Assessing Global Pandemic Risks from Emerging Infectious Diseases and High Containment Laboratory Leaks: A Country Level Spatial Network SIR Model Analysis

Ross J. Tieman

ross.tieman@gmail.com

a Alliance to Feed the Earth in Disasters (ALLFED), Fenner School of Society and Environment, Australian National University https://orcid.org/0000-0003-4639-1628

Pedro A. Oliboni

Global Priorities Institute, University of Oxford

Simeon Campos

Research Article

Keywords: global catastrophic risk, biosecurity, SIR model, travel networks, emerging infectious disease

Posted Date: May 3rd, 2024

DOI: https://doi.org/10.21203/rs.3.rs-4343122/v1

Additional Declarations: The authors declare no competing interests.

Assessing Global Pandemic Risks from Emerging Infectious Diseases and High Containment Laboratory Leaks: A Country Level Spatial Network SIR Model Analysis

Ross J. Tieman [a,b *], *Pedro A. Oliboni* [c,1], *Simeon Campos* [1]

[a] *Alliance to Feed the Earth in Disasters (ALLFED)*
[b] Fenner School of Society and Environment, Australian National University
[c] Global Priorities Institute, University of Oxford
* Corresponding author contact: ross@allfed.info

Abstract

Future pandemics could arise from several sources, notably, Emerging Infectious Diseases (EID); and lab leaks from High Containment Biological Laboratories (HCBL). Recent advances in infectious disease, information technology and biotechnology provide building blocks to reduce pandemic risk if deployed intelligently. However, the global nature of infectious diseases, distribution of HCBLs, and increasing complexity of transmission dynamics due to travel networks, make it difficult to determine how to best deploy mitigation efforts. Increasing understanding of the risk landscape posed by EID and HCBL lab leaks could improve risk reduction efforts.

The presented paper develops a country level spatial network Susceptible Infected Removed (SIR) model based on global travel network data and relative risk measures of potential origin sources, EID and lab leaks from Biological Safety Level 3+ and 4 labs, to explore expected infections over the first 30 days of a pandemic. Model outputs indicate that for EID and lab leaks India, the US and China are most impacted at day 30. For EID, expected infections shift from high EID origin potential countries at day 10 to the US, India and China, while for lab leaks the US and India start with high lab leak potential. With respect to model uncertainties and limitations, results indicate several large wealthy countries are influential to pandemic risk from both EID and lab leaks indicating high leverage points for mitigation efforts.

1. Introduction

This paper explores the role of global travel networks in propagating risk from two origin sources for Pandemic Potential Pathogens (PPP): Emerging Infectious Disease (EID), and Lab leaks, escape of pathogens from Biological Safety Level 3+ and 4 laboratories, through the development of a spatial network Susceptible Infectious Removed (SIR) model.

Novel infectious disease has historically been a major cause of human harm. Salient examples include the black death which is argued to have caused 200M deaths over three major outbreaks (Perry & Fetherston, 1997) and the Spanish flu killing between 17.4 million (Spreeuwenberg et al., 2018) and 50 million (Johnson & Mueller, 2002) people or between 0.95% and 2.7% of the population. Pandemics, that is infectious disease outbreaks

[1] Authors made equivalent significant contributions

that spread globally, are considered a potential source of existential risk (Millett & Snyder-Beattie, 2017). COVID-19, a comparatively mild pathogen, highlights humanity's ongoing vulnerability to infectious disease and the potential damage that can be caused by pandemics. Estimates of existential risk from natural pandemics are considerable ranging from a 1 in 2000 to 1 in 20000 per century, though these estimates should be considered with skepticism due to methodological issues and challenges in estimating the probability of low frequency events (Millett & Snyder-Beattie, 2017; Sandberg & Bostrom, 2008). Estimates of existential risk from pandemics become considerably higher if engineered pathogens are considered, with order of magnitude estimates being 1 in 30 per century (Ord, 2020).

Historically, transmission of novel pathogens from animals to humans has been the major source of Emerging Infectious Disease (EID) and consequently pandemic risk. More recently human interaction with Pandemic Potential Pathogens (PPP) in laboratory settings for purposes such as diagnostic testing, vaccine development and development of biological weapons has yielded a new reservoir of PPP. Though most work with PPP is undertaken in designated Biological Safety Level (BSL) 3 and 4 laboratories, highly secure laboratory facilities specially developed to prevent high risk pathogens from being released into the environment, many records of pathogen escape, henceforth referred to as *lab leaks*, from BSL 3 and 4 facilities exist. Notable examples of lab leaks include: The 2007 UK outbreak of Foot and Mouth Disease from the Pirbright facility a BSL 4 labs (DEFRA, 2008), and Venezuelan equine encephalitis virus (VEEV) in 1995 which was identical to a lab strain from 1963 (Brault et al., 2001). Consequently BSL 3 & 4 labs are a significant source of pathogen risk, arguably of higher concern than EID due to by design selecting for high risk pathogens and being located in urban areas (Merler et al., 2013).

In addition, modern societies global air travel networks facilitate rapid dispersion of pathogens (Holmes et al., 2018), further increasing the risk posed from the emergence of new diseases or lab leaked pathogens. Even if a particular country isn't likely to give rise to a new emerging infectious disease or be the source of a lab leak, it might be strongly connected to other countries that do face such a risk.

Reducing the risk of extinction or long term impacts from pandemics should be considered a high priority (Millett & Snyder-Beattie, 2017), but which countries are most exposed to risk from future pandemics? Which countries contribute the most to creating such risks? And how given this information should one prioritize mitigation efforts? Without sufficient understanding of pandemic risk, one could go about selecting interventions and preventive measures in a way that assumes risk is uniformly distributed. However, pandemic risk is clearly not uniformly distributed. As such, a good enough estimate of how risk varies according to geographical and socio-economic factors would allow targeting interventions and preventive measures where they would help the most. These are difficult but important questions, whose answer can be used to prioritize how limited resources are used for mitigation efforts.

This paper contributes to understanding of pandemic risk by exploring the interaction of travel networks with two sources of PPP, EID and lab leaks from BSL 3+ and 4 labs. This is achieved by developing a spatial network Susceptible Infected Removed (SIR) model that

uses a relative distribution of risk of outbreak from two potential sources of PPP, EID and lab leaks, to explore how risk propagates through travel networks. Knowledge of how a hypothetical pathogen may be distributed globally provides information for prioritizing mitigation efforts for pandemics e.g. high priority locations for screening of people who present symptoms that cannot be diagnosed easily (Holmes et al., 2018). To our knowledge there has been no attempt to aggregate origin risk from EID and lab leaks with travel networks in assessing pandemic risk. Overlaying the risk distribution EID and lab leak sources can shed light on commonalities in risk propagation between different origin sources to better target intervention deployment.

The presented model is a coarse grained first pass at assessing the distribution of pandemic risk. Given the complexity of disease dynamics and the associated challenge of prioritization, it is perhaps too early to tell whether results are robust enough to be actionable. It is the authors hope that this work can provide some preliminary insights, and direct future work to reduce the uncertainties faced in understanding the highly complex phenomena involved in future risks from pandemics.

2. Methods

To investigate the role of travel networks a Susceptible Infected Removed spatial network model using countries as nodes was developed, where the network is derived from travel data. The model is run with each country as a pandemic origin source for a set of disease parameters. Outputs of the model are then scaled according to the relative risk of origin in a given country. Origin sources are EID, or lab leak. The SIR model structure, method of aggregation and EID and lab leak origin potential measures are described in more detail below.

