American college of technology

Department of computer science

Preventive measures during pandemic in data mining

Hanna Meressa

Problem relevance

Pandemics are large-scale outbreaks of infectious disease that can greatly increase morbidity and mortality over a wide geographic area and cause significant economic, social, and political disruption. Evidence suggests that the likelihood of pandemics has increased over the past century because of increased global travel and integration, urbanization, changes in land use, and greater exploitation of the natural environment.

The international community has made progress toward preparing for and mitigating the impacts of pandemics. The 2003 severe acute respiratory syndrome (SARS) pandemic and growing concerns about the threat posed by avian influenza led many countries to devise pandemic plans; even now, a time Covid-19 is one off the big issue in our world.

I select this topic

Why we are not control the pandemics disease? In addition, why we are minimize the risk predict before the diseases killing more people. This is why I choose this title. Then, I am interested in selected for "preventive measures during the pandemic in data mining" because if you have measured preventive the data, you will protect and reduce the risk.

Relevance (local relevance is a plus)

In our country, we have many problem and then we must measure preventively during pandemic. Because if we doing that the coming caciques will be lease and indicated organization are to prepper materials, medicine, and other things are ready everything.

Background

Pandemic is define as “an epidemic occurring over a very wide area, crossing international boundaries, and usually affecting a large number of people”. Their geographic scale rather than the severity of illness therefore, identify pandemics. For example, in contrast to annual seasonal influenza epidemics, *pandemic influenza* is define as “when a new influenza virus emerges and spreads around the world, and most people do not have immunity”. I do not want consider endemic diseases because pandemic and endemic are constantly present in particular locations or regions.

Endemic diseases are far more common than pandemics and can have significant negative health and economic impacts, especially in low- and middle-income countries with weak health systems. This pepper does specifically consider how to reduce the risks, impacts, and mitigation of pandemics as well as knowledge gaps.

Risks

- Pandemics are occur throughout history and appear to be increasing in frequency, particularly because of the increasing emergence of viral disease from animals.
- Pandemic risk is driven by the combined effects of spark risk (*where* a pandemic is likely to arise) and spread risk (*how likely* it is to diffuse broadly through human populations).
- Some geographic regions with high spark risk, including Central and West Africa, lag behind the rest of the globe in pandemic preparedness.
- Probabilistic modeling and analytical tools such as exceedance probability (EP) curves are valuable for assessing pandemic risk and estimating the potential burden of pandemics.
- Influenza is the most likely pathogen to cause a severe pandemic. EP analysis indicates that in any given year, a 1 percent probability exists of an influenza pandemic that causes nearly 6 million pneumonia and influenza deaths or more globally.

Impacts

- Pandemics can cause significant, widespread increases in morbidity and mortality and have disproportionately higher mortality impacts on LMICs.
- Pandemics can cause economic damage through multiple channels, including short-term fiscal shocks and longer-term negative shocks to economic growth.
- Individual behavioral changes, such as fear-induced aversion to workplaces and other public gathering places are a primary cause of negative shocks to economic growth during pandemics.
- Some pandemic mitigation measures can cause significant social and economic disruption.
- In countries with weak institutions and legacies of political instability, pandemics can increase political stresses and tensions. In these contexts, outbreak response measures such as quarantines have sparked violence and tension between states and citizens.

Mitigation

- Pathogens with pandemic potential vary widely in the resources, capacities, and strategies required for mitigation. However, there are also common prerequisites for effective preparedness and response.
- The most cost-effective strategies for increasing pandemic preparedness, especially in resource-constrained settings, consist of investing to strengthen core public health infrastructure, including water and sanitation systems; increasing situational awareness; and rapidly extinguishing sparks that could lead to pandemics.
- Once a pandemic has started, a coordinated response should be implemented focusing on maintenance of situational awareness, public health messaging, reduction of transmission, and care for and treatment of the ill.

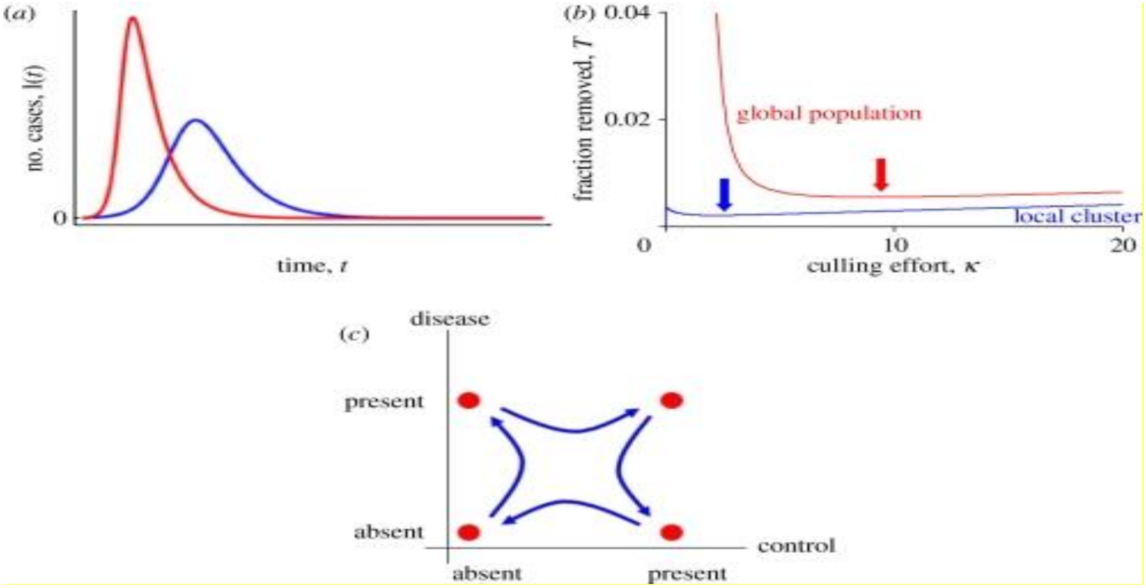
- Successful contingency planning and response require surge capacity—the ability to scale up the delivery of health interventions proportionately for the severity of the event, the pathogen, and the population at risk.
- For many poorly prepared countries, foreign aid providers will deliver surge capacity likely. This is a tenable strategy during localized outbreaks, but global surge capacity has limits that likely will be reach during a full-scale global pandemic as higher-capacity states focus on their own populations.
- Risk transfer mechanisms, such as risk pooling and sovereign-level catastrophe insurance, provide a viable option for managing pandemic risk.

