SECTION A: RESEARCHER / RESEARCH TEAM INFORMATION

Lead Applicant Details

Title	Professor
First name (given name)	Ann
Surname (family name)	Nicholson
Gender	Female
Position at employing organisation/ institution	Dean (Research), Professor
ORCID ID (https://orcid.org) or URL to academic profile	https://orcid.org/0000-0002-2269-9823
Email	[redacted]

Employing Organisation/Institution

*Institution with a remit including health, research or academic pursuit, and with legal status which includes the scope to sign the **Data Transfer Agreement.***

Institution name	Monash University
City, Country	Melbourne, Australia
Does your institution agree to execute the Data Transfer Agreement? (if your application is approved)	**YES**

Co-applicants

ALL individuals accessing the data must be listed. Any additions must be notified to the COVID-19 Data Access Committee. Add rows as necessary.

1. Name	Tom Snelling
1. Position / Role in analysis	Professor of Public health and Infectious Diseases Physician - Clinical Advisor
1. Organisation/Institution	University of Sydney
2. Name	Steven Mascaro
2. Position / Role in analysis	Senior Research Fellow - Mathematical Modeller
2. Organisation/Institution	Monash University
3. Name	Yue Wu
3. Position / Role in analysis	Senior Research Officer- Mathematical Modeller
3. Organisation/Institution	University of Sydney

4. Name	Owen Woodberry
4. Position / Role in analysis	Research Fellow
4. Organisation/Institution	Monash University - Mathematical Modeller
5. Name	Ross Pearson
5. Position / Role in analysis	Project Manager
5. Organisation/Institution	Monash University
6. Name	Jessica Ramsay
6. Position / Role in analysis	Research Assistant - Project Coordinator
6. Organisation/Institution	Telethon Kids Institute
7. Name	Todd Cooper
7. Position / Role in analysis	Project Officer (Data Science)- Data Management
7. Organisation/Institution	University of Sydney

Conflicts of Interest

List details of any existing or perceived conflicts of interest (financial or non-financial) that exist relating to the use of the requested data by the data requestor and/or co-applicants (see ICMJE.org for the definition of conflicts of interest)

Applicants declare no conflicts of interest.

COVID-19 Data Platform - Data Access Application Form

Please review the Data Access Guidelines and the Data Transfer Agreement before completing this form. A complete application should address all of the Review Considerations outlined in the Data Access Guidelines. Note that the details of all approved applications will be made publicly available on the COVID-19 Data Platform website.

Complete all sections of this form fully and return to covid19@iddo.org.

SECTION B: RESEARCH PLAN

Title of Proposed Research	COVID-Intelligence, a Bayesian Network Clinical Decision Support System for COVID-19
Is this a re-submission of a previous application to the COVID-19 DAC? If yes, provide the submission date of the previous application.	No. Approval was granted by the ISARIC 4C independent data and materials access committee (IDAMAC) for a collaborative analysis of 3000 consented participants. Access is progressing slowly due to the limitations of the EU data sharing agreement.

Summary of Research in Lay Language *(suggested ~ 100 words)*

With the aim to improve the diagnosis and management of COVID-19, we have developed mathematical models that describe the COVID19 disease process, from infection through to the outcome (recovery or death). The models have been built using the knowledge volunteered by a large number of medical experts from various fields, and will now be developed further using data from known and suspected cases of COVID-19. Patient data will be used to test, strengthen and confirm our individual patient predictions of COVID-19 disease progression. Ultimately, we anticipate the model will be used to understand COVID19 disease and guide individual patient management.

Summary of Research Objectives and Scientific Value *(suggested maximum 400 words)*

We propose to use individual patient data to parameterise and validate the COVID-19 causal Bayesian Networks (BN) we have developed with subject domain experts. If validated, the models will 1) aid our understanding of the pathophysiology of COVID19, thereby guide the focus of future research efforts, and 2) allow real time decision support for decision making at both the individual (clinical) level, and at the population (public health) level.

COVID-Intelligence has harnessed expert knowledge elicited from a range of clinical domain experts to develop BN models of COVID-19 infection, diagnosis, disease pathophysiology and clinical progression. Our expectation is that validated BN models will be ultimately used to provide real time decision support for individualised patient care, by utilising clinical, demographic and laboratory information available at the point-of-care.

