nature machine intelligence

MATTERS ARISING

https://doi.org/10.1038/s42256-020-00253-3

Check for updates

Replication of a mortality prediction model in Dutch patients with COVID-19

Marian J. R. Quanjel[1]✉, Thijs C. van Holten[2], Pieternel C. Gunst-van der Vliet[3], Jette Wielaard[4], Bekir Karakaya[1], Maaike Söhne[5], Hazra S. Moeniralam[6] and Jan C. Grutters[1,7]

ARISING FROM Yan et al. *Nature Machine Intelligence* https://doi.org/10.1038/s42256-020-0180-7 (2020)

There is a clear need for a simple mortality prediction model to help guide clinical decision making for patients with COVID-19. Yan et al. demonstrated the strong predictive capacity of a decision rule consisting of three readily available laboratory measures for COVID-19 mortality: lactate dehydrogenase (LDH), high-sensitivity C-reactive protein (hs-CRP) and percent lymphocytes[1].

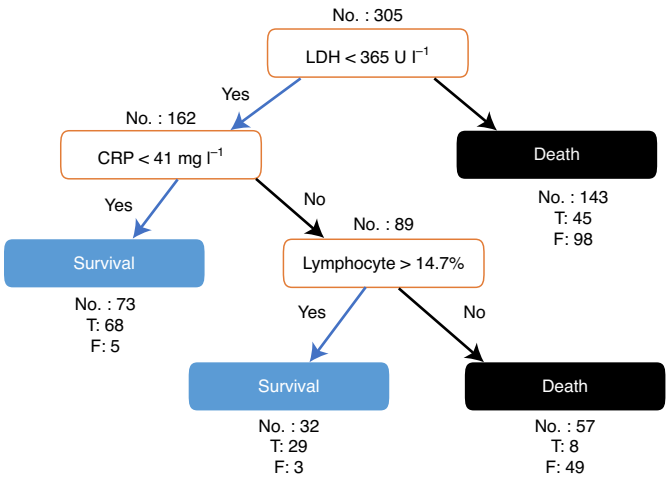
We performed an independent replication of their model using data from our large general hospital (St Antonius Hospital, Nieuwegein, the Netherlands). We included 305 patients over 18 years of age who presented to the emergency room with a clinical suspicion of COVID-19 between 19 March and 4 May 2020, with a positive SARS-CoV-2 polymerase chain reaction (PCR) result and for whom data for the three biomarkers were available at presentation. Mortality data were collected on 14 May 2020 at an average follow-up duration of 41 days.

The average age of the patients was 62.7 years and 188 (62%) were male. Of these patients, 61 died (at 1–33 days after admission, with a median of 7 days). We found that the model has 92% survival prediction accuracy but only 27% mortality prediction accuracy. This compares to a 100% survival prediction and 81% mortality prediction found by Yan et al.

In our population, 42 out of 303 patients were admitted to the intensive care unit (ICU). Of these patients, 36 (86%) had an unfavourable outcome from the decision rule, but only 15 (36%) of them died (all with LDH > 365 U l^{-1}).

We conclude that, in Dutch patients, a favourable outcome of the decision rule was indeed a good predictor of non-admission to the ICU and of survival. Although an unfavourable outcome of the decision rule could have been interpreted as a warning sign, the majority of our patients thus classified still survived. We hypothesize that this discrepancy between our data and those of Yan et al. may be due to genetic differences in the expression of the presented biomarkers. For example, LDH expression has been reported to display substantial genetic heterogeneity between Asians and Caucasians[2]. Alternatively, differences in treatment protocol or in baseline characteristics of the patients may have influenced the outcome.

In conclusion, our analysis supports the high survival prediction accuracy of the decision rule proposed by Yan et al., but fails to confirm its high mortality prediction accuracy. The identification of patients with COVID-19 with a low risk of mortality can be useful to inform the level of surveillance within or outside the hospital.



The decision rule using three key features and their thresholds in absolute value. Num, the number of patients in a class; T, the number of correctly classified; F, the number of misclassified patients.

Reporting Summary

Further information on research design is available in the Nature Research Reporting Summary linked to this Article.

Data availability

The data that support the findings of this study are available in the Supplementary Information.

Received: 18 May 2020; Accepted: 1 October 2020;
Published online: 12 November 2020

References

1. Yan, L. et al. An interpretable mortality prediction model for COVID-19 patients. *Nat. Mach. Intell.* **2**, 283–288 (2020).
2. Lv, J. et al. Prognostic value of lactate dehydrogenase expression in different cancers: a meta-analysis. *Am. J. Med. Sci.* **358**, 412–421 (2019).

Author contributions

M.J.R.Q. conceived of the presented idea. M.J.R.Q., T.C.v.H., P.C.G.-v.d.V. and J.W. extracted and analysed the data. All authors discussed the results and contributed to the final manuscript.

