# Verisk COVID-19 Projection Tool

## Methods and Assumptions

April 14, 2020

AIR

## Copyright

## Trademarks

## Contact Information

If you have any questions regarding this document, contact:

AIR Worldwide Corporation
131 Dartmouth Street
Boston, MA 02116-5134
USA

Tel:	(617) 267-6645
Fax:	(617) 267-8284

# Table of Contents

# Introduction

The information the Verisk COVID-19 Projection Tool provides is designed to help the global community in their assessment of this event. Users of this information should be aware that it is based on a model. All models are simplifications of the real world and will not inherently capture all theoretical aspects of the subject in question. Relying on this information as purely factual is inappropriate. AIR suggests you use this tool and the information it provides in concert with other data and models.

The information provided by this tool is based on: 1) the simulated results from the AIR Pandemic Model; 2) the latest reported confirmed cases and deaths, and 3) AIR's judgment regarding underreporting, healthcare infrastructure, overall assessment of the event, and other aspects of this outbreak.

Users of this tool should be aware that the assumptions that inform it are subject to change at any moment. AIR’s intent is to make as reasonable an estimated projection of cases and deaths as possible, based on the latest data and research. Therefore, if new data, research, or improvements in method suggest changes should be made to the projections, AIR will make those changes if we deem them necessary.

To develop this information, AIR estimated the actual cases and deaths for each country, given the latest information available at the time of running the model. The low, moderate, and high scenarios are based on various levels of $R_0$ (i.e., reproductive number at time = 0), effectiveness of containment measures, estimates of case fatality rate, estimates of global and local travel patterns, and other metrics. Containment measures and procedures are not uniform across countries and across U.S. states, so developing assessments of those containment measures across large geographical, political, social, and economic domains presents challenges. In addition, these metrics are not independent of one another. For example, some cultures’ general practices may mitigate or exacerbate disease transmission and potentially affect regional fatality rates of the pandemic. Therefore, even countries or regions under very similar government policies may experience differences in their transmission dynamics. This may have a knock-on effect on the mortality rates in multiple ways. Most notably, countries inundated with cases may have their healthcare system overtaxed, which may increase the mortality rate experienced by that country.

Therefore, to model these complex relationships, AIR relied on the AIR Pandemic Model’s inherent country-specific factors, the current literature on this event, and our judgment to develop a more robust and quantifiable adjustment process. In some cases, this process may result in an over- or underestimation of a country’s risk, but at the same time provides more stability in the projections; we believe this trade-off is balanced.

Through the duration of this tool's availability, AIR will incorporate new data at regular intervals and make updates to our assumptions as we learn more about the pandemic. Therefore, no assumption here should be considered permanent, as updates will likely occur. These changes may have a material impact on our model's estimates.

Finally, this document is not intended to provide an all-encompassing set of assumptions and methods, but rather an overview of the methods used. The results are propagated based on the AIR Pandemic Model, some of whose details are regarded as AIR's intellectual property.

# Range of Projections

The low, moderate, and high scenarios are not all-encompassing. It is best to think of these as projections of what the future could look like. This pandemic combined with our ability to respond is unprecedented in human history. This results in a great deal of uncertainty about what will happen in the future. Therefore, it is possible that the actual results will fall outside the bounds of the tool's modeled projections.

The scenarios listed below are intended as guideposts for the event; they do not represent an absolute best- or worst-case scenario. The results provided in the following tabs contain a range of assumptions and are based on sets of simulations that provide a midpoint, given their respective scale of severity. The $R_0$ parameter and $R_c$ parameters are defined below. In either case, we convert these to $R_e$ (Effective Reproductive Number) by modifying relative to the percentage of the population still susceptible for the given polygon, contact matrix, age, and sex.

- **Low scenario** – Designed to capture the potential trajectory of the pandemic under the condition that the mitigation efforts in each country are very effective, the containment continues to be effective, and underreporting is relatively low.
- **Moderate scenario** – Designed to capture the potential trajectory of the pandemic under the condition that the mitigation efforts have a moderate impact on the long-term rate of transmission, the containment measures don't change substantively in the future, and underreporting is average.
- **High scenario** – Designed to capture the potential trajectory of the pandemic under the condition that the mitigation efforts have a modest impact on the transmission rates over the duration of the event and underreporting is relatively large.
    - The current state of mitigation and our response to it only represents the current moment in time. Unfortunately, modeling a country's long-term response to "lockdown" is challenging. Countries may lift economic and social constraints early, individual adherence to policies may wane over

time, and seasonality may affect the disease dynamics and our response effectiveness.