The BN models will be implemented using consumer and clinician-friendly interfaces, making accessible the insights that BNs can provide about key disease processes, the likely effects of treatments, and possible outcomes, providing clear and useful guidance to clinical decision makers. While 'black box' modelling approaches reduce data to just a summary of associations, BNs are grounded in causal inference which allows relationships between relevant factors to be represented explicitly and graphically as a network of cause-effect relationships.

Patient-specific information (e.g. demographic, symptoms, laboratory and imaging results) can be added as updated 'evidence' as it comes to hand, and the BN will be used to reason

diagnostically or prognostically, by computing updated probabilities about the likelihood of infection and/or the likelihood of a range of potential outcomes. Crucially, BNs not only model statistical correlations (i.e. what happens on average in patients who receive treatment A), they can help to predict what will happen when we intervene (i.e. what would happen if that same patient instead received treatment B). The ability to apply causal reasoning when considering these counter-factual scenarios is critical for decision-making.

Research objectives:

- To use external (international) historic COVID-19 datasets, to train (parameterise) and verify (validate) our expert-elicited BN models , to be able to predict the probability of poor disease outcomes (death and/or need for ventilatory or haemodynamic support) for future patients with suspected or confirmed COVID-19 disease with limited data.
- To determine the performance characteristics of the developed COVID-19 model

Primary and Secondary Outcome Measures *(suggested maximum 200 words)*

On advice from clinical stakeholders, COVID-Intelligence models aim to predict the following outcomes of COVID-19 in a single network model:

- Hospitalisation
- Need for intensive care
- Survival to hospital discharge
- Hypoxia as measured by low blood oxygen saturation and/or need for supplemental oxygen therapy
- Multi-organ failure defined as need for dialysis, bilirubin > 5xULN, need for coagulation factor replacement.
- Thromboembolic complications
- Need for high flow oxygen therapy
- Need for invasive ventilation
- Need for haemodynamic support
- Need for steroids/anti-inflammatory therapy

Proposed Methodology and Statistical Analysis Plan *(suggested maximum 400 words)*

Methodology: In preparation for employing the ISARIC dataset, provisional BN model structures have been developed by the BN-modellers in partnership with a range of clinical specialists and disease experts.

Every variable (factor) in a (discrete) BN is associated with a table of probabilities called a conditional probability table (CPT), so-called because the probability distribution of the variable is specified *conditional* on the specific states of 'parent' variables that it is causally dependent upon (which can be thought of as subgroups for the variable). For example, such a table might specify the probability of the need for hospitalisation, *conditional* on being infected, as well as age, sex, comorbidity, etc. These CPTs can be populated (learned) directly using patient data.

Since all such tables only contain probabilities for subgroups, the CPTs (and in turn the BN as a whole) can be populated effectively with deidentified patient data. If necessary, learning can be done in a highly compartmentalised way, by splitting patient data into independent datasets (by cases, or by patient variables so long as they remain coupled to their immediate causes).

While each variable is parameterised separately, BN software packages (such as GeNIe) use efficient 'belief-updating' algorithms which allow us to calculate the probability of any variable or outcome (e.g. admission to hospital, need for mechanical ventilation) as a consequence of available information on other variables in the network. This allows the BNs to be validated in depth using the ISARIC dataset itself, other datasets (including context specific datasets), published reports, and expert knowledge. We will use standard methods for evaluating the model against data, including cross-validation using standard metrics such as *predictive accuracy* and area under *ROC curves* (AUC), as well as probability-sensitive metrics such as *log loss*. While it is complicated to compare BNs with existing models due to the novelty of both the disease and the data collected, we will compare the performance of our BN models to other modelling approaches, including standard multivariable logistic regression and naïve Bayes models, using the same datasets.

Importantly, we will also utilise our range of experts to validate the networks to ensure that the BNs are not only producing sensible and useful outputs, but also that the inferences produced by the BNs are explained well by the causal structure of the BNs. This type of validation will occur with experts in person; in addition, the BN structure will be compared with causal structures and relationships in the published literature (both implied and explicit).

Analysis: The overall performance characteristics of the COVID-19 model in predicting outcomes will be reported in the form of calibration plots, receiver-operating characteristic (ROC) curves and area under the curve (AUC) and log losses for each predicted outcome. We will also compare the model outputs with those of a logistic regression model with all BN variables included as covariates, an equivalent Naïve Bayes (NB) model, as well as Tree Augmented NB models (TAN models) that reincorporate dependencies between the covariates.