[1]St Antonius ILD Center of Excellence, Department of Pulmonology, St Antonius Hospital, Nieuwegein, The Netherlands. [2]Department of Clinical Chemistry, St Antonius Hospital, Nieuwegein, The Netherlands. [3]Department of Information and Intelligence, St. Antonius Hospital, Nieuwegein, The Netherlands. [4]Department of Medical Physics, St Antonius Hospital, Nieuwegein, The Netherlands. [5]Department of Internal Medicine, St Antonius Hospital, Nieuwegein, The Netherlands. [6]Department of Internal Medicine and Intensive Care, St Antonius Hospital, Nieuwegein, The Netherlands. [7]Division of Heart and Lungs, UMC Utrecht, Utrecht, The Netherlands. ✉e-mail: m.quanjel@antoniusziekenhuis.nl

Competing interests
The authors declare no competing interests.

Additional information
Supplementary information is available for this paper at https://doi.org/10.1038/s42256-020-00253-3.

Correspondence and requests for materials should be addressed to M.J.R.Q.

Reprints and permissions information is available at www.nature.com/reprints.

Publisher's note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

nature research

Corresponding author(s): NATMACHINTELL-MA20051777

Last updated by author(s): May 20, 2020

Reporting Summary

Nature Research wishes to improve the reproducibility of the work that we publish. This form provides structure for consistency and transparency in reporting. For further information on Nature Research policies, see our Editorial Policies and the Editorial Policy Checklist.

Statistics

For all statistical analyses, confirm that the following items are present in the figure legend, table legend, main text, or Methods section.

n/a	Confirmed	
☐	☒	The exact sample size (*n*) for each experimental group/condition, given as a discrete number and unit of measurement
☐	☒	A statement on whether measurements were taken from distinct samples or whether the same sample was measured repeatedly
☒	☐	The statistical test(s) used AND whether they are one- or two-sided *Only common tests should be described solely by name; describe more complex techniques in the Methods section.*
☒	☐	A description of all covariates tested
☒	☐	A description of any assumptions or corrections, such as tests of normality and adjustment for multiple comparisons
☒	☐	A full description of the statistical parameters including central tendency (e.g. means) or other basic estimates (e.g. regression coefficient) AND variation (e.g. standard deviation) or associated estimates of uncertainty (e.g. confidence intervals)
☒	☐	For null hypothesis testing, the test statistic (e.g. *F*, *t*, *r*) with confidence intervals, effect sizes, degrees of freedom and *P* value noted *Give P values as exact values whenever suitable.*
☒	☐	For Bayesian analysis, information on the choice of priors and Markov chain Monte Carlo settings
☒	☐	For hierarchical and complex designs, identification of the appropriate level for tests and full reporting of outcomes
☒	☐	Estimates of effect sizes (e.g. Cohen's *d*, Pearson's *r*), indicating how they were calculated

Our web collection on statistics for biologists contains articles on many of the points above.

Software and code

Policy information about availability of computer code

Data collection	EPIC Caboodle. MicrosoftSQL server 2016 Management Studio v 13.0.16106.4
Data analysis	Excel

For manuscripts utilizing custom algorithms or software that are central to the research but not yet described in published literature, software must be made available to editors and reviewers. We strongly encourage code deposition in a community repository (e.g. GitHub). See the Nature Research guidelines for submitting code & software for further information.

Data

Policy information about availability of data

All manuscripts must include a data availability statement. This statement should provide the following information, where applicable:

- Accession codes, unique identifiers, or web links for publicly available datasets
- A list of figures that have associated raw data
- A description of any restrictions on data availability

The data that support the findings of this study are available in the supplementary information.

Field-specific reporting

Please select the one below that is the best fit for your research. If you are not sure, read the appropriate sections before making your selection.

☒ Life sciences ☐ Behavioural & social sciences ☐ Ecological, evolutionary & environmental sciences

For a reference copy of the document with all sections, see nature.com/documents/nr-reporting-summary-flat.pdf

Life sciences study design

All studies must disclose on these points even when the disclosure is negative.

Sample size	305 patients
Data exclusions	No data is excluded
Replication	No replication
Randomization	no randomization
Blinding	no blinding

Reporting for specific materials, systems and methods

We require information from authors about some types of materials, experimental systems and methods used in many studies. Here, indicate whether each material, system or method listed is relevant to your study. If you are not sure if a list item applies to your research, read the appropriate section before selecting a response.

Materials & experimental systems

n/a	Involved in the study
☒	☐ Antibodies
☒	☐ Eukaryotic cell lines
☒	☐ Palaeontology and archaeology
☒	☐ Animals and other organisms
☒	☐ Human research participants
☐	☒ Clinical data
☒	☐ Dual use research of concern

Methods

n/a	Involved in the study
☒	☐ ChIP-seq
☒	☐ Flow cytometry
☒	☐ MRI-based neuroimaging

Clinical data

Policy information about clinical studies

All manuscripts should comply with the ICMJE guidelines for publication of clinical research and a completed CONSORT checklist must be included with all submissions.

Clinical trial registration	There was no registration needed
Study protocol	It's not available because it was a retrospective cohort study
Data collection	Retrospective cohortstudy, data extraction from files
Outcomes	Primary outcome death, secondary outcome survival