Knowledge Gaps

- Spending and costs specifically associated with pandemic preparedness and response efforts are poorly track.
- There is no widely accepted, consistent methodology for estimating the economic impacts of pandemics.
- Most data regarding the impacts of pandemics and the benefits and costs of mitigation measures come from high-income countries (HICs), leading to biases and potential blind spots regarding the risks, consequences, and optimal interventions specific to LMICs.

QUANTITATIVE MODELS

Making the kinds of predictions listed in requires some kind of 'model', even if this is only a mental model based on previous experience, expert opinion or a back-of-the-envelope calculation. For some questions, this may be adequate. For other questions, these informal approaches can be highly unreliable, the reason being that infectious diseases have nonlinear dynamics. One way to express this is that the biggest risk factor for acquiring an infection is the presence of infectious individuals, which introduces positive feedback into epidemic processes, in turn making the expected trajectory of an epidemic or the likely impact of control measures considerably more difficult to predict than is the case for non-communicable diseases such as stroke, cancer or obesity. On occasion, these nonlinearities can make infectious disease dynamics counterintuitive.



Non-linearities in infection dynamics. (a) Force-of-infection and duration of an outbreak. Reduced force-of-infection can increase the expected duration of an outbreak, as illustrated by two numerical realizations of the standard susceptible–latent–infectious–recovered (SLIR) model. Both have mean latent period = 1 time unit and mean recovery period = 1 time unit but the per capita transmission rate is halved from high (red) to low (blue), resulting in a smaller but longer lasting outbreak. (b) Impact of pre-emptive culling. Analysis of the impact of increased pre-emptive culling effort on the total loss of livestock farms during a FMD epidemic. Fraction of the global population removed (red line) and fraction of the global population removed within a single local cluster of 50 farms (blue line) are shown as functions of the number of pre-emptive culls per case. Parameter values used approximate those for the 2001 UK FMD epidemic. The culling effort minimizing global losses (red arrow) is almost 4× higher than that minimizing local losses (blue arrow). Figure re-drawn from. (c) Relationship between the presence of disease and the implementation of control. A local host population is shown moving (blue arrows) in sequence between four states (red dots): first, disease is introduced; then control is implemented; then disease is eliminated; then control ceases. This describes the expected sequence of events when control is implemented reactively and locally. Depending on how many local populations are in each of the four states at a given time point, a cross-sectional study could generate a positive or zero correlation between levels of disease and control effort as easily as a negative one (the naive expectation), even if control is fully effective.

Methods

APPROACHES TO PREDICTION

Whatever methodology is use, a key challenge in making any kind of prediction is to establish the extent to which the past is likely to be an accurate guide to the future. There are different levels of predictability that maybe taken to apply, expressed here in terms of the structure and parameter estimates of a mathematical model fitted to a previous FMD epidemic but now having to be adapted to make predictions about a new epidemic.

i. No change. The input data, parameter values and the model itself are all judged applicable to the new epidemic.
ii. The input data change. As a simple example, there may be changes to the location, size and species composition of livestock farms. Such changes, if known, would be readily incorporate into a new model.
iii. The parameter values change. This would be the case if, for example, a different strain of FMD virus was introduced, perhaps with different transmission characteristics. Or if farming practices had altered in ways, such as improvements in biosecurity, which changed FMD transmission rates. These kinds of changes would be difficult to quantify *a priori*, and new parameter values may need to be estimated from early epidemic data.
iv. The model changes (i.e. the original model structure/assumptions are incorrect for the new epidemic). This could be the case if, for example, the strain of FMD virus introduced was much more liable to airborne transmission, requiring that this feature be built into the models.

Addressing (ii) is straightforward, (iii) is reliant on methods for rapid estimation and re-estimation of parameter values as data accumulate (see, for example, for a state-of-the-art application), but (iv) will most probably depend on timely input from disease experts.

In practice, various approaches have been used for making predictions about future disease risks. These include expert opinion, statistical methods, simulation modelling, and risk modelling.

a) Expert opinion

There is a now a substantial literature on methodologies for systematically surveying expert opinion. One example of their application to future infectious disease risks derives from a 2006 UK government Foresight project .There were two components to this study: the identification and ranking of future disease risks (or, more correctly, ‘hazards’) and the identification of factors involved in changes to these risks in the future (so-called ‘drivers’ of risk). As with all studies of this nature, the results necessarily reflect the expertise, interests and geographical locations of the participants as well as, crucially, the precise questions they were asked.

b) Statistical methods

The statistical workhorse of risk factor analysis is the generalized linear model (GLM) .The methodology is well established and its application routine, but it has its limitations for the analysis and prediction of infectious disease data. As discussed in §2, the risks of individuals becoming infected are dynamic and not independent of one another; this introduces spatial and temporal autocorrelations that may be difficult to account for .This problem is exacerbated when the intention is to predict future risks, perhaps under different scenarios such as a range of possible intervention strategies. Extrapolations of statistical models have limited value in this context. Hence, as discussed below, dynamic models have often been preferred.

The outputs of a risk factor analysis are routinely expressed in terms of odds ratios associated with each of the risk factors in the model, identifying the main drivers of risk in the study population. Often, however, it is also useful to calculate the risk (e.g. the probability of being infected) for each individual in the population, based on their individual risk factors.

The pandemic disease outbreaks had been caused by heterogeneous factors that involved vast area of study and tremendous data collection activities. One of the data collection activities can be done by using the Electronic Health Record and WHO.

Data collection and analysis

Qualitative data was extract from papers included in the review. The data extracted include specific details about countries experience in responding SARS and COVID-19. Finally, textual data were extract from papers included in the review.

Data Collection and Preparation

The first challenge is the restriction on data collection. Social media providers use unknown and undocumented sampling filtration algorithms that allow for collecting only a sample of the overall data. In addition, there are restrictions on some private data that may be needed for the detection process. In addition, users may not include some other important pandemics information. This may lead to inaccurate results produced by the tools of disease trend detection.

The pandemic disease outbreaks had been caused by heterogeneous factors that involved vast area of study and tremendous data collection activities. One of the data collection activities can be done by using the Electronic Health Record and WHO.

