2020 19th International Symposium on Distributed Computing and Applications for Business Engineering and Science (DCABES)

Prevention and Control of Emerging Infectious Diseases in Human Populations

Sophie Khaddaj
Artemedis International, London, EC 14 8PD, UK
sophiekhaddaj@gmail.com

Hussain Chrief
Oxford University Hospital, Oxford, HA0 1QW, UK
h.chreif@nhs.net

Abstract— **The prevention, control and prediction of emerging infectious diseases are vital in order to effectively manage their spread and impact. Over the years many modelling techniques have been developed for the management of infectious diseases. However, emerging diseases are linked to selective pressures caused by humans, for example environmental pressure such as urbanisation and habitat fragmentation. In this paper we present a new approach, which combines human behavioural factors together with advanced mathematical modelling and machine learning, for preventing, monitoring and predicting future epidemics. This will help medical professionals and policy makers to optimize, in real-time, response efforts to major outbreaks.**

Keywords: Selective Pressures, Emerging Infectious Diseases, Human Factors, Mathematical Modelling, Machine Learning.

I. Introduction

During the past decades there have been major scientific and technological advances both in biological and medical science, and in information and communication technology. These have impacted many branches of healthcare including disease control and management, particularly the management of infectious diseases. Such advances present a major opportunity for both original research and real-life applications to improve infectious disease management. These developments can considerably contribute to detecting and tracking the spread of infectious diseases. This is a global challenge, with severe consequences, particularly as the impact and the speed of infection spreading, and the potential for outbreaks have increased dramatically in the past few years [1-3]. This has strongly manifested itself with the current outbreak of coronavirus (COVID-19). As we have seen, such outbreaks can affect the human population on an unprecedented scale and across many geographical locations [1]. Therefore, managing infectious diseases in real time in a global world, is becoming crucial to avoid the spread of infections, which require new infectious diseases management approaches.

The rapid increase in the availability and reliability of health information systems, provide support for many communities, particularly during pandemics. These include health professionals, particularly in improving diagnosis provisioning, policy makers and managers in their resource allocation and scheduling, and patients in the sharing of health information and the encouragement of self-care [4]. However, health information systems rely heavily on the efficient use of electronic medical records, and electronic health data stored in various formats, such as texts, images etc. The stored data contains a wealth of information that can help not only in improving individual patient care, but it can also be used in many aspects of public health particularly in research; for example, in identifying patterns and potential associations between the different stored entities [5]. Thus, health data can provide a good source for the detection and tracking of infectious and non-infectious diseases, which presents a significant step in public healthcare management and improvement in decision-making processes.

Although modern technology can assist in devising new management solutions, it is still the case that in many situations infection management systems rely extensively on manually collected data. This has serious implications on real time monitoring, tracking and predicting the spread of infectious diseases due to the slowness and cost associated with manual systems. The ability to promptly and accurately predict and track infectious diseases is extremely important. Thus, there has been a number of works that apply data mining and machine learning algorithms in an attempt to accurately predict the spread of diseases [6-7]. However, new sources of electronic health data are emerging which include social media, Internet, mobile phones, satellite, remote sensing and radio-frequency sensors.

In addition, to the technological advances and complex mathematical modelling, direct human factors and actions contribute enormously to the spread of infectious diseases. Humans' modern lifestyle has diverse and far-reaching influences on the evolution of pathogens, such as viruses and bacteria. It has created considerable pressures on our environment through massive urbanisation, globalisation, extensive travels, energy generation, food production and intensive agriculture, which contribute to climate change and pollution, among other things.

However, the management of emerging infectious diseases requires novel solutions that are based on integrating human behaviour with advanced predictive modelling techniques in order to address and mitigate the impact of infectious diseases. Therefore, in this work an integrated approach is proposed that considers both behavioural science and advanced analytics for outbreak mitigation, detection and risk assessment. The aim is to prevent future infectious disease threats, and to predict, control and quantify their impact and risks at local and global levels.

We start with a discussion on the evolution of infectious diseases, and the commonly used techniques for their detection and management. Then, aspects of the complex human–nature relationship including human behaviour and environmental issues and factors are

2473-3636/20/$31.00 ©2020 IEEE
DOI 10.1109/DCABES50732.2020.00092

presented. These are addressed by proposing an approach that combines human behavioural and lifestyle factors with long-term risk predictions and impact analysis of human induced pressures, which are important to our lives, and even to our existence. We conclude with some suggestions for future work.