A sensitivity analysis will be performed to assess model performance (in the form of the measures above), assessing the model sensitivity to parameter values, state space choices and parameter selection. This will include exploration of different discretisation thresholds for categorising continuous variables, examining the merits of incorporating newly collected variables and varying the probabilistic relationships between parent and child nodes. We will perform one-way and two-way parameter sensitivity analyses (examining the effect of varying one and two individual parameters at a time), as well as a variance-based sensitivity analysis (VBSA) that allows the distribution over many input parameters to be investigated simultaneously.

Ethics *(suggested maximum 300 words)*
Provide details of any ethical considerations relating to the research proposal.
Additionally, list any approvals required by your institution to undertake this work, list reference numbers of any approved proposals, or explain why no approvals are required.

This project has been submitted to Monash University Human Research Ethics Committee for review (Project ID: 26942) under the negligible risk pathway and will proceed only in accordance with the approved study protocol.

Data used for this project will be de-identified. All data acquired for this project will be stored in secure and confidential conditions in accordance with the Monash University Research Data Management Policy, the approved study protocol and the Data Transfer Agreement.

Publication and Dissemination Plan *(suggested maximum 300 words)*
Provide details of plans for authorship/acknowledgement of data contributors.
Provide details of timelines for publication and dissemination of research findings.

Results of this project will be disseminated via peer-reviewed publications. Findings will also be submitted as abstracts to relevant local, national and international conferences as appropriate. Models and research findings will be made available as soon finalised (and within 12 months of commencement), through pre-print publication and/or as part of the Open Science Framework, a free, open source web application that enables easy and rapid sharing of research findings. There is potential for validated models and associated decision support tools to be incorporated into jurisdictional, national and international management guidelines.

All publications arising will adhere to the Australian Government's National Health and Medical Research Council guidelines for the dissemination of research findings. The project lead will coordinate dissemination of data from this project. All publications (e.g., manuscripts, abstracts, oral/slide presentations, book chapters) based on this project will be provided to the project lead and each project investigator for review prior to submission. Authorship will be determined in line with the Uniform Requirements for Manuscripts Submitted to Biomedical Journals published by the International Committee of Medical Journal Editors. In summary, authorship will be limited to those who have:

- Contributed substantially to the conception and design of the project; or the acquisition, analysis or interpretation of data for the work; AND
- Drafted the work or revised it critically for important intellectual content; AND
- Provided final approval of the version to be published; AND
- Agreed to be accountable for all aspects of the work in ensuring that questions related to the accuracy or integrity of any part of the work are appropriately investigated and resolved.

Acquisition of funding or general supervision of the research group alone does not constitute authorship. The final decision on authorship of any publication will be the responsibility of the project lead.

Research Priorities Addressed (*suggested maximum 300 words)*
Provide details of how this research aligns with nationally or internationally set research priorities.

This project aims to further the understanding of the risk factors for, and pathophysiology of, COVID-19 infection to inform future research direction, and individual patient management. It will therefore address a number of knowledge gaps outlined in the WHO's Coordinated Global Research Roadmap. This includes clinical considerations such as: groups at high risk of severe disease, pathophysiology of severe disease, clinical prognosis associated with various immunological biomarkers. The models also aim to support the optimal strategies for supportive care interventions and use the available data to optimise standard of care.

We intend for our models to be openly shared and to have clinical application around the world. As these models are intended to address knowledge gaps and improve clinical management, research objectives for this project therefore align with the highest ranked priorities of The African Academy of Sciences, research and development goals for COVID-19. In particular, our research aims to address the following priorities for clinical management: determine best clinical practice strategies to improve the processes of care e.g. develop criteria for early diagnosis, when it is safe to discharge, when to use adjuvant therapies for patients and contacts, identify

prognostic factors for severe disease, define the natural history of COVID-19 infection through careful standardised and comprehensive clinical and laboratory description of cases.

Collaboration and Knowledge Sharing *(suggested maximum 300 words)*
Provide details of how this research will collaborate, support and/or share knowledge with appropriate partners. The platform is particularly interested in research that builds capacity in low-resource settings.

COVID-Intelligence models have been developed in partnership with clinical experts across Australia and the UK. These models will be implemented across Australia via the COVID19 Clinical Data Analytics Platform (CDAP) which is supported by the Australian Government, jurisdictional health departments, and a number of institutions across Australia.