2.1 SIR model overview

The model is an adaptation of the SIR network model over cities described by Muroya et al. 2013, chosen as inspiration due to its simplicity (Muroya et al., 2013). The model considers the world population N, partitioned into n countries labeled by j=1,2,...,n . The population within the jth state is partitioned into susceptibles (Sj), infectives (Ij), and removed (recovered and dead) (Rj) .

The equations that describe the system are as follows:

$$\frac{dS_i}{dt} = -S_i \sum_{j \neq i} m_{ij} - S_i \beta_i \frac{I_i}{N_i} + \sum_{j \neq i} m_{ji} S_j \quad (1)$$

In words, the change in Si depends on how many susceptibles migrate out, how many susceptibles get infected, and how many susceptibles from other countries migrate in.

$$\frac{dI_i}{dt} = S_i \beta_i \frac{I_i}{N_i} - I_i \gamma_i - I_i \sum_{j \neq i} m_{ij} + \sum_{j \neq i} m_{ji} I_j \quad (2)$$

In words, the change in *Ii* depends on how many susceptibles get infected, how many infectives removed[2], how many infectives migrate out and how many infectives from other countries migrate in.

$$\frac{dR_i}{dt} = \gamma_i I_i - R_i \sum_{j \neq i} m_{ij} + \sum_{j \neq i} m_{ji} R_j \quad (3)$$

In words, the change in *Ri* depends on how many infectives are removed, how many removed migrate out, and how many removed from other countries migrate in.

Notation:

m_{ij} is the rate of travel from *i* to *j* at each time step;

γ_i is the rate at which infected people in country *i* are removed from the disease pool;

β_i is the transmission rate, which indicates the rate at which S_i individuals are infected by each individual in I_i.

The model is mechanistic, allowing the dynamics of spread of infectious diseases to be investigated. Importantly, the model allows for the study of indirect connections (travel from country A to B to C) and the variation of disease characteristics without the addition of too many moving parts.

The global spatial network Susceptible Infected Removed (SIR) structure models infectious disease dynamics as a phenomenon that happens in a geo-political network with countries as nodes and travel flows as directed links. The crucial simplification this model makes is of abstracting away from the network of people interactions to a network of country interactions that are mediated by flows of people. This abstraction is important especially in regards to lab leak as it averages out influence of population density and co-location of BSL labs in urban centers (see discussion).

Network SIR models for inter-city (cities as nodes) spread of infectious diseases are common in the literature (Muroya et al., 2013; Pujari & Shekatkar, 2020). Surprisingly, we found no attempts at SIR-like network models for global disease spread which use EID and lab leaks as a source of pathogen origin. This may indicate that the level of complexity of global relations cannot be appropriately captured with a network of countries. However, there are models for forecasting disease dynamics based on networks of states. For instance, Sharma et al. 2012 develops a SEIRD model (an elaboration of the SIR model) over a network of states in India (which are larger in terms of population than many

[2] Removed refers to those removed from the pool of available hosts, removed captures recovered and assumed immune hosts or hosts who have died. This model misses more complex dynamics such as decay of immunity.

countries) (Sharma et al., 2021). The existence of such a model of a network of large states may suggest that a global network of countries could be a reasonable approximation depending on one's goals.

2.1.1 Global Travel Data

The Global Transnational Mobility dataset is used to develop the network characteristics (Recchi et al., 2020). This dataset describes how many 'trips' there have been between any two countries in the world from the years 2011 to 2016 (1 trip = 1 person traveling)[3]. These estimates of trips account for multiple modes of transport and have been put together by aggregating global statistics on tourism and air passenger traffic (Recchi et al., 2019). The dataset reports travel as trips per year for the purpose of this model travel was converted to average trips per day

2.1.2 SIR Disease Parameters

The model uses transmission and removal rate parameters (β_i and γ_i). Transmission and removal rates are also assumed the same for all countries. For simplicity a single set of disease parameters equivalent to SARS-Cov2 are used for all countries for both EID and lab leaks (see Discussion). This maintains focus on the geographic, regulatory and socio-economic factors captured in origin potential measures and role of travel networks, and avoids determining plausible distributions of disease parameters likely to arise from different origin sources, which would be an extensive piece of work. Future iterations of the model could explore distributions of disease parameters equivalent to anticipated PPP informed by historical data[4].

SARS-Cov2 parameter values used are from Amiri Mehra et al. (Table 1.) (Amiri Mehra et al., 2020). It should be noted that parameter values for a given disease are influenced by many factors not captured within the model, so although informed by empirical data the parameters are ultimately abstractions.

Table 2.1-1 SIR model parameters for SARS - Cov2 (COVID19) calculated by Amiri Mehra et al. 2020

Parameter	β **(transmission rate)** $(person.\,day)^{-1}$	γ **(removal rate)** $(day)^{-1}$
Value	1	0.223

For each country with pandemic origin potential (EID or lab leak) the model is run once with an initial outbreak of 1000 infected individuals originating in that country. Each model run consists of 30 days. A country has origin potential if it has an EID vulnerability

[3] Dataset can be explored interactively (European Commission, n.d.)

[4] In the instance of synthetic pathogens e.g. pathogens developed through gain of function or other synthetic biology processes, determining plausible parameters would be even more complex as such pathogens would not be bound by proposed virulence trade-off.

score in Moore et al 2017, or has at least one lab as reported by the Global Biolabs Project (Koblentz et al., 2021; Moore et al., 2017)

2.2 Relative risk of pandemic origin

Considering all countries *C*. Given a pandemic being initiated, the probability that the pandemic originated in a given country *c* is defined by:

$$Pr(c) = \frac{r_c}{\sum_{r \in C} r_c'} \quad (4)$$

Where r_c is the origin source potential of country *c*. Origin potential measures the relative strength of contributing factors to a pandemic being initiated in a given country.

Relative risk of pandemic origin is considered separately for EID and lab leaks.

2.2.1 Emerging Infectious Disease origin potential

Origin potential of EID is based on the Infectious Disease Vulnerability Index developed in Moore et al 2017. The Infectious Disease Vulnerability Index aims to identify countries that are most vulnerable to outbreaks of infectious diseases with potential for transnational spread to inform preemptive actions that mitigate spread and consequences of such transnational infectious disease outbreaks (Moore et al., 2017). The Vulnerability Index is derived from a variety of factors in the following seven domains: demographic, public health, economic, disease dynamics, health care, political-domestic, political-international. These factors are then weighted according to elicited expert opinions.

Vulnerability scores in Moore et al. 2017 were transformed by subtracting the score from 1. This is because the original scores are out of [0,1] where a lower score means a country is more at risk. But for the purpose of this model, a country that is more at risk should be given a higher rather than a lower weight.

In particular to Moore et al. 2017 it is important to note that migration (average annual number of migrants per 1,000 people) is part of the vulnerability score. So travel has been included in some (albeit minimal) way in their assessment.

2.3.2 Lab leak origin potential

A simple measure of lab leak origin potential from BSL 3+ and 4 labs for a given country is given by:

$$r_c = b_c \sum_{i=1}^{n_c} a_{ic} \quad (5)$$

Which can be determined from the following information:

1. The number of BSL-4 and BSL-3+ labs in each country, n_c.

2. Area of BSL-4 labs and BSL-3+ in each country, a_{ic}.
3. Global Biolabs Biorisk Management Scorecard (GBBMS), (Koblentz et al., 2021) score for each country, b_c.