For all three scenarios, AIR modeled the impact of the outbreak by changing the following three parameters: $R_0$ (reproductive number under normal conditions, assuming a fully susceptible population), $R_c$ (reproductive number with containment), and underreporting (the difference between reported cases and deaths and actual cases and deaths). AIR ran hundreds of different scenarios by changing these parameters. AIR then selected the low, moderate, and high scenarios that represented both our assessment for a country to mitigate the outbreak, included real-time data to better select parameters for the given country, and fit within our assessment of a country’s mitigation efforts.

We will be updating the projections on a regular basis to stay consistent with the reporting and containment procedure updates to capture the stochastic variability that is inherent in this process. Users of this data should be cognizant of the breadth of projections AIR is providing. Given this very challenging task, users should be mindful that some projections and some individual metrics may not comport with current reported data. AIR strives to limit this, but the nature of this event is dynamic.

## Underreporting

Underreporting is a conditional term for each country and scenario, comprising the healthcare capacity of the given region, similar pathogens’ estimated historical underreporting, current literature on the outbreak and subject matter in general, and our own judgment. It is important to understand that the term underreporting has a unique meaning for each country. It is well known and reported in the epidemiological literature that the number of cases and deaths are underreported for nearly all major disease outbreaks. Estimating the amount of underreporting presents challenges; estimates vary based on myriad factors that include but are not limited to: pathogen characteristics, region, tracking and reporting on the event, and demographic profile.

AIR has many years of experience in studying disease transmission, reviewing outbreak data, and considering actual vs reported cases. In the case of the current pandemic, AIR has devoted a great deal of time to performing our own research on the topic and reviewing the most respected literature on the subject. In spite of our efforts, our assessment may still materially under- or overestimate the actual underreporting. AIR will continue to work to improve our estimates in this tool.

To understand why underreporting exists, we point you to the study by Gibbons et al, 2014. In that study, authors provide a framework for how to understand underreporting and why it may occur. Figure 1 is from this study and provides an example of why there may be 1,000 reported cases, but in reality, 3,300 cases exist.

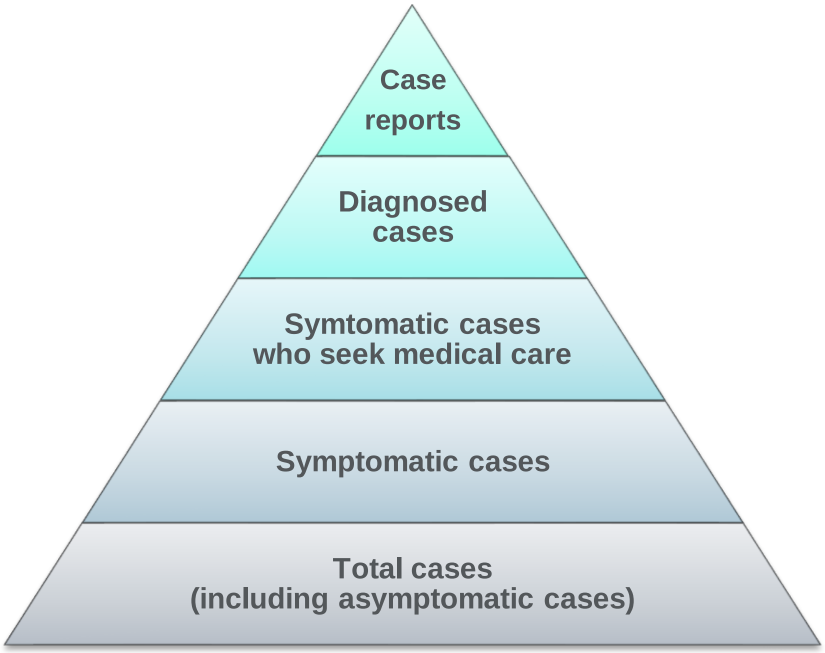

**Figure 1. Inspired by Gibbons et al., 2014, this pyramid illustrates how reported cases are only a fraction of the disease spread dynamic within a population.**