The models developed for this project will be licenced under a Creative Commons Attribution-ShareAlike 4.0 International Licence and widely promoted and shared. Models will be described through rapid pre-print publication and the models made available through the Open Science Framework for public access once finalised and we will encourage their use and validation across a number of different contexts across the globe. Given the wide global scope of the ISARIC/IDDO COVID-19 dataset, we expect our models will be applicable to a number of countries across the globe including low-resource settings. Our team is willing to work with groups to facilitate the local implementation of our models.

Funding *(suggested maximum 100 words)*
Provide details of how this research will be funded/resourced. Please name the source of funding.

Funding for this project has been provided by the Digital Health Cooperative Research Centre and the Snow Medical Research Foundation.

Scientific Review *(suggested maximum 200 words)*
If the project has been scientifically reviewed, please provide details. This could be by your institution, a funder/donor or review committee.

This project was reviewed by the Digital Health CRC scientific review panel.

SECTION C: DATA

Data Variables
*Provide a list of the **data variables** required to achieve the research objectives.*
Note: Please go to www.iddo.cognitive.city to explore the interactive COVID-19 data inventory and to identify the variables, populations and data volumes required for your analysis. You can select the data variables from this inventory and copy it to this section.

BNs calculate the probability of an outcome, such as thrombosis, based on any information available across other nodes (such as patient age, time since symptom onset, white blood cell count, oxygen saturation, etc). Making use of Bayes' rule, BNs 'propagate' new information across a network while taking into account prior information (e.g. observable information available at the time). Once developed, a BN is flexible with the information it accepts and the outcomes it reports. New evidence for any variable can be used to produce updated predictions for any other variable and providing more evidence to the BN generally produces more certain predictions.

Our model explicitly describes how a SARS-CoV-2 infection may progress over time (i.e., pre-hospital, at admission and daily assessments in hospital), from exposure to respiratory tract and pathology in other organs resulting in elevated inflammatory markers and thrombosis. The variables included in our models have been intentionally based on, and intended for, application to the data variables collected in the ISARIC CRF.

For these reasons, the majority of the ISARIC variables will help parameterise the model. Draft, non-parameterised models, can be viewed within our open science site.
We have provided some more detailed information below about the application of each variable within the COVID-Intelligence models:

We wish to request **all time related variables** to inform the dynamic relationships and disease progress within the models.

Disease pathway and progression

The variables in this section have been identified by our clinical domain experts as important to recreating a faithful causal picture of COVID-19's pathways and progression. All of the following variables are indicators of disease status and appear in different submodels of our larger causal model in such a way as to inform disease progression.

Our model of disease progression is subdivided into several submodels, encompassing lung, vascular, cardiac and immune functions and associated processes (including coagulation, oxygenation, metabolic function and other organ functions). The functioning of these different systems inform the prospect that complications arise and in turn what the expected outcomes are likely to be.

In almost all cases, the removal of a variable listed here would reduce the predictive power, explanatory capability and accuracy of the model. In particular, without access to all of these variables, we would not be in a position to assess how much our models might be affected by missing data in applied settings.

Variable:	**Role and submodels:**[1]
month and year of admission relative	indicator for progression of disease within all models
dates for all subsequent events variables (e.g. day 2, day 4, day 20 from admission)	indicator for progression of disease within all models
GCS01-Total Score	indicator for neurological status (complications submodel)
RASS01-Total Score	indicator for neurological status (complications submodel)

[1] All variables in this section play the role of indi
cators that are used to inform the status of some system, organ or function that in turn inform disease progression

AVPU01-Responsiveness	indicator for neurological status (complications submodel)
SAS01-Total Score	indicator for neurological status (complications submodel)
Oxygen Saturation	indicator for lung function (respiratory submodel)
Partial Pressure Oxygen	indicator for lung function (respiratory submodel)
Fraction of Inspired Oxygen	indicator for lung function (respiratory submodel)
PP Arterial O2/Fraction Inspired O2	indicator for lung function (respiratory submodel)
Oxyhemoglobin	indicator for lung function (respiratory submodel)
Carboxyhemoglobin	indicator for lung function (respiratory submodel)
Deoxyhemoglobin	indicator for lung function (respiratory submodel)
Methemoglobin	indicator for lung function (respiratory submodel)
Carbon Dioxide	indicator for lung function, acidosis (respiratory and complications submodels)
Partial Pressure Carbon Dioxide	indicator for lung function, acidosis (respiratory and complications submodels)
pH	indicator for acidosis (complications submodel)
Cholesterol	indicator for metabolic function (complications submodel)
Base Excess	indicator for metabolic status (complications submodel)
Glucose	indicator for metabolic function (complications submodel)
Hemoglobin A1C	indicator for metabolic function (complications submodel)
Urate	indicator for metabolic function (complications submodel)
Magnesium	indicator for electrolyte balance (complications submodel)