II. Infectious Diseases Management

Emerging infectious disease outbreaks rank among the key challenges facing humanity as they have the potential to be more destructive than many other natural disasters such as earthquakes and hurricanes. In the age of globalisation, fast communication and transport, and high human mobility, infectious diseases can spread globally and at an unprecedented pace. Therefore, they can lead to major epidemics or pandemic outbreaks, as we have seen with COVID-19, which have potentially devastating consequences. Historically, infectious diseases are major causes of global death causing around 22% of deaths, and it is higher in the African region, where 56% of deaths are caused by infectious diseases (along with maternal, neonatal and nutritional conditions) [8]. Over the years, the World Health Organization (WHO) detected and verified a large number of outbreaks of potential international concern [8]. Therefore, it is vital to be able to effectively control infectious diseases and manage their spread and impact on communities.

However, the control and management of infectious disease outbreaks relied on traditional surveillance systems, which are not suitable for real time monitoring and prediction as they are slow, expensive and in many cases they lack accuracy. But, recent technological advancement particularly in health informatics can assist in managing the control of outbreaks in a timely and reliable manner. Thus, there have been a number of projects which deployed information technology for early detection and prediction of outbreaks, based on data collection and analysis which can improve the management and monitoring of infectious diseases [9-10].

In fact, the number of studies, that have considered the application and use of information technology for surveillance monitoring and predictive analysis of infectious diseases, have accelerated in the past few years, partially influenced by the recent outbreaks, and partially due the fast technological advancement, in web, mobile phone, social networks etc. The need for the improvement of surveillance systems and the role of health informatics have been considered by many research projects. These include the improvement in data gathering and analysis, the design of new systems for surveillance monitoring, and new methods and techniques for improving the accuracy of prediction and monitoring [10-11].

Moreover, the role and evaluation of new technologies such as the web for infectious disease surveillance have become essential. In fact, web-based tools have become indispensible in infectious disease surveillance [12], which when combined with outbreak intelligent analysis have provided valuable information to many health institutions both locally and on a global scale. In addition, social media has emerged as a potential source for harvesting health data, and applying data mining techniques in outbreak management, which has been discussed in [6-7]. Social media data have been shown to accurately track disease spread after major outbreaks and it can be used to complement both new and more traditional monitoring techniques.

These technological advances have resulted in many changes not only in data acquisition and storage, but also in data analysis and the delivery of knowledge. Data analysis is at the core of any infectious disease management system, as it can be used to identify implicit and previously unknown patterns, which can help in the prediction processes. Data analysis techniques range from simple query and reporting services, to statistical analysis and data mining.

Data mining is used for discovering patterns in data sets, which involves using many methods and algorithms from machine learning. The set of algorithms includes clustering, classifiers, rule association etc., which can be used for data classification so that the data elements are put into groups. This offers a better understanding and helps in creating new knowledge for decision-making processes, which can guide policy makers and health professionals. Data mining can also be used to give information about available data, in a descriptive way. In addition, many visualisation techniques can be deployed for either data descriptive or predictive modes.

The application of data mining techniques in health informatics has been considered in a number of works, including the monitoring of infectious diseases and detecting early outbreaks [13]. Statistical methods for the detection outbreaks of infectious diseases have be reviewed in [14] together with their applicability to surveillance data. In addition, a number of prediction models have been developed using electronic record data as well as other available public and commercial databases [15-16].

The real time outbreak management requires a high degree of flexibility and accuracy with many challenges for both infrastructure and data mining techniques in terms of scalability and correctness. The other challenge is associated with response time to the target audience, which requires speedy real time data gathering, processing, forecasting and visualisation, i.e. it is also a performance issue.

Overall, there have been a number of initiatives and collaborative projects that have been trying to address many of the outbreak monitoring problems. Some of the reported work relies on gathering many media sources, both in electronic and hard copies, including internet, reports and blogs, which are then analysed by human experts using statistical analysis in order to produce alert [11-12]. Other projects introduced some elements of automation, which relies on RSS feed to improve surveillance with integrated web presence and reporting and other media sources to produce an alert system, and human intervention, if needed, in an automated system. More outbreak specific systems have also been developed with information made available as and when required.

Other specialised infectious diseases surveillance frameworks and systems, for public health, have been developed particularly in conjunction with well-managed health systems where electronic records are readily available. Surveillance tracking systems and pattern identification has also been considered in [6-7], [17] with elements of automation of outbreak detection processes and an alert system. With the availability of electronic data records, statistical analysis and data mining techniques have been applied for the prediction of potential outbreaks, with potentially good results [11]. However, the use of technologies and modelling techniques are not enough on their own, as they require an active response and actions not only from policy makers, but also from the general population. Thus, human behaviour and actions are extremely important to successfully manage infectious diseases. Without a large-scale involvement of the general population, it will be impossible to deal with future threats posed by infectious diseases.

