Article title: Application of Inherent Risk of Contagion (IRC) framework and modelling to aid local Covid-19 response and mitigation
Authors: Helmi Zakariah[1], Fadzilah bt Kamaluddin[2], Choo-Yee Ting[3], Hui-Jia Yee[4], Shereen Allaham[5], Logan Manikam[6]
Affiliations: 1. AI4Good, Kuala Lumpur, Malaysia 2. Artificial Intelligence for Medical Epidemiology (AIME), Kuala Lumpur, Malaysia[1], N/A[2], Faculty of Computing and Informatics, Multimedia University, Cyberjaya, Malaysia[3], 4. Department of Epidemiology and Public Health, University College London Institute of Epidemiology and Health Care, London, UK 5. Aceso Global Health Consultants Limited, London, UK[4]
Orcid ids: 0000-0001-5288-3325[6]
Contact e-mail: logan.manikam.10@ucl.ac.uk

Preprint statement: This article is a preprint and has not been peer-reviewed, under consideration and submitted to UCL Open: Environment Preprint for open peer review.
Links to data: Helmi Zakariah
Funder: Selangor State Government
DOI: 10.14324/111.444/000053.v1
Preprint first posted online: 20 July 2020
Keywords: COVID-19, Risk of Contagion, Machine Learning, Risk Ranking Area, Georeference, Environmental modelling, Sustainable development, Sustainability, Sanitation, health, and the environment, Sustainable and resilient cities

Dear Editorial Team,

We are pleased to submit our manuscript titled “Application of Inherent Risk of Contagion (IRC) framework and modelling to aid local Covid-19 response and mitigation”. We believe that the manuscript is appropriate for publication in UCL Open: Environment megajournal as it represents an important consideration of a new analytical framework to measure the inherent risk of contagion (IRC) and provide georeferenced data to protect the health system, aid contact tracing, and prioritise the vulnerable in the community during the COVID-19 response. This paper will represent an aspect of environment-related research.

This paper builds upon the machine-learning-powered IRC analytical framework that, through the geo-referencing of COVID-19 cases in the affected region, is able to provide support to operational platforms from which response and mitigation activities can be planned and executed in Malaysia. The machine-learning-powered IRC analytical framework is now operational in Selangor, Malaysia as part of their COVID-19 local plan, and could be validated for other high and low- middle-income countries to aid government response and promote date-guided decision-making.

In particular, this paper will provide information about a framework that currently used in Malaysia to identify the district critical risk area, the number of vulnerable and their distribution within specific area and then estimate the resources. Thereafter, a cross-check with the pension fund and welfare fund databases will be employed to identify contact details of the vulnerable population. Moreover, the same risk-ranking will be utilised to plan aid and food supplies distribution if the government decided to put the critical risk zone under total lockdown.

This manuscript has not been published, nor under consideration for publication in another journal. All authors have approved the manuscript for submission. We have no conflict of interest to interest to declare.

We thank you for your consideration.

Your Faithfully,
Dr Logan Manikam

Application of Inherent Risk of Contagion (IRC) framework and modelling to aid local Covid-19 response and mitigation

Dr Helmi Zakariah [1,2]; Dr Fadzilah bt Kamaluddin; Dr Choo-Yee Ting[3]; Hui-Jia Yee [1,2]; Shereen Allaham[4,5]; Dr Logan Manikam[4,5]

Institutions

1. AI4Good, Kuala Lumpur, Malaysia
2. Artificial Intelligence for Medical Epidemiology (AIME), Kuala Lumpur, Malaysia
3. Faculty of Computing and Informatics, Multimedia University, Cyberjaya, Malaysia
4. Department of Epidemiology and Public Health, University College London Institute of Epidemiology and Health Care, London, UK
5. Aceso Global Health Consultants Limited, London, UK

Corresponding author: Logan Manikam
Postal address: 3, Abbey Terrace, London, SE2 9EY
Email: logan.manikam.10@ucl.ac.uk
Tel: +447868721152

Abstract

The current outbreak of coronavirus disease 2019 (COVID-19) caused by the novel coronavirus named SARS-CoV-2 represents a major global public health problem threatening many countries and territories. Mathematical modelling is one of the non-pharmaceutical public health measures that has the potential to play a crucial role for mitigating the risk and impact of the pandemic. A group of researchers and epidemiologists have developed a machine learning-powered inherent risk of contagion (IRC) analytical framework that, through the geo-referencing of COVID-19 cases in a particular region, is able to provide support to operational platforms from which response and mitigation activities can be planned and executed. This framework dataset provides a coherent picture to track and predict the COVID-19 epidemic post lockdown by piecing together preliminary data on publicly available health statistic metrics alongside the area of reported cases, drivers, vulnerable population, and number of premises that are suspected to become a transmission area between drivers and vulnerable population. The main aim of this new analytical framework is to measure the IRC and provide georeferenced data to protect the health system, aid contact tracing, and prioritise the vulnerable.

Keywords: COVID-19; Risk of Contagion; Machine Learning; Risk Ranking Area; Georeference

Introduction

The current outbreak of Corona Virus Disease 2019 (COVID-19) caused by the novel coronavirus named SARS-CoV-2 represents a major global public health problem threatening many countries and territories[1]. As this pandemic continues to expand, governments worldwide have implemented tight preventive measures, including partial or complete lockdowns with the exception of essential services (e.g. hospitals, clinics, convenience stores that provide daily necessities) that aim to limit virus transmission.