Data extraction and management the following data was abstracted to a Microsoft Excel spreadsheet.

Qualitative data was extracted from papers included in the review. The data extracted include specific details about countries experience in responding pandemic. Finally, textual data were extract from papers included in the review.

Measurement

Accurate, affordable, and valid measurements “are the basis for quality of care assessments” (Peabody and others 2004, 771). For too long, routine measures of quality in LMICs (ow- and middle-income countries) relied on structural elements (rosters, catalogs, and inventories of coverage and access), giving little thought to how these elements improve health. Such elements are relatively easy to count and measure, but are only remotely linked to better outcomes. Improving quality requires measurement of the care process—that is, what providers do when they see patients (Ansong-Tornui and others 2007; Peabody, Taguiwalo, and others 2006; Peabody and others 2011).

Measurement of the care process is critical, creating awareness of deficits in practice, gaps in care, and accountability at the individual and system levels, which improves focus and motivation. To serve as an instrument of change and accountability, provider-level measurement needs to be ongoing and cyclical. Transparency of results can increase knowledge and change intentions, but requires a supportive context to be effective (National Patient Safety Foundation 2015).

When coupled with useful feedback and done in a timely manner, measurement is the foundation for improving quality. If the measures are reliable, affordable, and anchored in valid, evidence-based criteria, quality of care can be followed over time and the impact of policy interventions can be assessed (Felt-Lisk and others 2012). Various quality measures have been developed, each with its own set of advantages and disadvantages. Although no measure is perfect, adequate measures exist, and every health system—from small clinics to national governments—can benefit from measurement. Feedback has the potential to promote improvement, but studies are limited, tending to focus on health care report cards (Baker and Cebul 2002; Dranove and others 2003; Kolstad 2013; Shaller and others 2003), which include public disclosure of quality scores that may not provide the same motivation to improve scores as when feedback is provided privately.

Table 10.1 Methods for Measuring Quality of the Care Process

Method	Advantage	Issues
Chart abstraction or review of medical record	• Nearly ubiquitous and theoretically could be obtained after the patient-provider encounter; in practice, record keeping in most LMICs is inadequate • Electronic medical record technology: improved uniformity, legibility, communication • Records of clinical events	• May lack relevant clinical details, especially when written for other purposes, such as legal protection • Poor record keeping and documentation lead to incomplete and inaccurate content • Illegibility of handwritten notes • Inaccuracies in the process of abstracting to produce data suitable for analysis • High costs involved in training medical abstractors • Variation in documentation practices across providers, facilities, and countries
Direct observation and recording of visits	• Records of clinical events • First-hand observation of actual encounters	• Ethical considerations • Need to inform providers and patients, which can induce the Hawthorne effect (bias when participant changes his or her behavior as a result of being evaluated) • High cost of training observers • Variations across observers
Administrative data	• Available in most facilities • Ubiquitous and inexpensive to collect when data collection system is in place	• Lack sufficient clinical detail • Inaccuracies in content • Poor data collection or management systems, especially in LMICs
Standardized patients	• The gold standard for process measurement • Captures technical and interpersonal elements of process • Reliable over a range of conditions, providing valid measurements that accurately capture variation in clinical practice among providers across patients	• Expensive • Not practical for routinely evaluating quality • Limited range of applicability (works best for adult conditions and conditions that can be simulated)
Clinical vignettes	• Can measure quality within a group of providers and evaluate quality at the population level • Responsive to variations in quality • Cases simulate actual patient visit and evaluate physician's knowledge • Validated against other methods and criteria for standard-of-quality measurement • Useful for comparison studies • Easy and inexpensive to administer • Ability to collect data independently	• Potential resistance of providers to complete the vignettes • Different methods for administering vignettes • Instrument validation • Link to patient-level data

Sources: Bertelsen 1981; Peabody and others 2004; Peabody and others 2011; Peabody, Nordyke, and others 2006.
Note: LMICs = low- and middle-income countries.

Modeling and Evaluation

Linear Regression

One of the data mining techniques that are used for prediction tasks is Linear Regression. In a problem with one predictor, this technique tries to find the best line to fit. That line could relate the predictor and prediction values. The extended version of this one predictor regression is called multiple linear regression and used for multiple predictor problems. This type of linear regression is utilized in this study.

Long Short-Term Memory (LSTM)

LSTM is an artificial recurrent neural network that is an effective model for the prediction of time series where data is sequential [9]. By storing the past in hidden states, they can predict the outputs more accurately. In this study, the aim was to estimate the number of positive COVID-19 cases through time; this is a well-suited task for the LSTM model. Thus, we used this model in this study.

The linear regression model and a 3-layer LSTM model are employed to predict the daily new cases. Rapid Miner Studio 9.3.001 and Python 3.7.3 are used for modeling, estimating, predict and evaluation.

Clean, remove noise, and avoid missing values

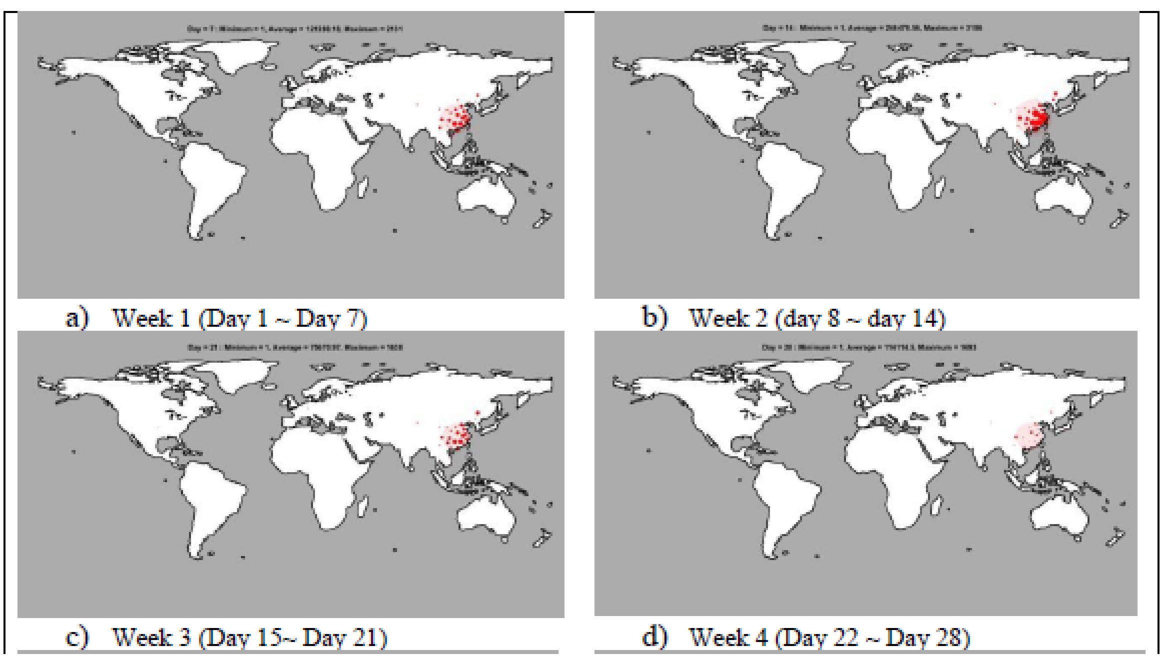
Accordingly, I am using WHO report and Google data as it is. Then first identifying missing values this means us writing the python code for standard and non-standard missing values next to that we must list missing values and separated unexpected missing value type, and summarizing missing values. Finally, we can remove noise and avoid values.