Calcium	indicator for electrolyte balance (complications submodel)
Bicarbonate	indicator for electrolyte balance (complications submodel)
Sodium	indicator for electrolyte balance (complications submodel)
Chloride	indicator for electrolyte balance (complications submodel)
Calcium, Ionized	indicator for electrolyte balance, kidney function (complications submodel)
Calcium, Ionized pH Adjusted	indicator for electrolyte balance, kidney function (complications submodel)
Potassium	indicator for electrolyte balance; kidney function (complications submodel)
Creatinine	indicator for kidney function (complications submodel)
Urea Nitrogen	indicator for kidney function (complications submodel)
Lactic Acid	indicator for haemodynamic status and acidosis (complications submodel)
Lactate	indicator for tissue oxygenation; haemodynamic status (complications submodel)
Direct Bilirubin	indicator for liver function (complications submodel)
Gamma Glutamyl Transferase	indicator for liver function (complications submodel)
Alkaline Phosphatase	indicator for liver function (complications submodel)
Alanine Aminotransferase	indicator for liver function (complications submodel)
Aspartate Aminotransferase	indicator for liver function (complications submodel)
Bilirubin	indicator for liver function (complications submodel)
Albumin	indicator for liver function; immune response
Iron	indicator for systemic immune response (immune submodel)

Ery. Mean Corpuscular Volume	indicator for systemic immune response (immune submodel)
Monocytes/Leukocytes	indicator for systemic immune response (immune submodel)
Erythrocytes	indicator for systemic immune response (immune submodel)
Basophils	indicator for systemic immune response (immune submodel)
Eosinophils/Leukocytes	indicator for systemic immune response (immune submodel)
Protein	indicator for systemic immune response (immune submodel)
Basophils/Leukocytes	indicator for systemic immune response (immune submodel)
Monocytes	indicator for systemic immune response (immune submodel)
Ery. Mean Corpuscular Hemoglobin	indicator for systemic immune response (immune submodel)
Eosinophils	indicator for systemic immune response (immune submodel)
Erythrocyte Sedimentation Rate	indicator for systemic immune response (immune submodel)
Neutrophils/Leukocytes	indicator for systemic immune response (immune submodel)
Lymphocytes/Leukocytes	indicator for systemic immune response (immune submodel)
Interleukin 6	indicator for systemic immune response (immune submodel)
Lymphocytes	indicator for systemic immune response (immune submodel)
Hemoglobin	indicator for systemic immune response (immune submodel)
Erythrocytes Distribution Width	indicator for systemic immune response (immune submodel)

C Reactive Protein	indicator for systemic immune response (immune submodel)
Leukocytes	indicator for systemic immune response (immune submodel)
Neutrophils	indicator for systemic immune response (immune submodel)
Procalcitonin	indicator for systemic immune response (immune submodel)
Ferritin	indicator for systemic immune response (immune submodel)
Hematocrit	indicator for systemic immune response and coagulation (immune, vascular and complications submodels)
Platelet Hematocrit	indicator for systemic immune response and coagulation (immune, vascular and complications submodels)
Platelets	indicator for systemic immune response and coagulation (immune, vascular and complications submodels)
Mean Platelet Volume	indicator for systemic immune response and coagulation (immune, vascular and complications submodels)
D-Dimer	indicator for coagulation (vascular and complications submodels)
Prothrombin Intl. Normalized Ratio	indicator for coagulation (vascular and complications submodels)
Activated Partial Thromboplastin Time	indicator for coagulation (vascular and complications submodel)
Activated PTT/Standard	indicator for coagulation (vascular and complications submodels)
Prothrombin Time	indicator for coagulation (vascular and complications submodel)
Thrombin Time	indicator for coagulation (vascular and complications submodel)
Fibrinogen, Functional	indicator for coagulation and vascular function (vascular and complications submodel)
Prothrombin Time Actual/Control	indicator for coagulation (vascular and complications submodel)

Fibrinogen	indicator for coagulation (vascular and complications submodel)
Creatine Kinase	indicator for myopathy; cardiac injury (cardiac submodel)
Troponin	indicator for cardiac injury (cardiac submodel)
Troponin I	indicator for cardiac function (cardiac submodel)
Troponin T	indicator for cardiac function (cardiac submodel)
Lactate Dehydrogenase	indicator for tissue damage (complications submodel)
Amylase	indicator for pancreatic injury; immune response (complications and immune submodels)

Signs and Symptoms

The following signs are also critical to assessing the progression of the disease, in much the same way as the indicators above.