Data for 1-3 are provided by the Global Biolabs Project (GBP) which represents the most developed dataset on BSL 3+ and 4 locations[5,6] (Koblentz et al., 2021, 2023). Missing data required for 1 -3 were inferred by the following actions:

- BSL 4 labs entries without lab area data were assigned the geometric mean of available BSL 4 lab area data[7] . BSL3+ labs entries which do not include area were assigned this value as well.
- Countries with a lab but no GBBMS score were assigned the arithmetic mean of all GBBMS scores.

BSL3+ numbers provided by GBP roughly match BSL 3 labs reported in *Mapping Biosafety Level-3 Laboratories by Publications* (Schuerger et al., 2022) which reports 57 BSL 3 locations globally, while GBP reports 55 active BSL 3+ labs. The closeness of the two numbers is somewhat suspicious given that BSL3+ is a higher classification of 3, intuitively one would expect a significantly larger number of reported BSL 3 labs. The distribution of BSL 3 labs reported by Schuerger et al. and BSL 3+ labs by Klobentz et al. do not match at a continental level. Discrepancies between numbers might be explained by differences in method, Schuerger et al. 's estimate uses Pubmed publications and would miss non-publishing labs such as diagnostic labs. The lack of standardization of BSL level characterization[8] may also impact lab numbers and characterization.

The GBBMS score is incorporated into the lab leak potential measure to incorporate how comparatively risky labs in a given country are. The GBBMS is out of 48 biosafety management indicators derived from standards and best practices endorsed by WHO, ISO 35001, NTI and other organizations. Indicators are scored positive if relevant statutory legislation (regulations, standards, and policies) is present in a country. The metric does not capture compliance or enforcement, decreasing its robustness.

The developed lab leak measure assumes risk of lab escape is proportional to lab area. This is based on the intuition that greater lab area implies a greater number of lab experiments and required supporting systems leading to greater potential for human error (Wurtz et al., 2016), mechanical failure (DEFRA, 2008), or other failures resulting in pathogen escape. For this version of the model a linear relation between lab area and risk

[5] BSL 3+ or BSL 3 advanced labs are BSL 3 designated labs that take on extra safety precautions to carry out research on riskier pathogens that do not require BSL 4 designation e.g. highly pathogenic influenza (Koblentz et al., 2023)

[6] Data can also be interacted with through the Global Biolabs interactive map (*Global Biolabs*, n.d.)

[7] Geometric mean was used over the median as lab area values indicated a lognormal distribution and sample size was relatively small n=51.

[8] The World Health Organization publishes guidelines and biosafety best practices concerning BSL 3 and 4 research; however, no globally body to regulate implementation exists. As such implementation of proper biosafety, is up to countries and even researchers further obscuring BSL 3 research 4 safety (*Laboratory Biosafety Manual, 4th Edition*, 2020).

is assumed[9]. There is uncertainty in this assumption, for instance an argument based on economies of scale would imply that bigger labs can afford better infrastructure thus reducing risk contribution per unit area. A more in depth investigation of the relationship of lab leaks and lab size are deferred to future work.

3. Results

Model results are aggregated to obtain an "average" number of cumulative infections[10], henceforth abbreviated to ACIs, for each pandemic initiation source (EID or lab leak). Each model run provides the number of susceptible, infected, and removed in each country in the world for 10, 20, and 30 days after the initial outbreak. Results of the SIR model are averaged independently for EID and lab leak risk as the relative risk scores (see 2.2) use different underlying measures that are not directly comparable. To average model results the relative risk of pandemic origin for each country is multiplied by the results of the model run that had the country as the origin of the pandemic and products are then summed.

This type of averaging approach gives a measure of relative risk to early stage pandemics. One interpretation of this measure is that it provides an estimate of the expected number of cumulative infections in each country conditional on a pandemic happening with a distribution of likelihood of pandemic origin from the considered origin source (EID or lab leaks) described in 2.2.

Other methods of aggregating and interpreting the results from this model are possible and would be valuable to explore in future work. It is not at all clear that average cumulative infections is appropriate to inform all prioritization problems. Different decision problems may ask for different risk measures. One may be interested, for instance, in preventing the largest number of countries from reaching a certain threshold of cases in an early stage pandemic. In this case, it could be that a measure that gave weights to each model run based on this threshold would be more appropriate - for example. Still, the risk measure of average cumulative infections does provide a general picture of risk from early pandemics that already begins to capture how travel networks affect the pandemic risk distribution.

Results from the aggregation described above can be illustrated with heatmaps of cumulative infections at 10, 20, and 30 days after the initial outbreak. In these heatmaps color intensity represents each country's share of global aggregate cumulative infections. EID and lab leak results are explored separately in sections below.

3.1 EID model results

Figures (3.1-1 - 3.1-3) demonstrate a shift in concentration of ACIs from high EID origin potential countries (Nigeria, Democratic Republic of Congo, Equatorial Guinea, Central

[9] As risk of origin is relative (i.e. the model assumes a pandemic has happened and the country's relative contribution is determined), a linear scaling of risk assumption according to lab area will not result in absurd results (e.g. probability being greater than 1).
[10] Cumulative infections is the sum of infected and recovered populations.

African Republic) to high travel nodes. At 10 days (fig 3.1-1) the distribution of ACIs is largely concentrated in Africa - which reflect EID origin potential scores. After 20 days we see that in Africa, ACI concentrates in Nigeria and the DRC, which already stood out after 10 days. At 20 days, ACIs start to spread to Asia. Interestingly, at 30 days the picture changes significantly with the US rising as the country with the largest number of average cumulative infections in the world, Africa bearing much less of the world's share of cumulative infections (although Nigeria is itself still prominent), and India, China, Russia and Europe concentrating much of the rest of ACIs at 30 days.

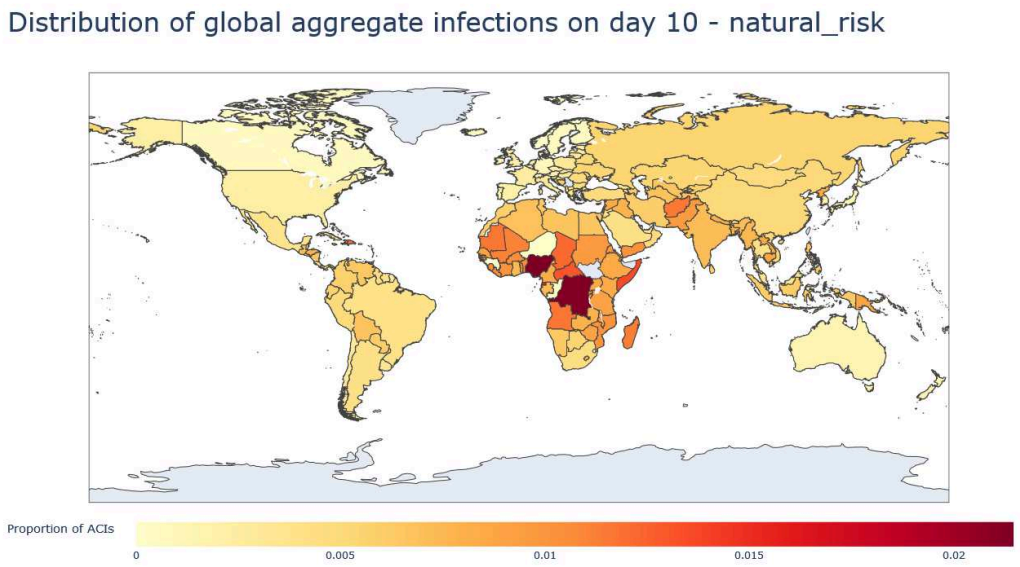

Figure 3.1-1. Heatmap of the distribution of ACIs from EID at day 10, measured as proportion of ACIs. Map shows concentration of ACIs in Central African Countries, notably: Nigeria, Democratic Republic of Congo, Equatorial Guinea, Central African Republic.