Data Mining Techniques/Algorithms

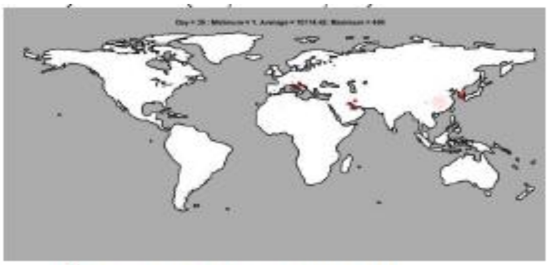
Logistic Regression (LR) Logistic regression (LR) is used to determine the association between categorical dependent variables against the independent variables [9]. LR is used when the dependent variable has two values such as 0 and 1, yes and no or true and false and thus it is called binary logistic regression.

Table 1. The results of the best models evaluated on coronavirus dataset (January 22 to April 12).

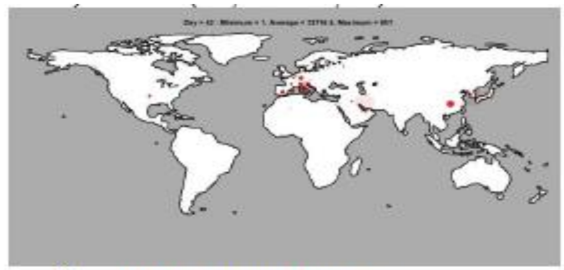
Maximum Number of confirmed cases in a day		Number of Neighbors	Interval of days [min,max]	MSE		MAE	
				Value	Percent	Value	Percent
<200	Train	1	[14,16]	9.4770	0.25%	1.3001	1.41%
	Test			463.1623	2.43%	11.5343	6.86%
[200,1000)	Train	4	[14,36]	18.7168	1.17%	2.3672	5.68%
	Test			187560	11.09%	106.1987	26.02%
≥1000	Train	1	[14,21]	709.1846	5.757e-05%	13.5065	0. 0384%
	Test			1.7469e+07	2.49%	2146	6.5%



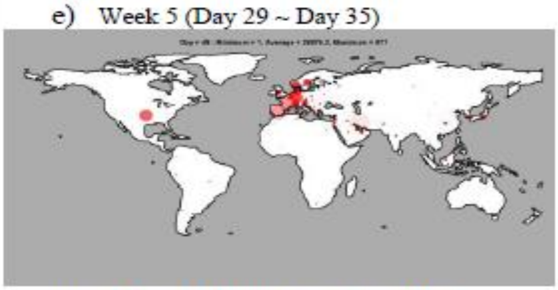
a) Week 1 (Day 1 ~ Day 7)
b) Week 2 (day 8 ~ day 14)
c) Week 3 (Day 15~ Day 21)
d) Week 4 (Day 22 ~ Day 28)



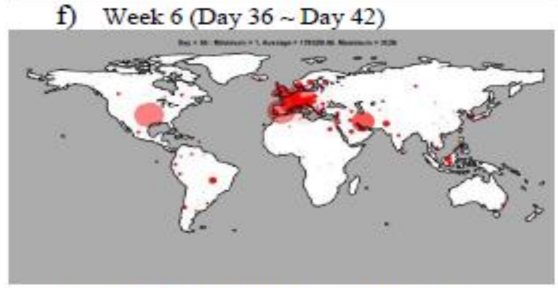
e) Week 5 (Day 29 ~ Day 35)



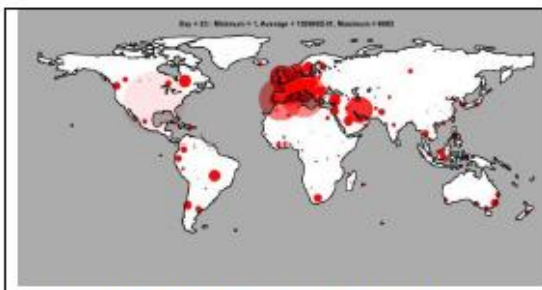
f) Week 6 (Day 36 ~ Day 42)



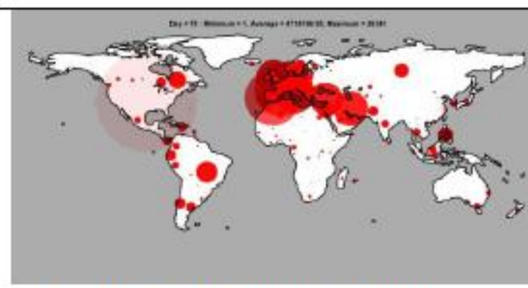
g) Week 7 (Day 43 ~ Day 49)