III. Human Factors

The human factors and pressures on the environment and resources play important roles in creating selective pressures on pathogens which are responsible for the spread of many infectious diseases. The relationship of humans and their surrounding environment is characterised by a high level of complexity, multidimensionality and governed by multiple biophysical and socioeconomic factors, which exerts a significant impact on our environment [19]. These include urbanisation, globalisation, deforestation, agriculture and energy (Figure 1), all creating environments in which infectious diseases can spread rapidly as they adapt to new evolutionary pressures, thus creating global threats to even the existence of humans [20].

In fact, humans often act against their own long-term best interests. Thus we need to change our behaviour and lifestyles in order to reduce environmental pressures. Behavioural science can help by offering insights into how and why people behave as they do, using research from many fields such as psychology and economics. In order to alter human behaviour, new policies and learning strategies have to be developed and implemented so to obtain a sustainable solution. Such long-term strategies will be centered on giving people more information about their actions and their impact.

A core component of human behavioural and lifestyle changes is the learning process in order to fully understand the implication of human behaviours, actions and the required changes to mitigate their impacts. For example, we need to slow down massive urbanisation programs, stop deforestation and have sustainable food production without harmful chemicals and safety measures to avoid unnecessary infection. Furthermore, clean energy production and pollution reduction can contribute to a healthy living environment, thus, reducing potential infections.

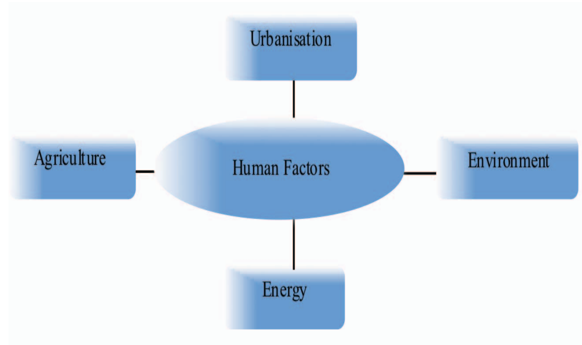

Figure 1. Human Pressures

IV. An Integrated Approach

In this section we present a new integrated approach for the management of infectious diseases that takes into account human behaviour as well as aspects of prediction and forecasting; thus helping in making intelligent decisions, not only on long-term prevention, but also on the prediction of potential outbreaks and the assessment of humans actions and their impact on the spread of infections. It enables the construction of infectious disease management systems with prediction and forecasting features; thus offering an integrated solution for targeting and containment, as well risk assessment and impact analysis.

The approach is based on two main components (Figure 2). Human pressure management requires changes in the attitudes and lifestyles of the population, particularly in terms of living environment, climate change, deforestation, urbanisation, pollution etc. This will need a comprehensive learning process about the impacts of the actions that we take, particularly in terms of heath and infectious disease prevention. This requires the intervention of policy makers and international organisations, and involves all living communities on a global scale.

More specifically, the approach aims to address some of the human induced selective pressures and attempt to balance them. This generic approach is based on the use of intelligent management of health, agriculture, energy, environment, with contributing factors such as transport, each addressing and managing one or more pressures. This can be achieved by emphasising the importance of evolving synergies between biophysical, socioeconomic and technological aspects. It involves aspects of direct human intervention through behavioural and living changes, and automation through the use of artificial intelligence [6-7], [17-18].

At the core of the approach is the use of the available massive computing power, which has been instrumental in genetic engineering, artificial intelligence, machine learning and advanced mathematical modelling are essential in the fight against infectious diseases. These can help in better understanding the evolutionary processes of pathogens, the infections they cause, and the prediction of their long-term impacts and risks.

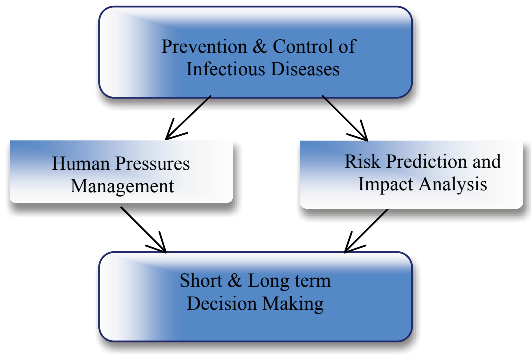

Figure 2. Management of Infectious Diseases

Modelling and simulation are essential as they can help in the prediction and assessment processes. These range from classical statistical models to more dynamic mathematical models and the use of artificial intelligence and data mining techniques (Figure 3). However, one of the foreseen problems in estimating the impact of our actions in terms of balanced selection processes is that minor shifts in initial conditions can have massive impacts, with a considerably disproportionate size in the future. In science there is a name for systems that operate like this - that is chaos. As traditional mathematical models are sensitive to initial conditions, machine learning is more appropriate as it might figure out how initial conditions from a very basic stage can affect the development of, and impact of environmental pressures and their patterns.