To develop our underreporting method, AIR relied on peer-reviewed/pending review literature, articles on news agencies, data from Verisk Maplecroft, and our own research. For this pathogen, AIR developed a quantifiable underreporting method that varies by country healthcare infrastructure, comparison of reported case fatality rate (CFR) to expectation, estimation of stress on the healthcare system, and sample error considerations. The underreporting factor also considers the fact that more than 30% of population is asymptomatic.

## Underreporting Metrics

Local healthcare infrastructure can have a significant impact on underreporting metrics. AIR uses multipliers of cases and deaths based on assumed reporting reliability by country.

Multiplier ranges of underreporting used are:

- Cases – Directly implemented
    - Great healthcare infrastructure: ~3 to 12
    - Moderate healthcare infrastructure: ~8 to 20
    - Poor healthcare infrastructure: >15
- Deaths – Result of many assumptions
    - Great healthcare infrastructure: ~1 to 10

  - Moderate healthcare infrastructure: ~3 to 15
  - Poor healthcare infrastructure: >8

Underreporting is relative to when data was pulled. As we project, underreporting ranges may change and therefore could fall outside the bounds noted above.

For deaths, underreporting can vary more, as death rates are based on many variable sets of data, and the final factor applied is a composite of many other factors. This results in a less strict relationship between healthcare infrastructure and underreporting factor.

# Severity

## Severity Definitions

To provide more information to the user, AIR has broken down the cases into various levels of morbidity and mortality.

- Mild = Asymptomatic to low-grade infection. Unlikely to seek medical treatment.
- Moderate = "Flu-like symptoms." May experience one or more of the following: fever, cough, shortness of breath, fatigue, muscle and joint pain, chills, headache, and other commonly associated respiratory symptoms. Moderate-to-high chance the person seeks medical treatment
- Severe = Person highly likely to seek medical treatment and would be expected to be admitted to a hospital for supportive care under normal circumstances
- Death = Fatality

## Severity Metrics

Morbidity and mortality estimates are based on estimating the number of cases that fall into various buckets. The list below provides the reasonable ranges that AIR has relied on to inform our judgment of the probability that a case belongs to a given stratification, or bucket:

- Mild = ~60 to 80% of total cases
- Moderate = ~15 to 25% of total cases
- Severe = ~5 to 15% of total cases
- Deaths = ~0.5 to 5% of total cases

# AIR Pandemic Model

All the information used in this analysis is derived from the AIR Pandemic Model, a stochastic modeling framework that can simulate the impact of a pandemic by age and sex for more than 10,000 tessels (similar to municipality). For coronavirus pathogens, the AIR Pandemic Model estimates the cases on a daily basis, using an SEIR (Susceptible-Exposed-Infectious-Removed) modeling framework. This framework allows us to more accurately model the disease dynamics, including but not limited to: travel patterns and restrictions; changes to mitigation efforts; introduction of pharmaceutical or nonpharmaceutical interventions; and impact on transmission due to changes in susceptible population. Specific to COVID-19 projection, AIR updates the estimated parameters frequently to develop near real-time forecasting.

The major variables (listed and explained below) influencing the severity of the event are: susceptible, exposed, infectious, and removed people by region; $R_0$; $T_c$Days; $R_c$; CFR; Country Specifics; Travel Patterns; Travel Restrictions; contact matrix; and city demographic data.

- $R_0$ = Reproductive number at time 0 assuming a fully susceptible population – Specific for each country
- $R_c$ = Reproductive number after containment – Specific for each country
- $T_c$Days = Days until international, national, and local communities respond in a substantive way to mitigate the spread (Assumed to have already started for each country for this pandemic)
- CFR = Case Fatality Rate – Specific to each country, based on the healthcare infrastructure of a given country
- Country-Specific Factors = GDP per capita, doctors per capita, hospital beds per capita, and other related healthcare metrics
- Travel Patterns = National and international travel between tessels (similar to municipalities).