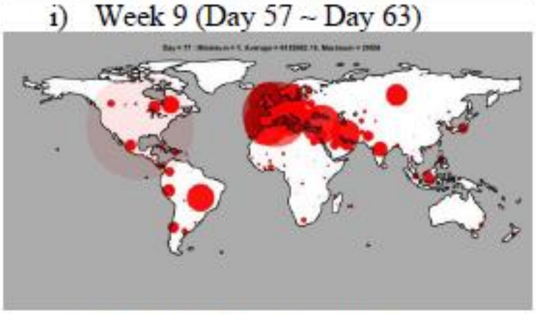
h) Week 8 (Day 50 ~ Day 56)



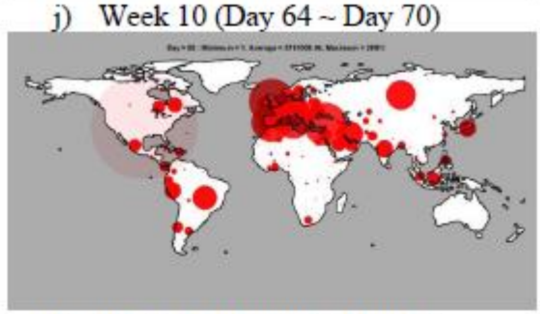
i) Week 9 (Day 57 ~ Day 63)



j) Week 10 (Day 64 ~ Day 70)



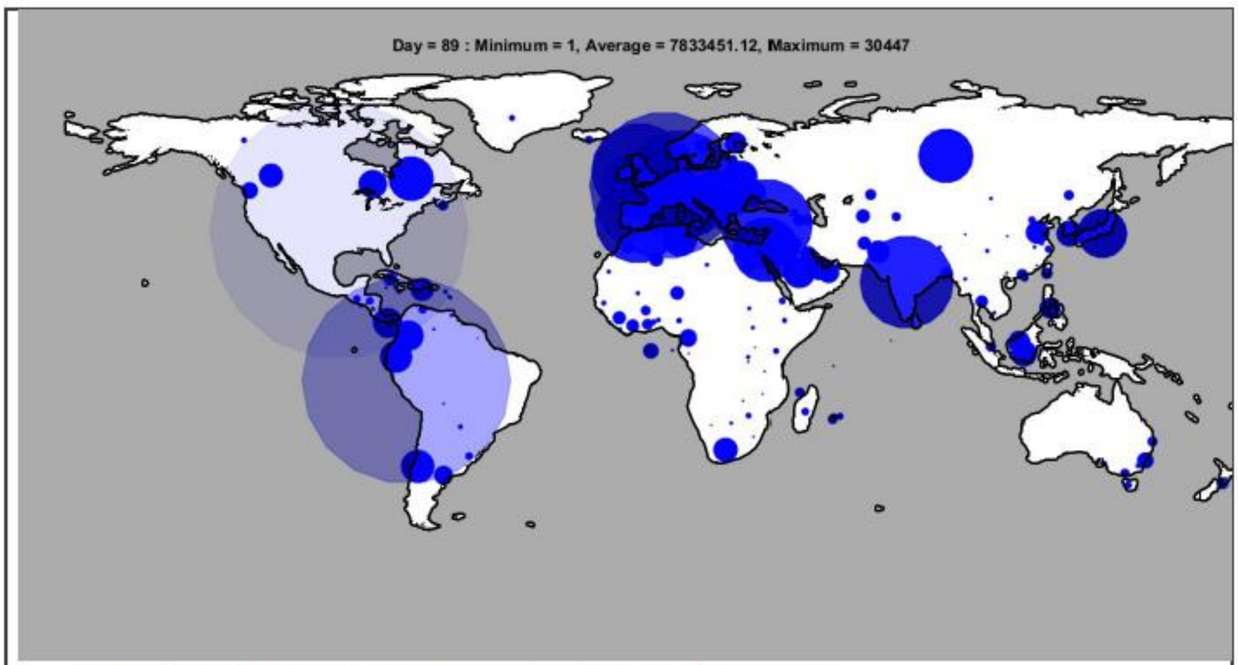
k) Week 11 (Day 71 ~ Day 77)



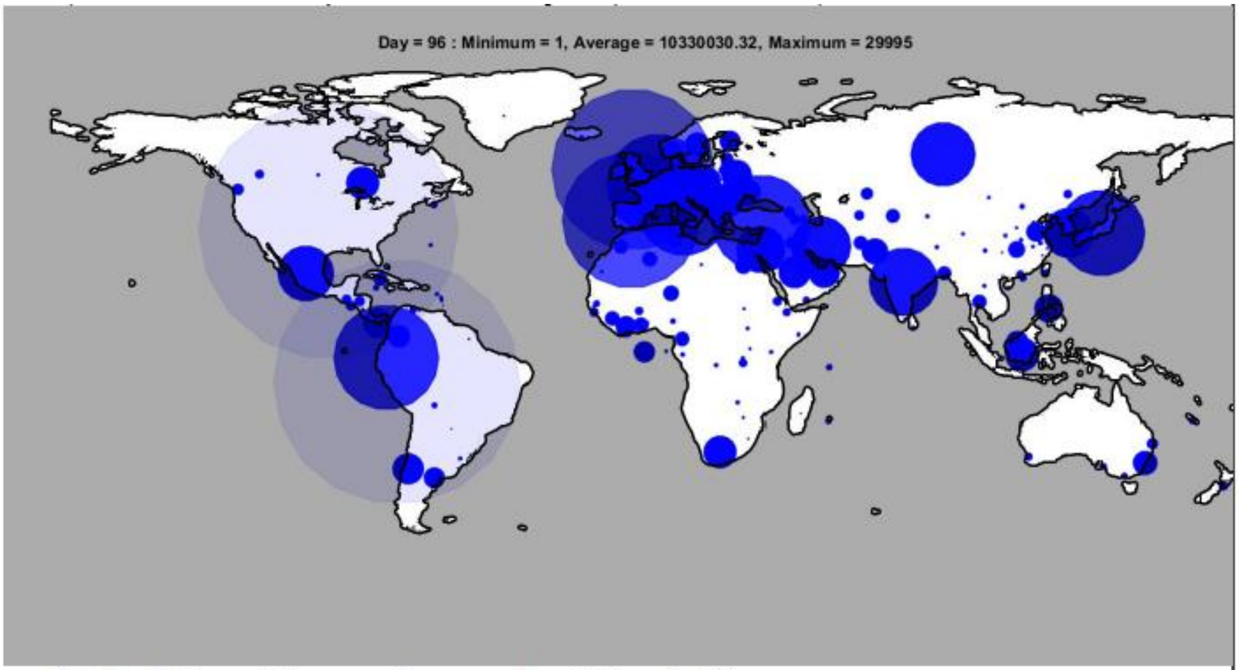
l) Week 12 (Day 78 ~ Day 82)

Table 1. Forecast the COVID-19 new cases for the next two weeks.