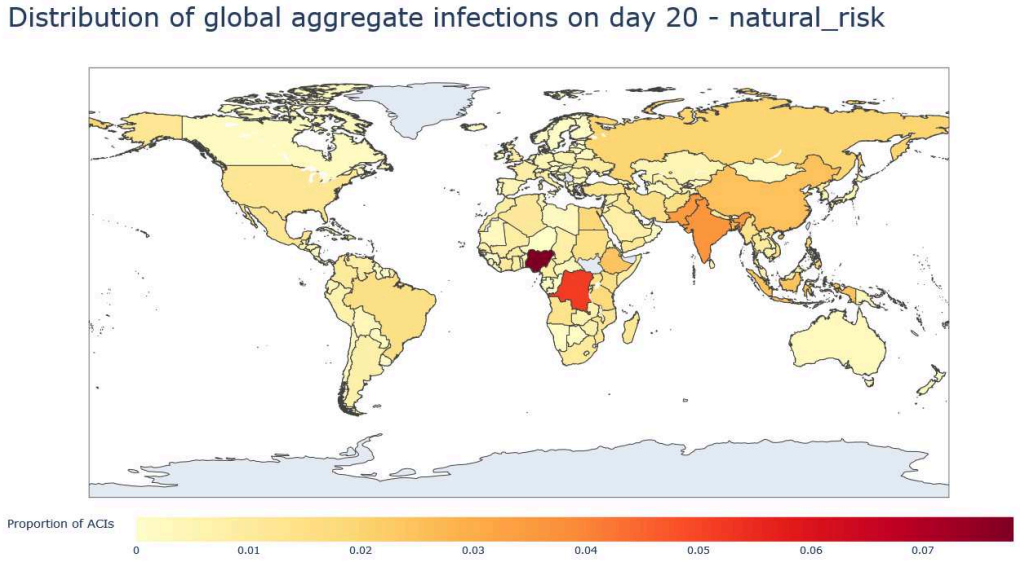

Figure 3.1-2. Heatmap of the distribution of ACIs from EID at day 20, measured as proportion of ACIs. Map shows a shift in concentration of ACIs from Central Africa towards Asia.

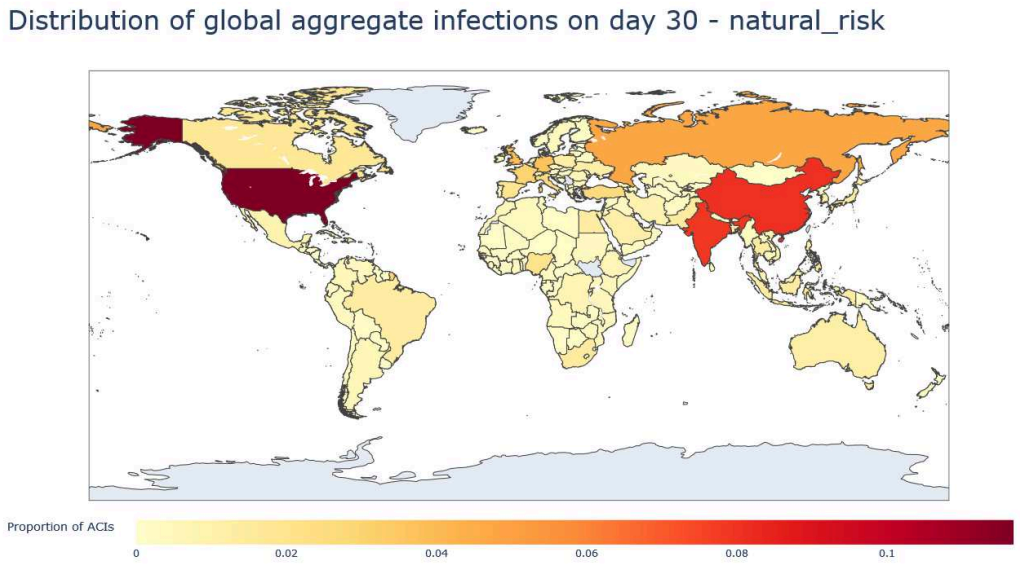

Figure 3.1-3. Heatmap of the distribution of ACIs from EID at day 30, measured as proportion of ACIs. Map shows a major jump in the proportion of ACIs located in the US and continued growth in India and China.

It is important to note that the average numbers may not match the spread of a particular disease such as SARS-Cov2 for any particular origin country. For instance, at 30 days the ACIs given by the model in the United States is about 30 million. Yet, after about 1.5 years of SARS-Cov2 there were only about 38 million confirmed infections in the US (CDC, 2020). However, looking at the model run with China as the country of origin, after 30 days there are about 5 million cumulative cases in the US. In the real world it took about 150 days for confirmed cases to reach this number (*Johns Hopkins Coronavirus Resource Center*, n.d.). If estimated unreported cases are considered it seems model results approximate cases for SARS-Cov2. Indeed, researchers estimated there were about 5 million cumulative cases in the US in April 4 2020 which is about one month after the 50th case was reported in the US (Lu et al., 2021).

Model results can also be displayed in rankings of countries most at risk according to the proportion of ACIs a country bears at a certain time.

Table 3.1-1. Day 10 top 10 countries according to proportion of ACIs

Country	I	R	Proportion of ACIs
Nigeria	5952	1704	0.0215
Congo (Democratic Republic of the)	5885	1688	0.0212
Equatorial Guinea	4142	1364	0.0154
Somalia	3740	1084	0.0135

Central African Republic	3583	1066	0.0130
Chad	3370	976	0.0122
Dominican Republic	3300	961	0.0119
Angola	3221	927	0.0116
Mauritania	3187	951	0.0116

Table 3.1-2. Day 20 top 10 countries according to proportion of ACIs

Country	I	R	Proportion ACIs
Nigeria	1202070	456668	0.0780
Congo (Democratic Republic of the)	727079	374007	0.0518
India	592721	177498	0.0362
Pakistan	542605	204129	0.0351
Bangladesh	419783	165663	0.0275
China	410423	122502	0.0251
Indonesia	382577	134817	0.0243
Ethiopia	348668	160521	0.0240
Russian Federation	301238	116703	0.0197
Egypt	260116	121677	0.0180

Table 3.1-3. Day 30 top 10 countries according to proportion of ACIs

Country	I	R	Proportion ACIs
United States of America	18546533	8172089	0.1169
China	11242643	7233951	0.0808
India	9627360	8398861	0.0789
Russian Federation	6971863	3883675	0.0475

United Kingdom of Great Britain and Northern Ireland	5580502	3171135	0.0383
Germany	5477127	2878455	0.0365
France	4546969	2273960	0.0298
Italy	3848005	2071099	0.0259
Turkey	3801141	1983450	0.0253
Nigeria	1190934	3417825	0.0202

The risk that each country poses to the world was also studied as measured by the share of average cumulative infections that each country is responsible for. In order to obtain this number we first obtain cumulative infections for each model run. Then, for each such model run cumulative infections is multiplied by the EID origin risk of this run's country of origin divided by the sum of EID origin risk. The resulting distribution of this measure of how much risk each country is responsible for is shown in figure 3.1-4. The ranking of countries that, according to this measure, impose the most risk is shown in Table 3.1-4.