Variable:	**Role and submodels:**
Signs and Symptoms	indicator for potentially many different aspects of the model
Respiratory Rate	indicator for lung function (respiratory submodel)
Mean Arterial Pressure	indicator for cardiac function (cardiac submodel)
Diastolic Blood Pressure	indicator for cardiac function (cardiac submodel)
Systolic Blood Pressure	indicator for cardiac function (cardiac submodel)
Heart Rate	indicator for cardiac function (cardiac submodel)
Pulse Rate	indicator for cardiac function (cardiac submodel)
Temperature	indicator for systemic immune response (immune submodel)

Response to management

To capture the natural history of the disease, it is critically important to know which treatments have been given, otherwise a highly distorted causal picture of the disease will be produced, rendering the models less clinically useful and potentially misleading. In addition, the individual effects (or non-effects) of the treatments need to be accounted for in the models if we, or others, hope to use the models to provide clinical decision support.

Variable:	**Role and submodels:**
HYDROCORTISONE	effect on progression and clinical outcomes (immune submodel)
HYDROXYCHLOROQUINE	effect on progression and clinical outcomes (immune submodel)
PREDNISONE	effect on progression and clinical outcomes (immune submodel)
DEXAMETHASONE	effect on progression and clinical outcomes (immune submodel)
CORTICOSTEROIDS	effect on progression and clinical outcomes (immune submodel)
PREDNISOLONE	effect on progression and clinical outcomes (immune submodel)
ORAL STEROIDS	effect on progression and clinical outcomes (immune submodel)
METHYLPREDNISOLONE	effect on progression and clinical outcomes (immune submodel)
NON-STEROIDAL ANTI-INFLAMMATORY AGENTS (NSAIDS)	effect on progression and clinical outcomes (immune submodel)
NON-STEROIDAL ANTI-INFLAMMATORY (NSAIDS)	effect on progression and clinical outcomes (immune submodel)
NAPROXEN	effect on progression and clinical outcomes (immune submodel)
COX-2 SELECTIVE INHIBITOR	effect on progression and clinical outcomes (immune submodel)
IBUPROFEN	effect on progression and clinical outcomes (immune submodel)

ACETYLSALICYLIC ACID	effect on progression and clinical outcomes (immune submodel)
INDOMETACIN	effect on progression and clinical outcomes (immune submodel)
ANTIBIOTIC AGENTS	effect on progression and clinical outcomes (respiratory and complications submodel)
AMPICILLIN	effect on progression and clinical outcomes (respiratory and complications submodel)
AMOXICILLIN	effect on progression and clinical outcomes (respiratory, and complications submodel)
SULFAMETHOXAZOLE AND TRIMETHOPRIM	effect on progression and clinical outcomes (respiratory, and complications submodel)
CEPHALOSPORINS - 1ST GENERATION	effect on progression and clinical outcomes (respiratory, and complications submodel)
CEPHALOSPORINS - 2ND GENERATION	effect on progression and clinical outcomes (respiratory, and complications submodel)
CEPHALOSPORINS - 3RD GENERATION	effect on progression and clinical outcomes (respiratory, and complications submodel)
CEPHALOSPORINS - 4TH GENERATION	effect on progression and clinical outcomes (respiratory and complications submodel)
CEPHALOSPORINS - 5TH GENERATION	effect on progression and clinical outcomes (respiratory and complications submodel)
CEFTRIAXONE	effect on progression and clinical outcomes (respiratory, and complications submodel)
ISONIAZID	effect on progression and clinical outcomes (respiratory, and complications submodel)
ANTIVIRAL AGENTS	effect on progression and clinical outcomes (respiratory, and complications submodel)
ACICLOVIR / VALACICLOVIR	effect on progression and clinical outcomes (respiratory, and complications submodel)

