

## Peer Review Report

# Review Report on The first 110 593 COVID-19 patients hospitalised in Lombardy: a regionwide analysis of case characteristics, risk factors and clinical outcomes

Original Article, Int J Public Health

Reviewer: Valentina Gallo

Submitted on: 14 Oct 2021

Article DOI: 10.3389/ijph.2022.1604427

### EVALUATION

#### **Q 1** Please summarize the main findings of the study.

This is an interesting paper reporting some statistics of about 92K COVID+ individuals admitted in hospital between February and December 2020 in Lombardy. The topic is important and the wealth of data can provide insights into one of the most affected regions in Italy, and in Europe.

#### **Q 2** Please highlight the limitations and strengths.

Three main methodological issues hamper the interpretations of the data as they stand; in addition, a few, some important, remarks on data interpretations are reported below.

Firstly, it is not clear how age was dealt with in the analysis, and what type of data was available. What is the underlying time variable in Cox models? If age in one year (of few years) age band are available, why only dichotomous data <70/70+ are presented in the final analysis? With the mortality from COVID so strongly associated with age, this is a main limitation in data interpretation. Also, I would like to be reassured that the Cox proportional Hazard assumption is met (I doubt it).

Secondly, it is not clear from the methods section how many and which comorbidities were identified on the database, and how they were dealt with. The notion that increasing the number of co-morbidities increases the mortality of COVID+ patients is well established now, and this analysis does not add much to the current body of evidence. Would it be possible to estimate the HR of mortality in diabetic vs. non diabetic for example? Hypertensives vs. no hypertensive? Etc. I think this would add considerably to the current discussion for example in priority in vaccine administration and how to identify most at risk categories.

Thirdly, mortality ascertainment. What have you done to try and capture those dying at home before reaching the hospital? Those dying at home after hospital discharge? Have you attempted a linkage with mortality records?

#### **Q 3** Please provide your detailed review report to the authors. The editors prefer to receive your review structured in major and minor comments. Please consider in your review the methods (statistical methods valid and correctly applied (e.g. sample size, choice of test), is the study replicable based on the method description?), results, data interpretation and references. If there are any objective errors, or if the conclusions are not supported, you should detail your concerns.

This is an interesting paper reporting some statistics of about 92K COVID+ individuals admitted in hospital between February and December 2020 in Lombardy. The topic is important and the wealth of data can provide insights into one of the most affected regions in Italy, and in Europe.

Three main methodological issues hamper the interpretations of the data as they stand; in addition, a few, some important, remarks on data interpretations are reported below.

Firstly, it is not clear how age was dealt with in the analysis, and what type of data was available. What is the underlying time variable in Cox models? If age in one year (of few years) age band are available, why only dichotomous data <70/70+ are presented in the final analysis? With the mortality from COVID so strongly associated with age, this is a main limitation in data interpretation. Also, I would like to be reassured that the Cox proportional Hazard assumption is met (I doubt it).

Secondly, it is not clear from the methods section how many and which comorbidities were identified on the database, and how they were dealt with. The notion that increasing the number of co-morbidities increases the mortality of COVID+ patients is well established now, and this analysis does not add much to the current body of evidence. Would it be possible to estimate the HR of mortality in diabetic vs. non diabetic for example? Hypertensives vs. no hypertensive? Etc. I think this would add considerably to the current discussion for example in priority in vaccine administration and how to identify most at risk categories.

Thirdly, mortality ascertainment. What have you done to try and capture those dying at home before reaching the hospital? Those dying at home after hospital discharge? Have you attempted a linkage with mortality records?

Once these issues are appropriately addressed, here there are a number of comments on the paper, its presentation, and the interpretation of the data.

1. The title is misleading. The unit of analysis is individual with COVID admitted to hospital, not the number of hospitalisations, please replace the figures in the title.
2. Abstract: more details of type of data collected are needed, the result section might be revised accordingly with methodological points above
3. Ageing population is probably one of the most important factors explaining higher mortality in Lombardy compared to other European regions. In the introduction (p.2, lines 48–53) please report median age (p25; p75) of the Lombardy populations instead of life expectancy
4. Introduction. When giving the local context (which is important, and nicely pictured) it is not possible to mention the relatively recent changes that Italy witnessed as consequence of the regionalisations and privatisation of the health system. An increasing wealth of evidence is pointing towards a crucial role of healthcare policy in COVID19 management and outcome
5. Is it possible from your DB to estimate something on long COVID? I believe that – even this is only a rough estimate – this type of evidence is desperately needed, and would be particularly welcome to the international scientific community
6. Methods. These need to be more detailed. Which variables were you able to identify and how? How were they coded? This is true for age (see above) but also for comorbidities (how many, defined how, etc.). Please state here that you do not have any relevant data to estimate socio-economic position or ethnicity (only mentioned at the end of the discussion)
7. Were data from outpatients clinics or diurnal hospitalisation (day hospital) available? Would these be useful for estimating long COVID?
8. In the methods report how the CFR was calculated
9. In the methods, please give some context on the hospitalisation procedures as these may vary across countries, i.e. anyone could self refer to the emergency room, the admission was free of charge, everyone is covered by the national health system, etc.
10. Results. I can see a lot of room for transforming Table 1 in a Figure. It would be important to add to the figure all the relevant trends, so to allow comparison. For example by superimposing two line graphs (with two separate axes if needed) and one or two histograms. For example, see below
11. What can you do and say to reassure the reader that after eliminating duplicates, the matching of hospital episode statistics with individuals was in fact correct? What type of quality checks have you run? How do the figures of re-hospitalisation compare with other research? Clearly, non-matched individuals would artificially lower the CFR, and if this is more likely systematically in one category than another, you would have introduced a bias which might bias your final results
12. Please do not read tables aloud in the text. Stick to very important figures only, and defer the rest of the data to the table/graph
13. It would be very useful to see the row data in Table 1 (hopefully to become Figure 1) also by type of hospital (private vs. public)
14. Is the (%) in the third column of Table 2 a CFR? If yes, please report this more explicitly
15. Interpretation. There are two results that need a much more careful interpretation in this paper: the public/private differences in mortality and the primary/secondary. It is alarming that these were not mentioned anywhere and analysed in depth in the discussion, as one potential explanation is bias or confounding by socio-economic status. I think that without attempting really dissect what is going on with these findings, for example bringing in some demographic data from other sources, the whole credibility of the paper is fundamentally hampered

16. Another important piece of information that you have but you have not commented upon is why mortality is higher when hospitalisation is higher. How do you explain that? With a biological effect of more severe infections during waves? With a healthcare systems overload (overwhelm) during waves? What can we learn from this very interesting piece of information? Moreover I find this result the most intriguing of them all, and relatively new on the scientific landscape.

#### PLEASE COMMENT

**Q 4** Is the title appropriate, concise, attractive?

1. The title is misleading. The unit of analysis is individual with COVID admitted to hospital, not the number of hospitalisations, please replace the figures in the title.

**Q 5** Are the keywords appropriate?

add epidemiology, hospital records, hospital admission statistics

**Q 6** Is the English language of sufficient quality?

yes

**Q 7** Is the quality of the figures and tables satisfactory?

No.

**Q 8** Does the reference list cover the relevant literature adequately and in an unbiased manner?)

yes

#### QUALITY ASSESSMENT

**Q 9** Originality



**Q 10** Rigor



**Q 11** Significance to the field



**Q 12** Interest to a general audience



**Q 13** Quality of the writing



**Q 14** Overall scientific quality of the study



#### REVISION LEVEL

**Q 15** Please make a recommendation based on your comments:

Major revisions.