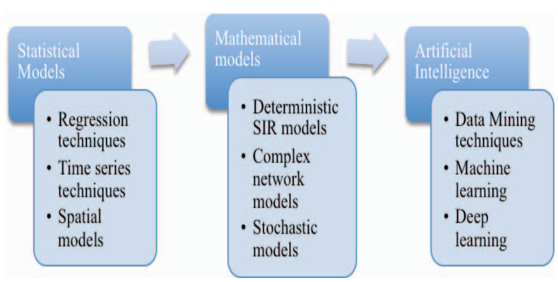

Figure 3. Analysis and Predictive Techniques

V. Conclusion

The fact is that despite our immense scientific achievements we continue to be highly vulnerable. This has been demonstrated by the outbreak of COVID-19, a highly adaptable virus found in us the perfect host, within a perfect environment, which we have created through massive urbanisation and globalisation. Therefore, we have investigated how human behaviour can be integrated with predictive solutions. The presented approach allows us to assess our influence on the evolution of pathogens, and what changes to our behaviour and lives we should make in order to reduce the potential threats of infectious diseases. In addition, by using modelling and simulation we can help in dynamically adapting our living conditions to mitigate any negative evolutionary influences on our lives. However, the quest for a comprehensive approach entails integrating and managing many techniques and methods ranging from behavioural analysis to machine learning, which have been touched briefly in this work and require considerable multidisciplinary research work.

References

[1] T. Constantino and T. Ugur, "Predicting COVID-19 Peaks Around the World," Frontiers in Physics , 8, 2020, pp. 1-6.

[2] P. Loubet, *et al.,* "Development of a Prediction Model for Ebola Virus Disease: A Retrospective Study in Nzérékoré Ebola Treatment Center, Guinea," The American journal of tropical medicine and hygiene, 95(6), 2016, pp. 1362-1367.

[3] C. Zanluca and C. dos Santos, "Zika virus – an overview," Microbes and Infection, 18(5), 2016, pp. 295-301.

[4] N. Archer et al., "Personal health records: a scoping review," Journal of the American Medical Informatics Association, 18(4), 2011, pp. 515–522.

[5] A. H. Barbara and J. M Drake, "Future directions in analytics for infectious disease intelligence," EMBO Rep., 2016, 17(6), pp. 785–789.

[6] G. S. Collins and K.G.M. Moons, "Reporting of artificial intelligence prediction models", The Lancet, 393(10181), 2019, pp. 1577-1579.

[7] Z. Zullaha, et al., "Applications of Artificial Intelligence and Machine learning in smart cities", Computer Communications, 154(15), 2020, pp. 313-323.

[8] World Health Organisation (WHO), Global Health Risks, https://www.who.int/healthinfo/global_burden_disease/GlobalHealthRisks_report_full.pdf, Accessed July 2020.

[9] C. G. Mónica, et al., "Computational Health Engineering Applied to Model Infectious Diseases and Antimicrobial Resistance Spread," Applied Sciences, 9(12), 2019, pp. 2486-2505.

[10] J. Li, *et al.,* An "Alarming and Prediction System for Infections Disease Based on Combined Models," APWeb 2016, Lecture Notes in Computer Science, 9932. Springer, Cham.

[11] B. Y. Reis, *et al*, "AEGIS: A Robust and Scalable Real-time Public Health Surveillance System", Am Med Inform Assoc, 2007,14(5), 2007, pp. 581-588.

[12] J. Choi, et al. "Web-based infectious disease surveillance systems and public health perspectives: a systematic review," BMC Public Health, 16, 2016.

[13] J. Li, *et al.,* "An Alarming and Prediction System for Infections Disease Based on Combined Models," Asia-Pacific Web Conference, 2016.

[14] S. Unkel, *et al.,* "Statistical methods for the prospective detection of infectious disease outbreaks: a review," Journal of the Royal Statistical Society, 175 (1), 2012, pp. 49-82.

[15] B. Michiels, *et al.*, "Influenza epidemic surveillance and prediction based on electronic health record data from an out-of-hours general practitioner cooperative" BMC Infect Dis. 17(84), 2017.

[16] C. D. Corley, et *al.,* "Disease Prediction Models and Operational Readiness," PLoS One, 9(3), 2014.

[17] M. J. Keeling and L. Danon, "Mathematical modelling of infectious diseases", British Medical Bulletin, 92, 2009, pp. 33–42.

[18] C.I Siettos and L. Russo, L. "Mathematical modelling of infectious disease dynamics", Virulence, 4(4), 2013, pp. 295–306.

[19] Seymour, V., 'The Human–Nature Relationship and Its Impact on Health: A Critical Review', Frontiers in Public Health, 4, 2016, pp. 260-276.

[20] A. P. Hendry, *et al.,* 'Human influences on evolution, and the ecological and societal consequences', Philosophical Transactions of The Royal Society B, Biological Sciences, 2017, pp. 1-13.