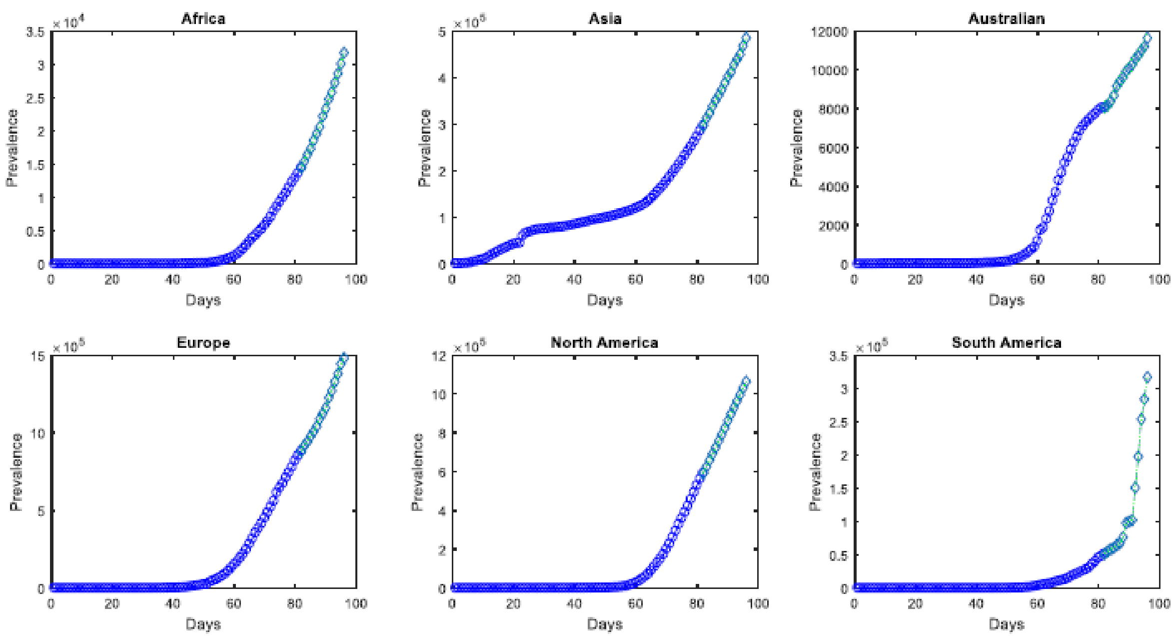
Date	*Continents*						*Total number of confirmed cases*
	Africa	Asia	Australian	Europe	North America	South America	
22 Jan~12-Apr	14591	297094	8074	881460	33574	51405	1286198
13-Mar	990	16144	110	38863	32897	4249	93253
14-Mar	1016	10385	199	22842	33692	2888	71022
15-Apr	877	15445	303	34517	31796	2005	84943
16-Apr	1287	12784	480	31087	33391	4430	83459
17-Apr	994	10254	263	32574	33269	2367	79721
18-Apr	1045	12262	182	49749	32997	9134	105369
19-Apr	1638	14087	331	34600	38903	21460	111019
20-Apr	1163	14112	142	36179	34797	2205	88598
21-Apr	1448	9241	195	64827	32164	1909	109784
22-Apr	1053	16079	303	43845	33590	48988	143858
23-Apr	1456	11806	167	58723	32599	46802	151553
24-Apr	1398	10233	257	51222	35186	56138	154434
25-Apr	1575	18489	194	64742	32600	30109	147709
26-Apr	1628	17000	442	37134	33574	32861	122639
Total	32159	485415	11642	1482364	505029	316950	2833559
Prevalence growth rate	54.63	38.80	30.65	40.54	93.35	83.78	54.61



a) Prediction of the prevalence on day 89 (end of week 12).



b) Prediction of the prevalence on day 96 (week 13).



One Figure Legends

Figure 1. Preprocessing steps on coronavirus dataset.
Figure 2. The structure of the proposed model.
Table 1. The results of the best models evaluated on coronavirus dataset.
Figure 3. Visualize the outbreak over the days.
Table 2. Forecast the COVID-19 new cases for the next two weeks.
Figure 4. Prediction of the prevalence in week 10 and 11.

Figure 1. Plot of COVID-19 prevalence over days.

CONCLUSIONS AND RECOMMENDATIONS FOR PRIORITIZING INVESTMENTS TO MITIGATE PANDEMIC RISK IN RESOURCE-LIMITED SURROUNDINGS

Preparing for a pandemic is challenging because of a multitude of factors, many of which are unique among natural disasters. Pandemics are rare events, and the risk of occurrence is influenced by anthropogenic changes in the natural environment. In addition, accountability for preparedness is diffuse, and many of the countries at greatest risk have the most limited capacity to manage and mitigate pandemic risk.

Unlike most other natural disasters, pandemics do not remain geographically contained, and damages can be mitigated significantly through prompt intervention. As a result, there are strong ethical and global health imperatives for building capacity to detect and respond to pandemic threats, particularly in countries with weak preparedness and high spark and spread risk.

Investments to improve pandemic preparedness may have fewer immediate benefits, particularly relative to other pressing health needs in countries with heavy burdens of endemic disease. Therefore, characterizing pandemic risk and identifying gaps in pandemic preparedness are essential for prioritizing and targeting capacity-building efforts. Thinking about risks in terms of frequency and severity, notably using probabilistic modeling and EP curves, can quantify the potential pandemic risks facing each country and clarify the benefit-cost case for investing in pandemic preparedness.

No single, optimal response to a public health emergency exists; strategies must be tailored to the local context and to the severity and type of pandemic. However, overarching lessons emerge after multiple regional epidemics and global pandemics. For example, because of their high spark and spread risks, many LMICs would benefit most from building situational awareness and health care coordination capacity; public health response measures are far more cost-effective if they are initiated quickly and if scarce resources are targeted appropriately.