One of the main questions during the COVID-19 lockdown is, what non-pharmaceutical public health (NPH) measures should be applied (e.g. targeted community-based screening and testing), in order to ensure a safe and gradual route back to normality (e.g. how to kickstart and allow crucial business to resume their operation) particularly in light of COVID-19's transmissibility and high asymptomatic infectivity periods.

Overview of the Method Used to Identify and Assess Inherent Risk of Contagion

Mathematical modelling plays a substantial role in answering these uncertainties whilst allowing for differing transmission periods and geographical areas[2,3] that takes into account; (i) population density, drivers of spread (e.g. factors favouring COVID-19 spread, including: climate, regional or cross-boarder travel, transport hubs, number of closely neighbouring areas) and vulnerable population who are at increased risk of suffering complications of the infection, such as older adults and/or anyone with a serious underlying health condition, (ii) precise geographical location of reported cases, and (iii) the number of premises that remain in operation during the lockdown, that could contribute as transmission sites[4]. Inherent risk of contagion (IRC) is known as the average accelerated degree of the spread of COVID-19 cases in a geographical unit (geo-unit), province and municipality[5]. Previous research aptly describes this approach as quantifying IRC[5].

A group of researchers and epidemiologists at AI4Good and AIME[6] have developed a machine-learning-powered IRC analytical framework that, through the geo-referencing of COVID-19 cases in the affected region, is able to provide support to operational platforms from which response and mitigation activities can be planned and executed in Malaysia. The main objective of this new analytical framework, however, is to not only track the current SARS-COV-2 virus, but to also use the calculated IRC to provide a prediction of how the infection could evolve in both the affected area and neighbouring regions in the near future. Preliminary data on publicly available health statistic metrics, including mortality and recovery rate, combined with the area of reported cases, drivers of spread, mapping of vulnerable population, and number of premises that are suspected to become a transmission site form the basis of this framework. Using this input, the framework dataset can provide a coherent picture to track and predict the COVID-19 epidemic post lockdown by a) creating an

area risk-ranking, b) conducting a vulnerability analysis, reaching out to those most in need as well as c) automate contact tracing through its collaboration with a telecommunication provider and Mobile App operator. The IRC calculation thus be used to make data-driven decisions on the scale of lockdown measures, organize the necessary resources (e.g. testing kits) , guide community screening practices and facilitate contact tracing. This may ultimately not only relieve pressure from the health care system but also allow for planning of welfare and aid distribution particularly to endangered regions and the vulnerable population (e.g. elderly, women of childbearing age and children exposed to hazard) living in them.
The dataset uses a similarity index to create a risk ranking area system by locality and district; which highlights areas in red, yellow, orange, and green (Figure 1) to indicate the relative risk of infection and spread in these areas. The risk was calculated using the following parameters; cases incidence rate, cases location, vulnerable population spread around the cases location, incubation period and serial interval time of infection, number and location of premises still operating under lockdown, and character similarity between neighbourhoods. As shown in Figure 1, by selecting the district of interest, in which we selected "Gombak" a district located in the state of Selangor in Malaysia, where this framework is currently being operationalised, all relevant residential areas within the district and their respective risk rating can be analysed. The figure shows that despite having a high-risk rating for "Gombak", not all residential areas in that district area are high-risk.

Figure 1 The IRC risk ranking by locality and district. Green zone (average risk), Yellow zone (heightened risk), Orange zone (high risk), Red zone (critical risk).

After identifying the district critical risk area, we identify the number of vulnerable and their distribution within the area and then estimate the resources (e.g. test-kit, medical personnel, number of testing tent/centre need to set-up). Thereafter, a cross-check with the pension fund and welfare fund databases will be employed to identify contact details of the vulnerable population. Moreover, the same risk-ranking will be utilised to plan aid and food supplies distribution if the government decided to put the critical risk zone under total lockdown.

Conclusion

The machine-learning-powered IRC analytical framework is now operational in Selangor, Malaysia as part of their COVID-19 local plan, and could be validated for other high and low-middle-income countries to aid government response and promote date-guided decision-making. We welcome such collaborations.

Acknowledgement

The authors would like to thank the Selangor State Government, Malaysia for funding the development of this framework. We would like to thank all members of AI4Good and AIME who contribute to the development and implementation of this approach.

Declaration of Interests

The authors declare no conflicts of interest with this work.

Ethics/Consent

N/A

References

1. Rao, A.S.S. and J.A. Vazquez, Identification of COVID-19 can be quicker through artificial intelligence framework using a mobile phone-based survey in the populations when cities/towns are under quarantine. Infection Control and Hospital Epidemiology, 2020: p. 1-18.
2. Peng, L., Yang, W., Zhang, D., Zhuge, C., and Hong, L. (2020). Epidemic analysis of COVID-19 in China by dynamical modeling. arXiv preprint arXiv:2002.06563.
3. Dye, C. and N. Gay, Modeling the SARS epidemic. Science, 2003. **300**(5627): p. 1884-1885.
4. Mossong, J., Hens, N., Jit, M., Beutels, P., Auranen, K., Mikolajczyk, R., ... and Heijne, J. (2008). Social contacts and mixing patterns relevant to the spread of infectious diseases. PLoS medicine, 5(3).
5. Ronsivalle, Gaetano Bruno and Foresti, Luca and Poledda, Gianfranco. (2020). A PROTOTYPE MODEL OF GEOREFERENCING THE INHERENT RISK OF CONTAGION FROM COVID-19. 10.13140/RG.2.2.30077.72163.
6. Dow Jones. 2020. A New Dawn of Artificial Intelligence - Preventing Epidemics - Dow Jones. [online] Available at: https://www.dowjones.com/predictive-modeling-to-prevent-epidemics/ [Accessed 09 April 2020].