? This statement needs to contain more direction than simply saying ‘may not be generalizable.’

**Thank you for this suggestion. To provide more direction, we updated the sentence (line 158-159) to “CNISP hospitals are predominantly large teaching hospitals, which may receive more severe cases as referral centres, and results may not be generalizable to all Canadian acute care facilities.”**

Tables and Figures:

13) Table 2 and Table 3: there is no explanation of what the p value represents. What statistical test was performed here? Based on methods, likely Fisher or Chi – please be explicit.

**Thank you for this suggestion. We have added a footnote to both Tables 2 and 3 to indicate which statistical tests were used.**

14) Table 2 is confusing in its total patient values. The n is listed as the first row for each age group and total. In the middle of the table, “Underlying Medical Conditions” – a small subset is listed for the n value. Is this because the smaller value represents the total number of people who had an underlying condition, or the total number of people where this information was available?

**Thank you for this suggestion. Tables 2 and 3 have been updated as follows: an n value is reported as the denominator for each data point to reflect the total number of patients for whom this information was available and the n value from the first row has been removed to avoid confusion.**

15) Figure 1: To improve visual representation, it may be helpful to place death as the lowest color rather than in the middle. Also, it may be helpful to collapse transferred and discharged into one category, as the fact that a patient is transferred is not very helpful information or different from discharged.

**Thank you for this suggestion. We have removed Figure 1 as these data are provided in Table 3.**

16) Figure 2: The legend for colors is missing Central and East

**Figure 2 (now Figure 1) has been updated as follows: the legend for Central and East regions are now visible**

Reviewer 2

Xiang Han

Institution	Department of Laboratory Medicine, The University of Texas MD Anderson Cancer Center, Houston, Tex.
General comments (author response in bold)	<p>Dr. Mitchell and colleagues from CNISP collected data on 1030 patients with COVID-19 in 31 sentinel acute care hospitals in Canada and performed an analyses of these patients. This study is straight-forward, and findings may inform decision making in the care for patients with COVID-19. This reviewer finds a number of opportunities, all minor and in data presentation and information flow, for quality improvement.</p> <p>1) Line 78: The use of 30% should be omitted here because it is redundant and causes confusion with the beginning 1.9%. Instead, it is better to use “six of the twenty children ....’  <b>Thank you for this suggestion. This sentence (line 83) was updated to “9 of the 37 children (24.3%) were less than 1 year of age and 10 (27%) were 1-4 years of age.”</b></p> <p>2) Line 79: The use of 30% should be omitted here too. In the next sentence, it is sudden here to a reader as to who are the 15 patients, instead of 20. Are they all symptomatic ones with COVID-19? Within the same sentence, the use of two patients with a 10% is also perplexing.  <b>Thank you for this suggestion. This sentence (line 83-85) was updated to “Fifty percent (18/36) had an underlying medical condition and 16.2% (6/37) were admitted to an intensive care unit (ICU)”</b></p> <p>3) Throughout the text, the denominators of various percentages change often and cause some confusion.  <b>Thank you for this suggestion. Data for some variables were not available for all cases. Therefore, a sentence (line 74-75) was added to the methods to indicate that missing and incomplete data were excluded from analyses and that denominators may vary.</b></p> <p>4) Line 118: This last sentence of the paragraph is better relocated to follow the first sentence on such all cause 30-day mortality. This allows smooth information flow and continuity.  <b>Thank you for this suggestion. This sentence “Thirty-day all-cause mortality was 19 times higher among patients with a reported underlying medical condition than those without (95.1% vs. 4.9%, p&lt;0.001)” was moved to follow the first sentence of the paragraph (line 120-121).</b></p> <p>5) Line 133-136: These sentences on pregnant women are better to be relocated to results to combine with the line 99 sentence to form a new paragraph, ideally following the paragraph on children.  <b>Thank you for this suggestion. These sentences were removed as they were no longer applicable when the data were updated.</b></p> <p>6) Table 1. A new row is better added to show the total of hospitals, cases and %.  <b>Thank you for this suggestion. The first row is a header which now indicates Province, Number of reporting hospitals, Number of cases, Proportion of cases.</b></p>

7) Table 2: It is better to show a new row of 'Lack of symptoms' as the first row of symptoms (above cough). This ensures readability of the table and information flow. If this is the case, brief narration of asymptomatic patients may be needed in the text.

**Thank you for this suggestion. In Table 2 the denominator data have been added for each data point to improve clarity. A row "any symptom" was added to Table 2 and discussed in the text regarding younger patients reporting few symptoms than older patients (line 107-110).**

8) Table 3: The numbers of death and attributed death do not add up to all ages. Why is that? Any way of correction or comment on?

**Thank you for this suggestion. We have added a footnote to both Tables 2 and 3 to indicate that age is missing for three cases, therefore the denominators for the age groups will not add up to the total (all cases).**

9) Figure 2: The PDF conversion of the graph likely lost legends for the orange line and gray line. The authors may need to be aware of this.

**Thank you for this note. Figure 2 has been updated as follows: the legend for Central and East regions are now visible**