Deriving and validating a risk prediction model for long COVID-19: protocol for an observational cohort study using linked Scottish data

SUPPLEMENTARY FILE

Diagnostic codes for Long COVID

Local codes introduced in Scotland[1]

Clinical Computer System	Code	Description
EMIS PCS	^ESCT1348648	Ongoing symptomatic COVID-19
EMIS PCS	^ESCT1348645	Post-COVID-19 syndrome
Vision	A7955	Ongoing symptomatic COVID-19
Vision	AyuJC	Post-COVID-19 syndrome

ICD-10 emergency use codes for conditions related to COVID-19[2]

ICD-10 Code	Description
U07.3	Personal history of COVID-19
U07.4	Post-COVID-19 condition
U07.5	Multisystem inflammatory syndrome associated with COVID-19

Daines L, *et al. BMJ Open* 2022; 12:e059385. doi: 10.1136/bmjopen-2021-059385

Tripod checklist: Prediction model development and validation[3]

Section/Topic	Item		Checklist Item	Page
Title and abstract				
Title	1	D;V	Identify the study as developing and/or validating a multivariable prediction model, the target population, and the outcome to be predicted.	1
Abstract	2	D;V	Provide a summary of objectives, study design, setting, participants, sample size, predictors, outcome, statistical analysis, results, and conclusions.	2
Introduction				
Background and objectives	3a	D;V	Explain the medical context (including whether diagnostic or prognostic) and rationale for developing or validating the multivariable prediction model, including references to existing models.	4/5
	3b	D;V	Specify the objectives, including whether the study describes the development or validation of the model or both.	5
Methods				
Source of data	4a	D;V	Describe the study design or source of data (e.g., randomized trial, cohort, or registry data), separately for the development and validation data sets, if applicable.	5/6
	4b	D;V	Specify the key study dates, including start of accrual; end of accrual; and, if applicable, end of follow-up.	5/6
Participants	5a	D;V	Specify key elements of the study setting (e.g., primary care, secondary care, general population) including number and location of centres.	5/6
	5b	D;V	Describe eligibility criteria for participants.	5
	5c	D;V	Give details of treatments received, if relevant.	NA
Outcome	6a	D;V	Clearly define the outcome that is predicted by the prediction model, including how and when assessed.	7
	6b	D;V	Report any actions to blind assessment of the outcome to be predicted.	NA
Predictors	7a	D;V	Clearly define all predictors used in developing or validating the multivariable prediction model, including how and when they were measured.	8
	7b	D;V	Report any actions to blind assessment of predictors for the outcome and other predictors.	NA
Sample size	8	D;V	Explain how the study size was arrived at.	5
Missing data	9	D;V	Describe how missing data were handled (e.g., complete-case analysis, single imputation, multiple imputation) with details of any imputation method.	9
Statistical analysis methods	10a	D	Describe how predictors were handled in the analyses.	12
	10b	D	Specify type of model, all model-building procedures (including any predictor selection), and method for internal validation.	12
	10c	V	For validation, describe how the predictions were calculated.	NA
	10d	D;V	Specify all measures used to assess model performance and, if relevant, to compare multiple models.	NA
	10e	V	Describe any model updating (e.g., recalibration) arising from the validation, if done.	NA
Risk groups	11	D;V	Provide details on how risk groups were created, if done.	NA
Development vs. validation	12	V	For validation, identify any differences from the development data in setting, eligibility criteria, outcome, and predictors.	NA
Results				
Participants	13a	D;V	Describe the flow of participants through the study, including the number of participants with and without the outcome and, if applicable, a summary of the follow-up time. A diagram may be helpful.	NA
	13b	D;V	Describe the characteristics of the participants (basic demographics, clinical features, available predictors), including the number of participants with missing data for predictors and outcome.	NA
	13c	V	For validation, show a comparison with the development data of the distribution of important variables (demographics, predictors and outcome).	NA
Model development	14a	D	Specify the number of participants and outcome events in each analysis.	NA
	14b	D	If done, report the unadjusted association between each candidate predictor and outcome.	NA
Model specification	15a	D	Present the full prediction model to allow predictions for individuals (i.e., all regression coefficients, and model intercept or baseline survival at a given time point).	NA
	15b	D	Explain how to the use the prediction model.	NA
Model performance	16	D;V	Report performance measures (with CIs) for the prediction model.	NA
Model-updating	17	V	If done, report the results from any model updating (i.e., model specification, model performance).	NA
Discussion				
Limitations	18	D;V	Discuss any limitations of the study (such as nonrepresentative sample, few events per predictor, missing data).	NA
Interpretation	19a	V	For validation, discuss the results with reference to performance in the development data, and any other validation data.	NA
	19b	D;V	Give an overall interpretation of the results, considering objectives, limitations, results from similar studies, and other relevant evidence.	NA
Implications	20	D;V	Discuss the potential clinical use of the model and implications for future research.	NA
Other information				
Supplementary information	21	D;V	Provide information about the availability of supplementary resources, such as study protocol, Web calculator, and data sets.	NA
Funding	22	D;V	Give the source of funding and the role of the funders for the present study.	

Daines L, *et al. BMJ Open* 2022; 12:e059385. doi: 10.1136/bmjopen-2021-059385

REFERENCES

1) Scottish Government. Management and recording of the long-term effects of COVID-19 (Long COVID). March 2020. Available from: https://www.scimp.scot.nhs.uk/wp-content/uploads/CMO-Letter-09.03.21.pdf (Accessed November 2021)

2) Public Health Scotland. Scottish Clinical Coding Standards. Number 27. https://www.isdscotland.org/products-and-services/terminology-services/clinical-coding-guidelines/Docs/Scottish-clinical-coding-standards-Feb-2021-No-27.pdf (Accessed November 2021)

3) Moons KG, Altman DG, Reitsma JB, et al. Transparent Reporting of a multivariable prediction model for Individual Prognosis or Diagnosis (TRIPOD): explanation and elaboration. Annals of internal medicine. 2015;162(1):W1-73.