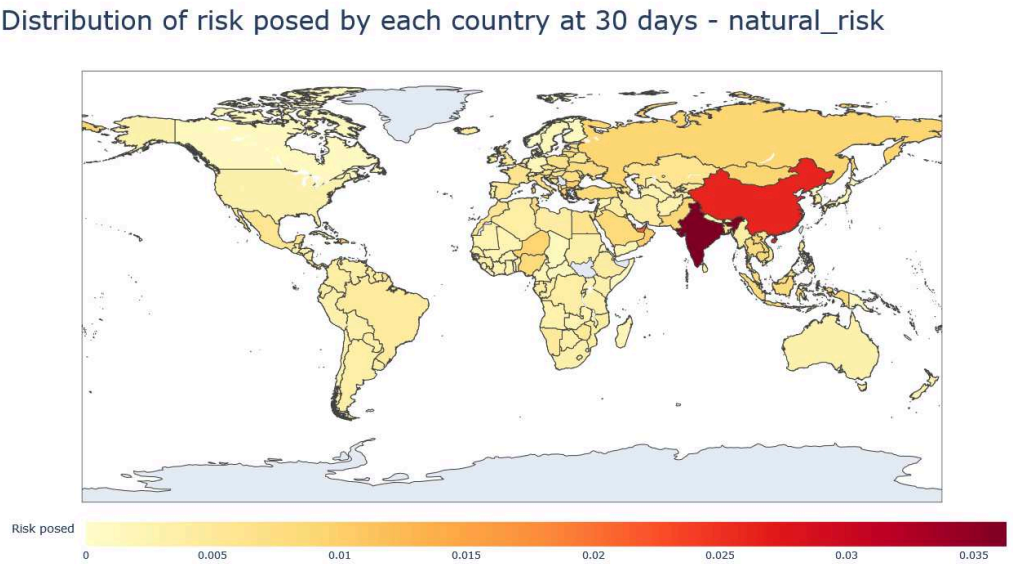

Figure 3.1-4. Distribution of risk posed by each country at day 30. India and China represent the largest risk posed according to the used metric.

Table 3.1-4. Ranking of risk posed to the world by each country at day 30

Country	Risk posed
India	0.0363
Maldives	0.0309
China	0.0262

United Arab Emirates	0.0212
Bahrain	0.0188
Seychelles	0.0165
Palau	0.0159
Cyprus	0.0139
Kuwait	0.0133
Malta	0.0118

Surprisingly, a few very small countries such as Bahrain, Seychelles, Cyprus, Palau, Malta and Singapore are very high up on this list. One potential explanation is because the Vulnerability Score from Moore et al 2017 doesn't appear to account directly for land area and total population, indicating that the risk measure misses an important factor. As such, the Vulnerability Scores likely give too much weight to very small countries since in reality more land area and population increase risk (*ceteris paribus*). These counter-intuitive results concerning small countries may thus be an artifact of the fact that the Vulnerability Score from Moore et al 2017 doesn't take into account area and population. This would suggest a more accurate model would need to transform EID origin potential so as to account for these two factors. A robustness check was run by re-obtaining the results about ACIs by considering only large countries (see Appendix 1. - Large countries robustness check). It appears that model results are robust to removing small countries with the overall ranking not changing significantly.

3.2 BSL 3+ and 4 Lab Leak Model Results

Figures 3.2-1 and 3.2.-2 show the proportion of ACIs in a country at day 10 and 30. The United States and India, both start with the majority ACIs at day 10 contributing ~⅙ of the total at this point in time (Table 3.2-1). Several other smaller countries also have a large amount of ACIs most notably Belarus and Gabon; this makes sense as Belarus possesses a large BSL 4 lab with area 1589 sqm and scores moderately on the GBBMS being 5th highest, while Gabon has two labs of (assumed average size) and possesses the worst GBBMS score of 0.92, which is over 3 times worse than the mean score.

By day 30 India has taken over being the country with the highest proportion of ACIs, with the United States being the second highest. By day 30 China has the 3rd highest, having tripled from day 10, which could be explained by China's large population allowing much greater growth in numbers of Infected and Removed.

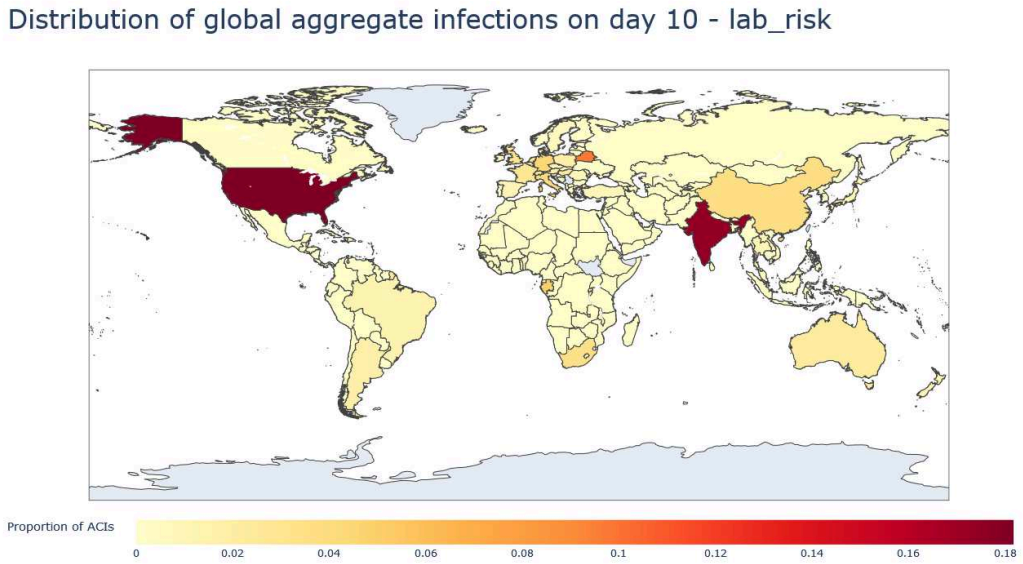

Figure 3.2-1 Heatmap of the ACIs from lab leak origin at day 10, measured as proportion of ACIs. Map shows the US and India as having the most ACIs.

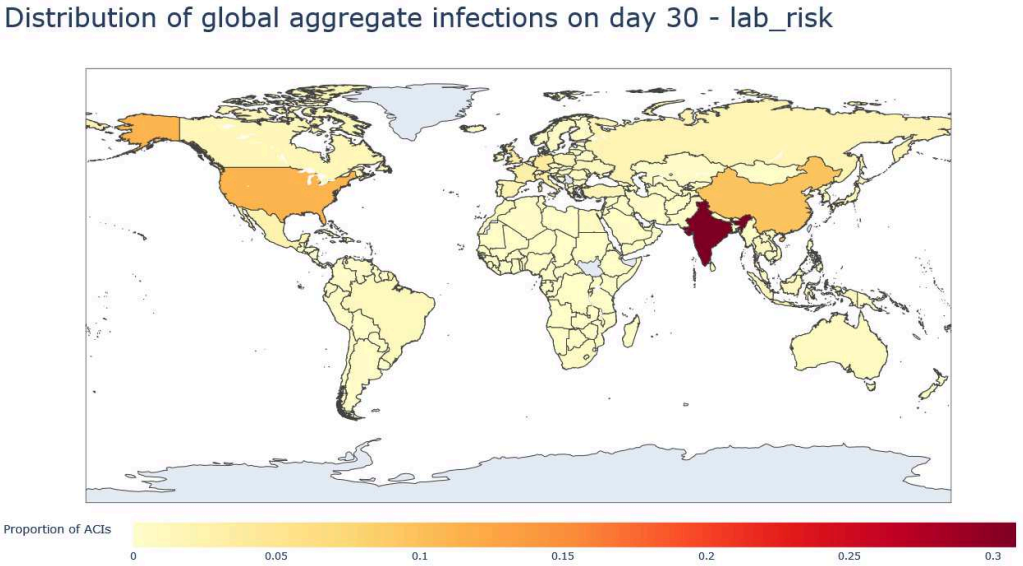

Figure 3.2-2 Heatmap of the ACIs from lab leak origin at day 30, measured as proportion of ACIs. Map shows India as having by far the greatest proportion of ACIs, the US' proportion has decreased from day 10, while China has increased.

