

Public Health and Epidemiology Informatics: Can Artificial Intelligence Help Future Global Challenges? An Overview of Antimicrobial Resistance and Impact of Climate Change in Disease Epidemiology

Alejandro Rodríguez-González^{1,2}, Massimiliano Zanin¹, Ernestina Menasalvas-Ruiz^{1,2}

¹ Centro de Tecnología Biomédica, Universidad Politécnica de Madrid, Madrid, Spain

² Escuela Técnica Superior de Ingenieros Informáticos, Universidad Politécnica de Madrid, Madrid, Spain

Summary

Objectives: To provide an overview of the current application of artificial intelligence (AI) in the field of public health and epidemiology, with a special focus on antimicrobial resistance and the impact of climate change in disease epidemiology. Both topics are of vital importance and were included in the “Ten threats to global health in 2019” report published by the World Health Organization.

Methods: We analysed publications that appeared in the last two years, between January 2017 and October 2018. Papers were searched using Google Scholar with the following keywords: public health, epidemiology, machine learning, data analytics, artificial intelligence, disease surveillance, climate change, antimicrobial resistance, and combinations thereof. Selected articles were organised by theme.

Results: In spite of a large interest in AI generated both within and outside the scientific community, and of the many opinions pointing towards the importance of a better use of data in public health, few papers have been published on the selected topics in the last two years. We identify several potential reasons, including the complexity of the integration of heterogeneous data, and the lack of sound and unbiased validation procedures.

Conclusions: As there is a better comprehension of AI and more funding available, artificial intelligence will become not only the centre of attention in informatics, but more importantly the source of innovative solutions for public health.

Keywords

Public health; drug resistance, microbial; climate change; artificial intelligence; big data

Yearb Med Inform 2019;224-31

<http://dx.doi.org/10.1055/s-0039-1677910>

1 Introduction

Artificial intelligence (AI), defined as a line of research at a workshop organized at Dartmouth College in 1956 [1], has been present in the scientific literature for many years, and has silently supported real-world applications spanning from recommendation systems, to fraud detection, and to video games. AI has nevertheless only recently hit the newspaper headlines, and only in the past few years has the general population become aware of concepts such as AI or machine learning (ML). New terms have also been coined, like big data analytics or data science, even though these build on top of ML. To further increase the confusion, in many cases these new terms are used interchangeably.

This increase in notoriety has also affected the healthcare domain. Although the use of AI and ML (as a subdomain of AI) in the healthcare sector is not new [2-6], the number of related research works has substantially increased during the last years [7-14]. In addition, the AI trend is not going to slow down soon, at least judging by the large number of papers discussing needed technological capabilities, limitations [15], and potential tasks where AI can provide better insights [16]. One question naturally arises: is this phenomenon just hype, or can it really be the foundation of a future healthcare?

On one hand, the massive use of terms like AI, ML, big data and data science can suggest hype; such interest may be driven not just from the recent advances in the field, but also from speculative and exaggerated news [17, 18] over the terms, results, and applications. This is nevertheless not an unusual dynamic, as can be seen in the Hype Cycle¹ analysis developed every year by the consultant company Gartner. Still, today's hypes may be the foundation of tomorrow's technologies. Thus, on the other hand, it is undeniable that AI and ML can provide a different point of view from which well-known problems can be tackled. Focusing on the manuscript theme, i.e., epidemiology and public health, scientists agree that it is becoming increasingly necessary to forge a partnership between the different disciplines of computer science and epidemiology, arguing that the latter can deeply benefit from the recent advances in the former [19]. Technology can then help merge different data sources together, with the emergence of unexpected synergies. This may be the case, for instance, between large -omics data sets (being them genomics, proteomics, transcriptomics, and so forth), clinical data produced in the hospitals and stored in electronic health records, and publicly

¹ <https://www.gartner.com/en/research/methodologies/gartner-hype-cycle>

available datasets with information about risk factors at both individual and macro environmental levels.

Merging data sources and creating extensive data sets are nevertheless only necessary conditions, not sufficient ones. As the authors of [14] highlighted, what is needed are capabilities to analyse and manipulate those data. The authors identified data mining (DM), which is mainly based on the application of ML techniques, as one of the key ingredients that will fuel future developments in the field. Although ML builds on statistics, it goes one step further. Where classical statistics mainly focuses on a descriptive level, ML can provide deeper insights in the analysis of the data and the creation of very different types of forecast models [20, 21]. For this reason, much interest has emerged in the epidemiology and public health areas. In [14], several examples that demonstrate the benefits of applying ML are discussed, but the validation and interpretation of the findings are also identified as major limitations.