This model is designed to allow a pathogen to spread throughout the world given the aforementioned variables.

The AIR model comprises many components, including but not limited to the following:

- The most up-to-date research on the modeled pathogen's characteristics
    - Virulence – the ability of a virus to produce disease and cause fatalities
    - Transmissibility – the effectiveness of a pathogen's ability to spread
    - Available countermeasures – the ability for humans to disrupt the spread of a pathogen

    - Spatial ignition frequency
    - Relationships between parameters
- A meta-population Susceptible-Exposed-Infectious-Removed ("SEIR") epidemiologic model is used to simulate local and global disease transmission from a simulated event, producing a spatio-temporal distribution of the disease incidence broken down per day by age and sex on a global model domain (also termed the “pandemic footprint”), for each modeled disease
- A short-range and long-range travel model that addresses the local, regional, and global spread of a given pathogen, as well as mitigation measures (such as vaccines, travel restrictions, containment procedures) that are modeled to take into account the means by which pathogens can spread and/or their spread can be stopped
- A morbidity and mortality model that runs concurrently with the calculation of the pandemic footprint and employs illness severity functions that estimate the probability of physician visit, hospitalization, and death, by age, sex, and location

The AIR Pandemic Model’s performance has been validated against illness, mortality, and financial loss data from several historical events and published research. The model is built to meet the wide spectrum of pandemic risk management needs of the insurance and reinsurance industry, government agencies, and other interested parties for each of the more than 200 countries included in it.

# Major Data Sources

The AIR model was constructed based on data and theories from more than 175 different sources, including medical, epidemiological, and scientific journal articles, international and national organizations, and historical firsthand accounts of given outbreaks. Table 2 contains a partial list of major data sources.

**Table 1. Major Data Sources**

| Source | Description |
|---|---|
| OAG Aviation Worldwide | Passenger origin and destination statistics |
| UN Population Division | Trends in international migrants by destination and origin |
| U.S. Census | Daily commuter rate statistics and demographic information for the United States |
| UN Food and Agriculture Organization | Global gridded density of domestic poultry and pig populations (~5-km resolution) |
| LandScan | Global gridded human population (~1-km resolution) |

| | |
|---|---|
| World Bank | Geographic distribution of age/sex cohorts and gross domestic product (GDP) |
| The World Health Organization | Geographic distribution of hospital beds and physicians per capita |
| Global Register of Migration Species (GROMS) | Migratory waterfowl population data |
| Global Cover | Land use/Land cover data |
| UN FAO Emergency Prevention System for Animal Health | Reported cases of influenza in wild birds, domestic birds, and swine |
| Scientific Literature | Epidemiological and biological research on: the system dynamics, disease spread, governing parameters (e.g., transmissibility, virulence, etc.), historical impact, mitigation impact, temporal data sets, etc., of the relevant pathogens |

# Active Containment

Both pharmaceutical and non-pharmaceutical mitigation measures are applied during each time step of the AIR SEIR meta-population model.

## Pharmaceutical Interventions

Pharmaceutical interventions vary by disease and include the administration of antivirals and vaccines. The efficacy and the location-specific availability of these pharmaceuticals are incorporated into the AIR Pandemic Model. Availability is directly related to factors such as GDP per capita and proximity to pharmaceutical manufacturing facilities, as well as the time elapsed since event ignition. For example, production and distribution considerations will likely result in a six-month delay in influenza vaccine availability after event ignition, while antivirals are assumed to be available immediately. In addition, the efficacy of these pharmaceutical mitigation measures is linked to the age of the infected persons receiving them, with reduced efficacy being more common among elderly patients (Osterholm et al. 2012).

## Non-Pharmaceutical Interventions

Non-pharmaceutical interventions, including travel restrictions, contact tracing, isolation, and quarantining are also explicitly incorporated into the model. The modeled impact of such measures varies by such socioeconomic factors as GDP per capita, doctors per capita, and hospital beds per capita, as suggested by historical data analysis.

The AIR Pandemic Model includes active containment measures, which can significantly affect the progression of a pandemic by reducing the probability of disease spread. To account for containment measures, AIR uses two variables $T_c$ and $R_c$. $T_c$ denotes time until active containment. This variable is designed to model the time, in days, it takes the international community to mount an effective response to the outbreak. The $R_c$ variable models the effectiveness of the response. It is applied by replacing the $R_0$ variable in the model.