Table 3.2-1 Day 10 top 10 countries according to proportion of ACIs

Country	I	R	Proportion of ACIs
United States of America	56171	16079	0.182
India	53546	15321	0.173

Belarus	30041	8763	0.098
Gabon	14879	4655	0.049
Germany	13784	3952	0.045
Italy	12775	3666	0.041
China	11124	3183	0.036
South Africa	10366	2975	0.034
United Kingdom of Great Britain and Northern Ireland	8965	2572	0.029
France	8504	2440	0.028

Table 3.2-3 Day 30 top 10 countries according to proportion of ACIs

Country	I	R	Proportion of ACIs
India	65145808	169866885	0.308
United States of America	29133532	56164609	0.112
China	31538291	41894582	0.096
Germany	15177979	13519508	0.038
United Kingdom of Great Britain and Northern Ireland	14810509	10177392	0.033
France	12492772	9271501	0.029
Italy	11297965	8759212	0.026
Mexico	11850272	5280472	0.022
Spain	7832971	6418677	0.019
Russian Federation	10621090	3540902	0.019

Considering the total risk posed by each country, which is measured by the share of average cumulative infections that each country is responsible for, India is the major contributor with a score of 0.3017. The United States is second with a score of 0.178 and

China third with a score of 0.0604 (table 3.2-4). Interestingly Singapore scores quite high being fourth overall, which given Singapore's relatively low lab risk, ranking 14, implies that it must be receiving a high number of infected and removed under a variety of pandemics when they are seeded in other countries. Speculatively, this might imply that Singapore's location in global travel networks might leave it highly exposed to risk of importing infections, though this would require determination of Singapore's network centrality to be confirmed.

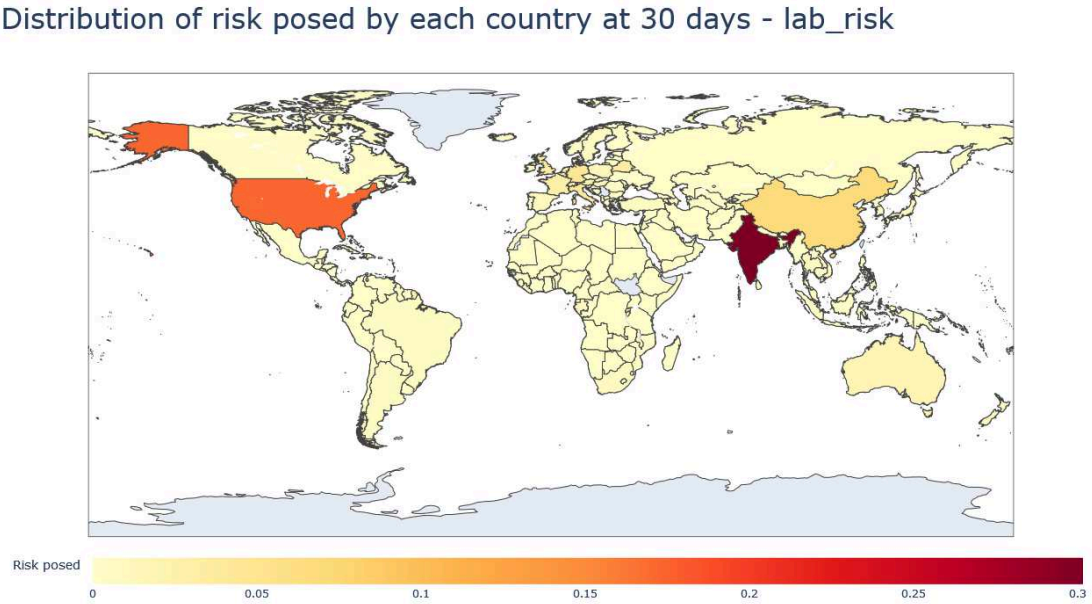

Figure 3.2-5 Distribution of risk posed by each country at day 30 from lab leaks. India represents the largest risk followed by the US and then China.

Table 3.2-4 Ranking of risk posed to the world by each country at day 30 for lab leaks.

Country	Risk posed
India	0.3017
United States of America	0.1728
China	0.0674
Singapore	0.0471
Germany	0.0459
Italy	0.0394
United Kingdom of Great Britain and Northern Ireland	0.0335

Belarus	0.0332
Switzerland	0.0295
France	0.0280

Lab leak model results appear to largely reflect the lab leak potential measure derived from the lab area and the GBBMS score of a country. One salient feature of the model result is that the risk stays concentrated in the US and India from day 10 to day 30. Both countries start with a relatively large concentration of lab leak risk and also have large populations, allowing large numbers of infected and removed to be accrued at a high rate over the 30 days. Given the limitations of GBBMS which only captures legislation requirements and does not consider implementation along with the inferred lab area data, should be considered highly uncertain.

3.3 Comparison

Relative risk scores for EID and lab leak models cannot be compared directly as the underlying risk metrics derived from EID vulnerability and GBBMS scores are not absolute measures of risk. That said some high level insights can be gleaned, most notably the difference in how risk propagates through the network of countries and the influence of initial risk distribution.

The most salient difference between the models appears to be how the proportion of aggregate infections moves from day 10 to day 30. For lab leaks, labs are most highly concentrated in the US and EU, whereas for the EID model, EID origin potential is highly concentrated in developing countries, most notably in Central Africa. This difference is important as in the lab leak model the highest risk countries also function as major travel hubs, whereas in the EID model impacted countries are less central in travel networks than the US and EU.

4. Discussion

The SIR model presented is best considered a preliminary investigation intended to guide future work. The model has several major limitations such as the country level of model abstraction, gaps in the data sets used, and assumptions pertaining to the EID and lab leak potential measures (see Appendix 4.). These introduce significant uncertainty and limit how strongly results should be used to inform prioritization. Considering results

with a healthy dose of skepticism, some high level insights can be gleaned and resulting implications for prioritization of mitigation efforts.

If you assume that natural pandemics (represented by the EID model) are the main source of danger, travel dynamics matter a lot because the main countries of EID origin as per EID vulnerability scores from Moore et al. are developing countries in Africa (Somalia, Central African Republic, Chad, South Sudan, Mauritania, Angola). These countries are only minimally connected and thus the main channels of transmissions from these poorly connected countries to the broader world is via other more connected countries. As a consequence the main dynamic at play is the travel network rather than the distribution of origin, as demonstrated in the results by the shift in average cumulative infections from Central African countries to major travel countries. One interpretation of this result relevant to prioritization of mitigation efforts is that the most connected countries are roughly both the most affected and the most important in global travel dynamics.