If DM for healthcare is not just hype, what steps have been taken in the last years, and what barriers are still to be overcome? The aim of this survey is to provide a view on the main opportunities, challenges, and practical implications. It is worth noting that reviews on epidemiology and public health within the AI/big data/data analytics areas are not in short supply, and that the previous two editions of the IMIA Yearbook have dealt with this specific subject [22, 23], providing a very interesting perspective of the worldwide efforts that have been done in the field in the last years. Therefore, and for the sake of providing a different perspective, we have decided to focus this paper on two specific areas which fall within the scope of public health and epidemiology, but, to the best of our knowledge, have not been previously addressed in depth. We choose this two areas among those stated by the World Health Organization (WHO) in its “Ten threats to global health in 2019” report²: artificial intelligence and big data

approaches applied to antimicrobial resistance (AMR), and the impact of climate change in disease epidemiology.

The rationale behind such selection includes the importance that both topics had in the last years, and especially will have in the near future. On one hand, if twenty years ago a death caused by a bacterial infection in the developed world was something unheard of, things have already started to change. If ways of reversing, or at least limiting, AMR are not soon found, a simple infected cut may lead to tragic consequences – a scenario called the “Post-Antibiotic Era”. On the other hand, climate change is a worldwide topic of political and scientific debate; yet, much less well known is that recent research works have shown that climate change is clearly influencing diseases, disease prevalence, and disease spreading. In both cases, the more challenging and complex techniques provided by AI/ML can yield solutions hitherto not explored, and hence provide interesting new research areas.

2 Methodology

The present review has been based on the analysis of scientific papers focusing on antimicrobial resistance and climate change in disease epidemiology, and that included contents about machine learning, data analytics, or artificial intelligence methods. Only publications that appeared between January 2017 and October 2018 have been considered. Papers were searched through Google Scholar, with keywords like “public health, epidemiology, machine learning, data analytics, artificial intelligence, disease surveillance, climate change, antimicrobial resistance”, and combinations thereof. Note that Google Scholar was chosen over other options, like PubMed, due to larger coverage.

A final manual curation has been carried out, firstly by selecting papers that seemed more appropriate, further looking for additional references in the bibliography of the same papers, and in the list of those papers that were citing them.

The final selection yielded a list of 40 papers, 27 for antimicrobial resistance and 13 for climate change.

3 Results

3.1 Antimicrobial Resistance

Since the discovery in 1928 of the first natural antimicrobial fungus by Alexander Fleming, the world has moved into the “Antimicrobial Era”. Antimicrobial therapies and agents allow preventing and blocking the transmission, and curing most infections. Not surprising, this led to an optimistic environment in the 20th century, with most physicians of the developed world not knowing the concept of an untreatable bacterial disease. Such optimism was nevertheless lessened by the appearance of infections resistant to available antimicrobials, as for instance the sulfonamide-resistant *Streptococcus pyogenes* in military hospitals in the 30s [24], or the penicillin-resistant *Staphylococcus aureus* in London in the 40s [25]. These mostly isolated cases have evolved, in the last decade, into a major concern, due to the interactions between three factors: the emergence of multi-drug resistance, proving that the use of multiple antimicrobials at the same time is not always a solution; the worldwide dissemination of some bacteria, particularly in countries with poor hygiene; and the reduced number of new drugs. Initiatives have been put in place to estimate the size of the problem. A 2014 report of the WHO estimated that 3.6% of new tuberculosis cases, and 20.2% of previously treated cases, were multidrug-resistant [26]. More alarming, *Klebsiella pneumoniae* has been found to be resistant to carbapenems in 54% of the cases; note that carbapenems are the last-resort solution to severe community and hospital-acquired infections [27]. The same report also estimated a yearly cost of microbial resistance to the US health system of US \$21 to \$34 billion, accompanied by more than eight million additional days in hospital – see also [28, 29] for alternative estimations. All of this led to the introduction of the concept of the “Post-Antibiotic Era”: the real possibility that in the 21st century death may be caused by common infections and minor injuries.