Building pandemic situational awareness is complex, requiring coordination across bureaucracies, across the public and private sectors, and across disciplines with different training and different norms (including epidemiology, clinical medicine, logistics, and disaster response). However, an appropriately sized and trained health workforce (encompassing doctors, nurses, epidemiologists, veterinarians, laboratorians, and others) that is supported by adequate coordination systems is a fundamental need—the World Health Organization has recommended a basic threshold of 23 skilled health professionals per 10,000 people (WHO 2013a).

Increasing the trained health workforce also will increase the capacity to detect whether any particular population (for example, human, farm animal, or wildlife) is suffering from a pathogen with high pandemic risk. Increasing the health workforce also will improve the overall resiliency of the health system, an improvement that can be applied to any emergency that results in morbidity and mortality shocks.

Additionally, building situational awareness will require sustained investment in infectious disease surveillance, crisis management, and risk communications systems. Investments in these capacities are likely to surge after pandemic or epidemic events and then abate as other priorities emerge. Hence, stable investment to build sustained capacity is critical.

Risk transfer mechanisms such as catastrophe risk pools offer a viable strategy for countries to manage pandemic risk. Further developing these mechanisms will allow countries to offload portions of pandemic risk and response that are beyond their immediate budgetary capacity. For this reason, risk transfer solutions should be designed with the needs and constraints of LMICs in mind. However, countries must have predefined contingency and response plans as well as the absorptive capacity to use the emergency financing offered by such solutions. Broad and effective use of pandemic insurance will require parallel investments in capacity building and emergency response planning.

Finally, researchers must address the significant knowledge gaps that exist regarding LMICs' pandemic preparedness and response. Improving the tracking of spending and aid flows specifically tied to pandemic prevention and preparedness is vital to tracking gaps and calibrating aid flows for maximum efficiency. Systematic data on response costs in low-income settings are scarce, including data regarding spending on clinical facilities, supplies, human resources, and response activities such as quarantines. Bridging these data gaps can improve pandemic preparedness planning and response through evidence-based decision-making and support efforts to prevent and mitigate epidemics and pandemics.

Summary of probable SARS cases with onset of illness from 1 November 2002 to 31 July 2003

Areas	Cumulative number of cases: Female	Male	Total	Median age (range)	Number of deaths^a	Case fatality ratio (%)	Number of imported cases (%)	Number of HCW affected (%)	Date onset first probable case	Date onset last probable case
Australia	4	2	6	15 (1-45)	0	0	6 (100)	0 (0)	26-Feb-03	1-Apr-03
Canada	151	100	251	49 (1-98)	43	17	5 (2)	109 (43)	23-Feb-03	12-Jun-03
China	2674	2607	5327^b	Not available	349	7	Not Applicable	1002 (19)	16-Nov-02	3-Jun-03
China, Hong Kong Special Administrative Region	977	778	1755	40 (0-100)	299	17	Not Applicable	386 (22)	15-Feb-03	31-May-03
China, Macao Special Administrative Region	0	1	1	28	0	0	1 (100)	0 (0)	5-May-03	5-May-03
China, Taiwan	218	128	346^c	42 (0-93)	37	11	21 (6)	68 (20)	25-Feb-03	15-Jun-03
France	1	6	7	49 (26 - 61)	1	14	7 (100)	2 (29)^d	21-Mar-03	3-May-03
Germany	4	5	9	44 (4-73)	0	0	9 (100)	1 (11)	9-Mar-03	6-May-03
India	0	3	3	25 (25-30)	0	0	3 (100)	0 (0)	25-Apr-03	6-May-03
Indonesia	0	2	2	56 (47-65)	0	0	2 (100)	0 (0)	6-Apr-03	17-Apr-03
Italy	1	3	4	30.5 (25-54)	0	0	4 (100)	0 (0)	12-Mar-03	20-Apr-03
Kuwait	1	0	1	50	0	0	1 (100)	0 (0)	9-Apr-03	9-Apr-03
Malaysia	1	4	5	30 (26-84)	2	40	5 (100)	0 (0)	14-Mar-03	22-Apr-03
Mongolia	8	1	9	32 (17-63)	0	0	8 (89)	0 (0)	31-Mar-03	6-May-03
New Zealand	1	0	1	67	0	0	1 (100)	0 (0)	20-Apr-03	20-Apr-03
Philippines	8	6	14	41 (29-73)	2	14	7 (50)	4 (29)	25-Feb-03	5-May-03

Republic of Ireland	0	1	1	56	0	0	1 (100)	0 (0)	27-Feb-03	27-Feb-03
Republic of Korea	0	3	3	40 (20-80)	0	0	3 (100)	0 (0)	25-Apr-03	10-May-03
Romania	0	1	1	52	0	0	1 (100)	0 (0)	19-Mar-03	19-Mar-03
Russian Federation	0	1	1	25	0	0	Not Available	0 (0)	5-May-03	5-May-03
Singapore	161	77	238	35 (1-90)	33	14	8 (3)	97 (41)	25-Feb-03	5-May-03
South Africa	0	1	1	62	1	100	1 (100)	0 (0)	3-Apr-03	3-Apr-03
Spain	0	1	1	33	0	0	1 (100)	0 (0)	26-Mar-03	26-Mar-03
Sweden	3	2	5	43 (33-55)	0	0	5 (100)	0 (0)	28-Mar-03	23-Apr-03
Switzerland	0	1	1	35	0	0	1 (100)	0 (0)	9-Mar-03	9-Mar-03
Thailand	5	4	9	42 (2-79)	2	22	9 (100)	1 (11)^d	11-Mar-03	27-May-03
United Kingdom	2	2	4	59 (28-74)	0	0	4 (100)	0 (0)	1-Mar-03	1-Apr-03
United States	13	14	27	36 (0-83)	0	0	27 (100)	0 (0)	24-Feb-03	13-Jul-03^e
Viet Nam	39	24	63	43 (20-76)	5	8	1 (2)	36 (57)	23-Feb-03	14-Apr-03
Total			8096		774	9.6	142	1706		

a. Includes only cases whose death is attributed to SARS.

b. Case classification by sex is unknown for 46 cases.

c. Since 11 July 2003, 325 cases have been discarded in Taiwan, China. Laboratory information was insufficient or incomplete for 135 discarded cases, of wh

d. Includes HCWs who acquired illness in other areas.

e. Due to differences in case definitions, the United States has reported probable cases of SARS with onsets of illness after 5 July 2003.