## Epidemiological Model and Event Footprint

In the AIR Pandemic Model, the local intensity of each simulated pandemic is defined in terms of a pandemic footprint for a region. Outbreak ignition is modeled on a gridded domain across the entire globe at a tessel resolution. Tessels are sub-national, population-based grid cells similar to municipalities. The model utilizes population data that contains age and sex distributions for each of these granular tessels. Each tessel is defined by a polygon. There are more than 10,000 tessels around the world, covering the entire global population. At the start of a simulated outbreak event, start parameters are drawn from distributions fitted to available historical records and scientific research. Starting parameters include $R_0$, CFR, age virulence profile, start day, length of outbreak, days until active containment, effectiveness of active containment, ignition location, and others. For every simulated event, morbidity and mortality counts are produced for each tessel on a daily basis. The morbidity and mortality counts are further broken down by age, sex, and the SEIR compartments. The resulting event footprint is a comprehensive daily representation of the simulated event. The sections below describe in more detail the different components of the epidemiological model which is used to produce a given event footprint.

**Figure 2. Global map of air travel density. (Source: AIR Worldwide)**

## Epidemiological Model: Event Ignition

For each emergent outbreak, epidemic, or pandemic in the AIR Pandemic Model’s stochastic catalog, a given geographic location is seeded with exposed individuals. After an incubation period, a portion of these exposed individuals become infectious. In SEIR modeling, this portion is determined using a statistical distribution that captures the incubation period of the modeled pathogen.

## Epidemiological Model: The Susceptible-Exposed-Infectious-Removed Compartments

The AIR Model simulates local and global disease transmission for each emergent pandemic in the AIR Pandemic Model’s stochastic catalog using a daily meta-population SEIR epidemiologic model as shown in Figure 3. This process yields, for each pandemic in the AIR Pandemic Model’s stochastic catalog, a complete pandemic footprint. At each of the aforementioned tessels the AIR model uses the generalized SEIR epidemiological model. In a generalized SEIR epidemiologic model, the population of each tessel is divided into four disease compartments to facilitate estimates of disease spread through the population (Anderson and May 1991, Keeling and Rohani 2008). This process allows the model to estimate how persons move from being susceptible, to exposed, to infectious, and finally to removed.

- Susceptible: Able to contract the pathogen
- Exposed: Has been in direct contact with the infectious agent
- Infectious: Someone who is currently symptomatic or asymptomatic and can infect susceptible people in the population
- Removed: Indicates someone who has recovered, died, or been successfully vaccinated

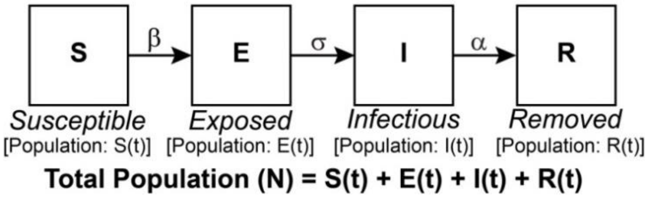

**Figure 3. Outline of a generalized SEIR model**

The SEIR model contains eight disease state compartments: one susceptible state; one exposed state; three infectious states (asymptomatic, symptomatic but can travel, and symptomatic but cannot travel); and three removed states (vaccinated, recovered, and dead), as shown in Figure 4. These eight compartments are each further divided into seven age categories and two sex categories, yielding a total of 112 categories in the complete SEIR model.