In order to tackle this problem, it is necessary to first understand what causes antimicrobial resistance. In simplistic terms, random mutations can make some bacteria immune to specific drugs; yet, the

² <https://www.who.int/emergencies/ten-threats-to-global-health-in-2019>

widespread use of these drugs, and especially their uncontrolled and superfluous use, introduces an evolutionary pressure, making these bacteria proliferate [30, 31]. In addition, genes with resistance traits can be transferred among bacteria of different taxonomic and ecological groups: in other words, resistance can spread [32]. Finally, this ability to resist eventually disappears in the absence of the antibiotic, but at a very slow pace – due to the minimal survival cost to the emerging resistant strains [33]. Given this complex scenario, how can AI help?

Peptides

If old antimicrobial drugs are not working, or are expected to lose efficacy in the near future, the obvious solution is to develop new ones. This is nevertheless not a simple process, and especially not a cheap one; yet, it could be made more efficient if a system was available to *in silico* screening which drugs are expected to be more effective, and hence to save *in vivo* tests.

One promising field is the use of peptides, i.e., short chains of amino acid monomers linked by peptide bonds, conceptually similar to but shorter than proteins [34]. Some of these peptides have been recognized as potent and broad-spectrum antibiotics, which could be effective where standard therapies are starting to fail. AI can help in two aspects: on one hand, given a peptide structure, to predict the degree of antimicrobial activity; and, on the other hand, to suggest new synthetic peptide structures.

The first option is not new in the literature, with examples being published back to 2007 [34], and the interest is increasing, as shown by the large number of papers published in the last two years. From a technical point of view, these papers can be organized in two groups. Several authors have applied standard data mining models, including Support Vector Machines (SVM) [36–38] and semi-supervised clustering [39]. Attempts of using more complex solutions, as Long Short Term Memory (LSTM) [40] artificial neural networks, have also been performed [41], supported by the rationale that sequences of amino acids conceptually resemble discrete time series. In both cases, results were remarkable, reaching an accu-

racy in the discrimination of antimicrobial vs. non-antimicrobial peptides of 91.9% [36], 94.76% [37], and 95.79% [41]. It has nevertheless been highlighted that special care should be used when interpreting these results, and especially avoid taking them at face value. Specifically, Porto *et al.* [42] have shown that most existing algorithms could not differentiate between two highly similar peptides, but with different antimicrobial activities. This seems to stem from three potential reasons related to the way test data sets were created: (i) there is no common database on non-antimicrobial peptides, (ii) the antimicrobial activity is usually evaluated on a very limited number of strains, and (iii) the antimicrobial activity is not differentiated between antibacterial, antifungal, and/or antiviral. Furthermore, Lee *et al.* [36] demonstrated that a SVM model can at first seem effective, but it can actually be assessing other overlapping properties – in this specific case, peptides were not classified by their antimicrobial activity, but instead by their membrane penetration. Readers will find interesting discussions on how to overcome such limitations in both papers.

Instead of testing known peptides, one may choose a different path and design brand new peptides; this is usually done by combining short sequences known to be part of effective peptides, or by modifying few elements of a given model peptide. Several artificial intelligence approaches have been tried in the last years, including artificial neural networks and their LSTM variant [43, 44], evolutionary algorithms [45], and genetic programming [46]. While all these contributions achieved interesting results, as the identification of large sets of new peptides with forecasted antimicrobial action, it is worth noting that the creation of new peptides faces the same problem as the *in silico* test of existing ones: until a good validation procedure is found, all results have to be tested *in vitro* before any claim could be made.

Dashboard

One of the most important conditions to be able to successfully tackle any problem is to have enough information about the problem itself. Being antimicrobial resistance no exception, several organizations already

recognized the importance of collecting up-to-date information about the appearance of new resistant bacteria [26, 47]. The first challenge clearly resides in obtaining such information: many countries, especially poor ones, do not track antimicrobial resistance cases, and some of those that do only provide data based on a small number of tests, thus increasing the uncertainty about their reliability [26]. On top of this, as it is well known in the Big Data community, collecting data is not tantamount to being able to understand them. This is why an increasing attention has been devoted to the design of dashboards, i.e., systems able to pre-process and integrate heterogeneous data, and present a synthetic view of the information to the user – in the form, for instance, of a trend analysis or of the geographical distribution of threads. Such information could then be used to better guide policy making [48], e.g., to interpret the data with an intersectoral approach (human, animal, environment), and to assess and predict the risk of transmission. Beyond many older works, which are reviewed in [48], it is worth highlighting a recent paper describing the design and evaluation of such a dashboard in northwestern Iranian hospitals [49], for going beyond a pure research work and being an example of an actual deployed solution.