LOPINAVIR AND ROTINAVIR	effect on progression and clinical outcomes (respiratory, and complications submodel)
OSELTAMIVIR	effect on progression and clinical outcomes (respiratory, and complications submodel)
ANTIFUNGAL AGENTS	effect on progression and clinical outcomes (respiratory, and complications submodel)
PRIMAQUINE	effect on progression and clinical outcomes (respiratory, and complications submodel)
ANTIMALARIAL AGENTS	effect on progression and clinical outcomes (respiratory, and complications submodel)
CHLOROQUINE	effect on progression and clinical outcomes (respiratory, and complications submodel)
NON-INVASIVE VENTILATION	effect on progression and clinical outcomes (respiratory, and complications submodel)
INVASIVE VENTILATION	effect on progression and clinical outcomes (respiratory, and complications submodel)
HIGH-FLOW NASAL CANNULA OXYGEN THERAPY	effect on progression and clinical outcomes (respiratory, and complications submodel)
PRONE POSITIONING	effect on progression and clinical outcomes (respiratory, and complications submodel)
NEUROMUSCULAR BLOCKING AGENTS	effect on progression and clinical outcomes (respiratory, and complications submodel)
OXYGEN THERAPY	effect on progression and clinical outcomes (respiratory, and complications submodel)
EXTRA CORPOREAL LIFE SUPPORT (ECLS/ECMO)	effect on progression and clinical outcomes (respiratory, and complications submodel)
INHALED NITRIC OXIDE	effect on progression and clinical outcomes (respiratory, and complications submodel)
TRACHEOSTOMY	effect on progression and clinical outcomes (respiratory, and complications submodel)
CASPOFUNGIN	effect on progression and clinical outcomes (respiratory, and complications submodel)
REMDESIVIR	effect on progression and clinical outcomes (respiratory, and complications submodel)

IVERMECTIN	effect on progression and clinical outcomes (respiratory, immune and complications submodel)
ANGIOTENSIN CONVERTING ENZYME INHIBITORS (ACE INHIBITORS)	effect on progression and clinical outcomes (respiratory, immune and complications submodel)
ANGIOTENSIN II RECEPTOR BLOCKERS (ARBS)	effect on progression and clinical outcomes (respiratory, immune and complications submodel)
AZITHROMYCIN	effect on progression and clinical outcomes (respiratory, immune and complications submodel)
VASOPRESSOR/INOTROPIC SUPPORT	effect on progression and clinical outcomes (complications submodel)
RENAL REPLACEMENT THERAPY (RRT) OR DIALYSIS	effect on progression and clinical outcomes (complications submodel)
HEPARIN	effect on progression and clinical outcomes (complications submodel)

Risk factors

Key risk factors, such as age, sex and comorbidities, are required both as a critical way of controlling for confounding, as well as to identify groups at increased risk of severe disease.

Variable:	**Role and submodels:**
Age	reported risk factor for severity of outcome
Sex	reported risk factor for severity of outcome
Country	inform prevalence and risk of infection (for diagnostic purposes); inform prior distributions over other variables (e.g. BMI, age, sex)
Comorbidities	critical risk factor for severity of outcome (respiratory, cardiac and complications submodels)
Intensive Care Unit	indicator of disease severity
Weight	reported risk factor for severity of outcome
Height	as part of BMI, reported risk factor for severity of outcome

Body Mass Index	assess nutritional status as a risk factor for severe disease (weight, height and BMI are of course overlapping)
Mid-Upper Arm Circumference	assess nutritional status as a risk factor for severe disease
Estimated Gestational Age	assess prematurity as a risk factor for severe disease
Pregnant Indicator	assess pregnancy status as a risk factor for severe disease

Infections/co-infections

Co-infections can function in a similar way to comorbidities; in addition, there is the potential for them to interact with the disease itself, and hence are required to properly categorise patient disease progression and severity.

Variable:	**Role and submodels:**
Severe Acute Resp Syndrome Coronavirus 2	viral load would inform the immune submodel, confirmation would inform the model overall
Coronaviridae	co-infection status (immune submodel)
Bacteria	co-infection status (immune submodel)
Haemophilus influenzae	co-infection status (immune submodel)
Microbial Organism Identification	co-infection status (immune submodel)
Adenoviridae	co-infection status (immune submodel)
Respiratory Syncytial Virus	co-infection status (immune submodel)
Other Respiratory Diagnosis Indicator	co-infection status (immune submodel)
Adenovirus	co-infection status (immune submodel)
Coronavirus	co-infection status (immune submodel)
Human Immunodeficiency Virus	co-infection status (immune submodel)