Conversely if you assume that lab leaks are the main source of risk, then the main countries are also quite globally connected countries (e.g. US, China, India etc.). Thus the distribution of pandemic origin ends up affecting the result at day 30. The list of most high risk countries then includes India and a few other countries that are not the richest but still heavily connected and with poor biosafety management as captured by the GBBMS. India is especially of note contributing ~⅓ of risk posed at day 30. This result is somewhat unsurprising given the presence of 5 BSL 3+ and 4 labs in the country and the second lowest biosafety management score of 0.77 behind only Gabon at 0.92, and being ~2.7 times worse than the average score. All else equal, efforts to improve biosafety management practices, and compliance in India may provide a good opportunity to reduce expected damages from lab leaks.

One relevant way to frame the results is in terms of countries being risk generators and risk importers. A risk generator is a country that contributes and exports significant risk; this is largely driven by the origin potential (EID or lab leak). A risk importing country is one that possesses low origin potential but ends up with a high proportion of ACIs. For many countries lab leak origin potential is 0 as they do not possess any BSL 3+ or 4 labs[11], and as a result they will be risk importers. This distinction is useful as it highlights how countries might wish to respond, and types of mitigation efforts that different groups might push for. For instance, risk importing countries may be best served to develop means to rapidly insulate themselves from major transport hub countries where risk is

[11] Though it should be noted that BSL 3 and even BSL2 labs are likely to contribute risk but are not represented in the data set.

most likely to accrue and then disperse e.g. US, India, China. Though social, political and economic factors may limit the reality of being able to achieve this.

Risk importing countries differ between the EID and lab leak model and interestingly the role of major travel hub countries differ between the EID and lab leak models. For the EID model US, India and China, are risk importers which concentrate cases between days 10 and days 30, whereas for lab leaks US, India and China are significant risk generators. This implies that lab leak origin pandemics should be especially of concern (as compared to EID origin) as the US, India and China are likely to generate outbreaks, resulting in large expected cumulative infections in these countries that can be distributed to the many countries that are directly connected by travel networks.

The concentration of risk in a small number of fairly rich countries (US, India, China) at day 30 in both the EID and lab leak model indicates potentially high leverage points for pandemic risk mitigation. These countries, in light of being central in the spatial network (derived from travel data) and/or contributing significant lab leak risk, have the power to significantly decrease the global risk of pandemic by implementing the right pandemic preparedness & response public policies.

Examples of public policies that might be beneficial to explore for such countries include:

- Deploying metagenomic sequencing in their main airports. Given that these countries are the nodes of the global traffic, this would increase the likelihood that spreading pathogens are identified rapidly by these airports, which can inform response actions.
- Having strict travel ban policies with pre-commitments. When travel bans are implemented under certain conditions e.g. the population density of the country of origin is lower than the population density of the country of arrival, they can be net positive. Implementing well designed travel bans with pre-commitments in central countries may help avoid worst case contagion dispersion via air travel network nodes.

5. Conclusion

The spatial network SIR model presented combines travel network data and two sources of potential pandemics (EID and lab leaks) to investigate early stage pandemic spread at the country level. Especially of interest is the BSL 3+ and 4 lab leak risk model which synthesizes a novel lab leak potential score from recently collated BSL 3+ and BSL 4 lab

data and country biosafety management practices. To the best knowledge of the authors this represents a first attempt to develop such a model.

With respect to the model limitations and uncertainties, results for both EID and lab leak models demonstrate that a handful of large, comparatively wealthy countries (US, India and China) are likely to be heavily impacted by day 30 from pandemics initiated by either source. For EID initiated pandemics, these countries appear to import risk from EID origin countries concentrated in Central Africa (Nigeria, Democratic Republic of Congo, Equatorial Guinea, Central African Republic), while for the lab leak model these countries especially the US and India generate the majority of risk. Consequently, mitigation efforts targeting these large, comparatively wealthy, highly connected countries may offer a better opportunity, all else equal, to reduce risk globally. The implication of this result runs somewhat counter to existing policy recommendations, such as Moore et al. 2017 that suggest focusing mitigation efforts on EID origin countries.

City level features such as population density, colocation of BSL 3+ and 4 labs in urban areas, and close proximity of highly connected travel nodes (major airports, ports etc.) that facilitate rapid global spread are likely to make certain cities disproportionately impactful to EID and lab leak risk. The country level of abstraction used by the model misses these features, decreasing model accuracy and limiting usefulness for targeting mitigation efforts. Exploring EID and lab leak risk with models that use cities as nodes would capture critical features, and allow more meaningful network analysis to be undertaken improving model results and usefulness for prioritization.

Data availability:

Model code and supplementary data can be made available upon request of the authors.

Acknowledgements:

Thanks to David Manheim and Steven Luby who provided feedback on an earlier version of the model and recommendations to incorporate BSL labs to create a more complete picture of risk. Thanks to the team at Global Biolabs for the work in collating information on the various BSL 3+ and 4 labs which made an invaluable resource for the project. Thanks to SERI who provided an initial opportunity for Simeon to explore aspects of pandemic risk and prioritization that informed work on the paper. Thanks to Sahil Shah who provided input in the preceding work and David Denkenberger who provided feedback on the conference paper. Thanks to Eric Thierry for help on improving the simulation code.

6. References

Amiri Mehra, A. H., Abbasi, Z., & Zamani, I. (2020). Parameter Estimation and Prediction of COVID-19 Epidemic Turning Point and Ending Time of a Case Study on SIR/SQAIR Epidemic Models. *Computational and Mathematical Methods in Medicine*, *2020*, e1465923. https://doi.org/10.1155/2020/1465923

Baum, S. D., Armstrong, S., Ekenstedt, T., Häggström, O., Hanson, R., Kuhlemann, K., Maas, M. M., Miller, J. D., Salmela, M., Sandberg, A., Sotala, K., Torres, P., Turchin, A., & Yampolskiy, R. V. (2019). Long-term trajectories of human civilization. *Foresight*, *21*(1), 53–83. https://doi.org/10.1108/FS-04-2018-0037

Beckstead, N. (2013). On the overwhelming importance of shaping the far future. *Rutgers University*.

Bostrom, N. (2002). Existential Risks: Analyzing Human Extinction Scenarios and Related Hazards. *Journal of Evolution and Technology*, *9*(1). https://www.nickbostrom.com/existential/risks.pdf

Bostrom, N. (2013). Existential Risk Prevention as Global Priority. *Global Policy*, *4*(1), 15–31. https://doi.org/10.1111/1758-5899.12002

Bostrom, N., & Cirkovic, M. M. (2008). *Global Catastrophic Risks*. Oxford University Press, Oxford.