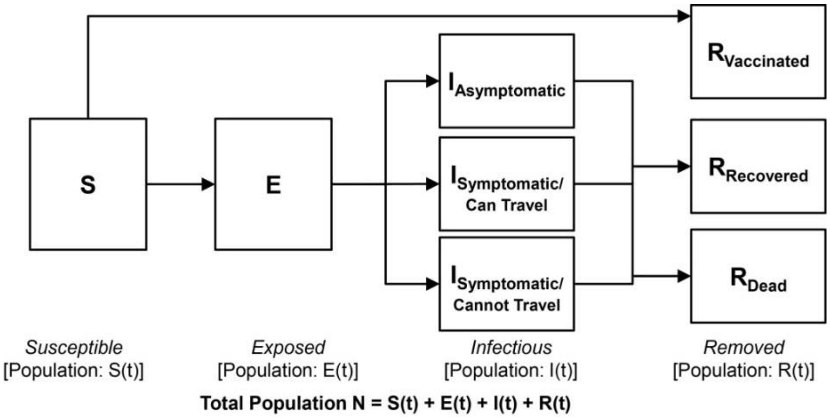

**Figure 4. Flowchart showing the eight disease state compartments of the AIR SEIR Model for airborne diseases**

The AIR model estimates the number of people within each of these states at daily time steps. At any point in time, each person exists in one of these disease states, therefore, the sum of all 112 categories is constant and equal to the number of people in each of the states at the time, t = 0. In the AIR SEIR model, the probability of disease transmission from infectious to susceptible individuals is calculated using the daily infectious contact rate. The daily infectious contact rate is obtained by dividing the $R_0$ value for the event by the mean illness duration. An age-based contact matrix is also used to account for disease spread due to mixing of individuals of different age groups.

The SEIR model interacts with the outbreak spread model to enable AIR to explicitly model how the pathogen can spread from one tessel to another. As more people become infected in a given region, yet are able to travel, the likelihood of disease spreading to other geographical regions increases. To properly account for this spread, the AIR model explicitly captures the travel by various means: international, regional, and local patterns. Each of these travel patterns are based on international data sets and used to understand the average flow between two points on the globe, either close in proximity or far apart. Areas with excessive travel or that are close in proximity to the outbreak will likely be the first to experience the spread of the pathogen and explicitly captured in the AIR model travel patterns.

Infections often increase during the winter months. To properly account for seasonality, the probability of disease transmission is adjusted using a seasonality modifier. The modifier is based on distance from the equator (i.e., latitude) and the day of the year relative to North/South cardinal directions. The model uses a modified sine function to scale the contact parameter, thus the farther north or south an individual is located, the more of an effect this will have. Increasing the contact parameters effectively increases the number of people who become exposed. In tropical regions, the probability of disease transmission is not adjusted for seasonality as climatic conditions stay fairly constant throughout the year.

### Epidemiological Model: Disease Spread

AIR breaks down the main drivers of disease events into epidemiological spread (i.e., disease transmission), meaning the fine-scaled progression of disease among a particular population (i.e., the SEIR model), and geographic spread, which refers to the movement of humans around the globe. Spread can occur from long-range travel, between neighboring countries, or within a country. AIR models each of these features using three distinct methodologies. Each of these methods allows the AIR model to capture the varying means by which pathogens can spread.

# Characteristics of Modeled Pathogens

Characteristics of infectious disease interact to determine how they spread throughout the human population. These characteristics allow the disease to spread more readily across the globe or cause greater impact in particular populations. Others influence the outcome severity in its human host but may limit the final size of the disease outbreaks. At AIR, research estimates the range of potential values of quantifiable variables by pathogen type in order to grasp the full scope of future possible disease events.

## AIR Model and Covered Perils

AIR models many infectious disease pathogens. The AIR Pandemic Model is designed to adjust with the disease in question and consider the disease-specific factors that affect the simulated size and severity of the event.

## Additional Information

Information can be provided for all modeled pathogens upon request. You may contact us for additional details.

# About AIR Worldwide

AIR Worldwide (AIR) provides risk modeling solutions that make individuals, businesses, and society more resilient to extreme events. In 1987, AIR Worldwide founded the catastrophe modeling industry and today models the risk from natural catastrophes, terrorism, pandemics, casualty catastrophes, and cyber incidents. Insurance, reinsurance, financial, corporate, and government clients rely on AIR's advanced science, software, and consulting services for catastrophe risk management, insurance-linked securities, longevity modeling, site-specific engineering analyses, and agricultural risk management. AIR Worldwide, a Verisk (Nasdaq:VRSK) business, is headquartered in Boston, with additional offices in North America, Europe, and Asia. For more information, please visit www.air-worldwide.com.

AIR

A Verisk Business