Figure L.2A Risk of SARS transmission to health-care workers exposed to tracheal intubation

Review: Aerosol Generating Procedures
Comparison: 02 Tracheal intubation
Outcome: 01 Exposed versus unexposed

Study or sub-category	Exposed n/N	Unexposed n/N	OR (random) 95% CI	Weight %	OR (random) 95% CI
Scales (2003)	3/5	3/14		16.86	5.50 [0.61, 49.54]
Fowler (2004)	6/14	2/62		22.81	22.50 [3.86, 131.06]
Loeb (2004)	3/4	5/28		14.23	13.80 [1.18, 161.71]
Raboud (2010)	12/144	14/480		46.10	3.03 [1.37, 6.70]
Total (95% CI)	167	584		100.00	6.56 [2.28, 18.88]

Total events: 24 (Exposed), 24 (Unexposed)
Test for heterogeneity: Chi² = 4.97, df = 3 (P = 0.17), I² = 39.6%
Test for overall effect: Z = 3.49 (P = 0.0005)

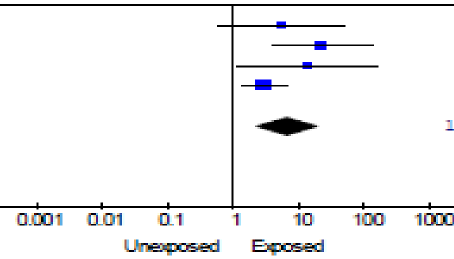
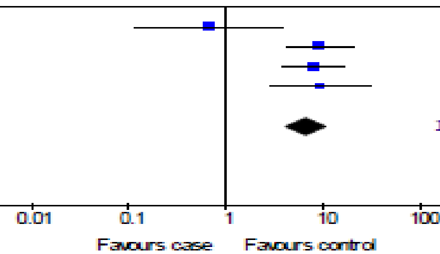


Figure L.2B Tracheal intubation as risk factor for SARS transmission

Review: Aerosol Generating Procedures
Comparison: 02 Tracheal intubation
Outcome: 02 Cases versus controls

Study or sub-category	Case n/N	Control n/N	OR (fixed) 95% CI	Weight %	OR (fixed) 95% CI
Teleman (2004)	2/36	4/50		26.73	0.68 [0.12, 3.91]
Pei (2006)	28/120	9/281		34.90	9.20 [4.19, 20.21]
Chen (2009)	16/91	17/657		28.81	8.03 [3.90, 16.56]
Liu (2009)	6/12	45/465		9.57	9.33 [2.89, 30.15]
Total (95% CI)	259	1453		100.00	6.60 [4.12, 10.55]

Total events: 52 (Case), 75 (Control)
Test for heterogeneity: Chi² = 7.78, df = 3 (P = 0.05), I² = 61.4%
Test for overall effect: Z = 7.87 (P < 0.00001)



CI, confidence interval; n, number of events; N, sample size; OR, odds ratio; SARS, severe acute respiratory syndrome

Abbreviations and acronyms

LMICs = low- and middle-income countries

EP = exceedance probability

SARS = severe acute respiratory syndrome

REFERENCES

1. World Health Organization (WHO). Natural ventilation for infection control in health care settings. Geneva, WHO, 2009 (http://whqlibdoc.who.int/publications/2009/9789241547857_eng.pdf).
2. American Society of Heating Refrigerating and Air-Conditioning Engineers, American National Standards Institute, American Society for Healthcare Engineering. Ventilation of health care facilities: ASHRAE/ASHE Standard. Atlanta, GA, American Society of Heating, Refrigerating, and Air-Conditioning Engineers, 2008.
3. A. I. A. Academy of Architecture for Health. Facilities Guidelines Institute (United States). Department of Health and Human Services. Guidelines for design and construction of hospital and health care facilities. Washington, D.C., American Institute of Architects, 2001.
4. Mayhall CG. Hospital epidemiology and infection control. Philadelphia, Lippincott Williams & Wilkins, 2004.
5. Wenzel RP. Prevention and control of nosocomial infections. Philadelphia, Lippincott Williams & Wilkins, 2003
6. World Health Organization (WHO). International health regulations (2005). Geneva, WHO, 2006 (http://www.who.int/csr/ihr/en).
7. Centers for Disease Control and Prevention (CDC). Guidelines for preventing the transmission of Mycobacterium tuberculosis in health-care settings, 2005. Morbidity and Mortality Weekly Report, 2005, 54(RR17).
8. Centers for Disease Control and Prevention (CDC). Guidelines for environmental infection control in health-care facilities. Morbidity and Mortality Weekly Report, 2003, 52(RR–10).
9. Roy CJ, Milton DK. Airborne transmission of communicable infection – the elusive pathway. New England Journal of Medicine, 2004, 350(17):1710–1712.
10. WHO Centre for Health Development (Kobe Japan). A glossary of terms for community health care and services for older persons. Kobe, Japan, World Health Organization Centre for Health Development, 2004 (http://whqlibdoc.who.int/wkc/2004/WHO_WKC_Tech.Ser._04.2.pdf).
11. Stedman TL. Stedman's medical dictionary. Philadelphia, Lippincott Williams & Wilkins, 2006.
12. Brankston G, Gitterman L, Hirji Z et al. Transmission of influenza A in human beings. Lancet Infectious Diseases, 2007, 7(4):257–265 (http://www.ncbi.nlm.nih.gov/entrez/query.fcgi?cmd=Retrieve&db=PubMed&dopt=Citation&list_uids=17376383).
13. Last J. A dictionary of epidemiology. New York, Oxford University Press, 2001.
14. World Health Organization (WHO). Guidelines for the use of seasonal influenza vaccine in humans at risk of H5N1 infection. Geneva, WHO, 2004 (http://www.who.int/influenza/human_animal_interface/virology_laboratories_and_vaccines/vaccine_use_h5N1_riskgroups/en).
15. World Health Organization (WHO). WHO guidelines for pharmacological management of pandemic (H1N1) 2009 influenza and other influenza viruses. Geneva, WHO, 2010 (http://www.who.int/csr/resources/publications/swineflu/h1n1_use_antivirals_20090820/en/index.html)