Brault, A. C., Powers, A. M., Medina, G., Wang, E., Kang, W., Salas, R. A., De Siger, J., & Weaver, S. C. (2001). Potential Sources of the 1995 Venezuelan Equine Encephalitis Subtype IC Epidemic. *Journal of Virology*, *75*(13), 5823–5832. https://doi.org/10.1128/JVI.75.13.5823-5832.2001

Caplan, B. (2008). The totalitarian threat. In B. Caplan, *Global Catastrophic Risks*. Oxford University Press. https://doi.org/10.1093/oso/9780198570509.003.0029

CDC. (2020, March 28). *COVID Data Tracker*. Centers for Disease Control and Prevention. https://covid.cdc.gov/covid-data-tracker

Cumulative Cases—Johns Hopkins Coronavirus Resource Center. (n.d.). Johns Hopkins Coronavirus Resource Center. Retrieved 19 November 2021, from https://coronavirus.jhu.edu/data/cumulative-cases

DEFRA. (2008). *Foot and mouth disease 2007: A review and lessons learned*. Department for Environment Food and Rural Affairs. https://assets.publishing.service.gov.uk/government/uploads/system/uploads/attachment_data/file/250363/0312.pdf

Denkenberger, D., Sandberg, A., Tieman, R. J., & Pearce, J. M. (2021). Long-term cost-effectiveness of interventions for loss of electricity/industry compared to artificial general intelligence safety. *European Journal of Futures Research*, *9*(1), 11. https://doi.org/10.1186/s40309-021-00178-z

Denkenberger, D., Sandberg, A., Tieman, R. J., & Pearce, J. M. (2022). Long term cost-effectiveness of resilient foods for global catastrophes compared to artificial general intelligence safety. *International Journal of Disaster Risk Reduction*, *73*, 102798. https://doi.org/10.1016/j.ijdrr.2022.102798

European Commission. (n.d.). *Dynamic Data Hub*. Retrieved 29 March 2023, from https://migration-demography-tools.jrc.ec.europa.eu/data-hub/

Global Biolabs. (n.d.). Global Biolabs. Retrieved 19 March 2023, from https://www.globalbiolabs.org

Holmes, E. C., Rambaut, A., & Andersen, K. G. (2018). Pandemics: Spend on surveillance, not prediction. *Nature*, *558*(7709), 180–182. https://doi.org/10.1038/d41586-018-05373-w

Johnson, N. P., & Mueller, J. (2002). Updating the accounts: Global mortality of the 1918-1920 'Spanish' influenza pandemic. *Bulletin of the History of Medicine*, *76*(1), 105–115. Scopus.

Jones, B. F. (2010). Age and Great Invention. *Review of Economics and Statistics*, *92*(1), 1–14.

https://doi.org/10.1162/rest.2009.11724

Koblentz, G. D., Ernhardt, B., Rodgers, J., Lentzos, F., Ameneiros, M., Houser, R., & Wingo. (2021). *Global BioLabs* [Policy]. https://www.globalbiolabs.org

Koblentz, G. D., Ernhardt, B., Rodgers, J., Lentzos, F., Ameneiros, M., Houser, R., & Wingo. (2023). *Global BioLabs Report 2023* (Global Biolabs) [Policy]. Schar School of Policy and Government. https://static1.squarespace.com/static/62fa334a3a6fe8320f5dcf7e/t/6412d3120ee69a4f4efbec1f/1678955285754/KCL0680_BioLabs+Report_Digital.pdf

Laboratory biosafety manual, 4th edition (No. 4; LABORATORY BIOSAFETY MANUAL, p. 124). (2020). World Health Organization. https://www.who.int/publications-detail-redirect/9789240011311

Lu, F. S., Nguyen, A. T., Link, N. B., Molina, M., Davis, J. T., Chinazzi, M., Xiong, X., Vespignani, A., Lipsitch, M., & Santillana, M. (2021). Estimating the cumulative incidence of COVID-19 in the United States using influenza surveillance, virologic testing, and mortality data: Four complementary approaches. *PLOS Computational Biology*, *17*(6), e1008994. https://doi.org/10.1371/journal.pcbi.1008994

MacAskill, W. (2022). *What we owe the future: A million-year view*. Oneworld.

Margolis, J., & Parfit, D. (1986). Reasons and Persons. *Philosophy and Phenomenological Research*, *47*(2), 311. https://doi.org/10.2307/2107444

Merler, S., Ajelli, M., Fumanelli, L., & Vespignani, A. (2013). Containing the accidental laboratory escape of potential pandemic influenza viruses. *BMC Medicine*, *11*(1), 252. https://doi.org/10.1186/1741-7015-11-252

Millett, P., & Snyder-Beattie, A. (2017). Existential Risk and Cost-Effective Biosecurity. *Health Security*, *15*(4), 373–383. https://doi.org/10.1089/hs.2017.0028

Moore, M., Gelfeld, B., Okunogbe, A., & Paul, C. (2017). Identifying Future Disease Hot Spots: Infectious Disease Vulnerability Index. *Rand Health Quarterly*, *6*(3), 5.

Muroya, Y., Enatsu, Y., & Kuniya, T. (2013). Global stability for a multi-group SIRS epidemic model with varying population sizes. *Nonlinear Analysis: Real World Applications*, *3*(14), 1693–1704. https://doi.org/10.1016/j.nonrwa.2012.11.005

Ord, T. (2020). *The Precipice: Existential Risk and the Future of Humanity*. Hachette Books.

Perry, R. D., & Fetherston, J. D. (1997). Yersinia pestis—Etiologic agent of plague. *Clinical Microbiology Reviews*, *10*(1), 35–66. https://doi.org/10.1128/CMR.10.1.35

Pujari, B., & Shekatkar, S. (2020). *Multi-city modeling of epidemics using spatial networks: Application to 2019-nCov (COVID-19) coronavirus in India*. https://doi.org/10.1101/2020.03.13.20035386

Recchi, E., Deutschmann, E., & Vespe, M. (2019). *Estimating transnational human mobility on a global scale* [Working Paper]. https://cadmus.eui.eu//handle/1814/62326

Recchi, E., Deutschmann, E., & Vespe, M. (2020). *Global transnational mobility dataset* [Data set]. Robert Schuman Centre for Advanced Studies. https://hdl.handle.net/1814/67634

Sandberg, A., & Bostrom, N. (2008). *Global catastrophic risks survey* [Technical Report (vol 1)]. Future of Humanity Institute, Oxford University.

Schuerger, C., Abdulla, S., & Puglisi, A. (2022). *Mapping Biosafety Level-3 Laboratories by Publications*. Center for Security and Emerging Technology. https://doi.org/10.51593/20220019

Sharma, N., Verma, A., & Gupta, A. (2021). Spatial network based model forecasting transmission and control of COVID-19. *Physica A: Statistical Mechanics and Its Applications*, *581*, 126223. https://doi.org/10.1016/j.physa.2021.126223

Shulman, C., & Thornley, E. (forthcoming). How Much Should Governments Pay to Prevent Catastrophes? Longtermism's Limited Role. In J. Barrett, H. Greaves, & D. Thorstad (Eds.), *Essays on Longtermism*. Oxford University Press.

Spreeuwenberg, P., Kroneman, M., & Paget, J. (2018). Reassessing the Global Mortality

Burden of the 1918 Influenza Pandemic. *American Journal of Epidemiology*, *187*(12), 2561–2567. https://doi.org/10.1093/aje/kwy191

Wurtz, N., Papa, A., Hukic, M., Di Caro, A., Leparc-Goffart, I., Leroy, E., Landini, M. P., Sekeyova, Z., Dumler, J. S., Bădescu, D., Busquets, N., Calistri, A., Parolin, C., Palù, G., Christova, I., Maurin, M., La Scola, B., & Raoult, D. (2016). Survey of laboratory-acquired infections around the world in biosafety level 3 and 4 laboratories. *European Journal of Clinical Microbiology & Infectious Diseases*, 35(8), 1247–1258. https://doi.org/10.1007/s10096-016-2657-1

Supplementary Files

This is a list of supplementary files associated with this preprint. Click to download.

- Preprintv1.1AppendixPandemicriskEIDandlableaksSIR1.pdf