Stewardship and Antimicrobial Susceptibility Tests

If one of the causes behind the appearance of antimicrobial resistance is the over-prescription of antimicrobials in general, and of antibiotics in particular, a simple solution stems from a more rational use of such drugs. This is nevertheless more complex than what may *prima facie* appears, as multiple elements may intervene in the definition of what an appropriate drug usage is, as for instance the identification (or not) of the infectious pathogen, or the general condition of the patient. These situations then impact specific antimicrobial choice, dose, dosing interval, and final de-escalation. The solution may come from the adoption of antimicrobial stewardships, i.e., decision support systems tailored to this specific problem. Specifically, it has repeatedly been suggested that the use of artificial intelligence strategies may help reach more optimal solutions, in terms of

what antimicrobial is better to prescribe, and in what dosage. For instance, Schouten and De Waele state in [50] that “the large amounts of data we collect in electronic medical records will increasingly be supported by analytic tools and artificial intelligence.”. Other works proposed similar messages [51, 52]. Several examples of this can be found in the literature, as for instance [53-55]. Yet, curiously enough, no research work on AI applications for antimicrobial stewardships has been published during the last two years.

Within the elements composing a stewardship system, one is especially susceptible to be improved by an artificial intelligence approach: the design of effective antimicrobial susceptibility tests (ASTs). These tests try to assess which antimicrobial agent is best suited to tackle a specific infection, thus reducing the need of prescribing large spectrum ones [56], and are therefore a perfect example of personalized medicine. While several testing methods are customarily used, usually based on the comparison of different *in vitro* cultures, these often take several days and cannot thus support real time decisions. Three different solutions have recently been proposed. Firstly, Yu *et al.*, [57] resorted to deep learning to analyze videos of freely moving bacterial cells in urine in real time, for then extracting several relevant phenotypical features; the wait was thus strongly reduced, and results were made available in less than 30 minutes. Secondly, Wu *et al.*, [58] also analysed images, but focused on how bacteria form clusters, and on extracting topological features from them through linear regressors, SVMs and artificial neural networks. Lastly, Athamanolap *et al.*, [59] combined machine learning algorithms and Polymerase Chain Reaction (PCR) to infer the exact bacterial strand from the corresponding genetic code. In this case, results were available in approximately 6.5 hours.

Infection Prediction

As a last topic, let us recall that medicine is always in favor of preventing a disease over curing it; and in order to reduce antimicrobial usage, the best solution would entail completely avoiding the infection, instead of trying to tackle it once it has emerged. Consequently, a last line of research focuses on using artificial intelligence tools to assess

the risk of an infection, especially in critical places like emergency departments, as a function of the patient’s characteristics or of the details included in his/her electronic health record. In the last two years, several alternatives have been tested, including SVM [60], decision trees (DT) [61, 62], random forests (RF) [63] and logistic regressions [64]. Each of these papers proposed a different approach, especially in terms of the data used – respectively biochemical markers, curated patient data, and electronic health records. Still, and in spite of such heterogeneity, AUC (Area Under the Curve) [65] scores spanned from 0.8 [60] to 0.962 [61], making it clear that such support algorithms may serve as a basis for deploying prophylactic measures, hence reducing the need of resorting to larger quantities of antimicrobial drugs once the disease has been contracted.

3.2 Impact of Climate Change in Disease Epidemiology

Human activities associated with our modern civilization are resulting in changes to Earth’s climate that are faster than what has been observed at any point in history. The impact of such changes has been widely measured, both from socioeconomic [66-69], environmental, and natural perspectives. There are issues like the geographical redistribution of plant and animal species globally [70], shrinkage of glaciers [71], increase in global average Earth’s surface temperature [72], diseases in tropical plantation crops [73], increment of greenhouse gas concentrations in ocean systems [74], increases in rain and its intensity in high latitudes, increase of the growing seasons length, early trees flowering, insects emergence, and egg-laying in birds, among others [75].

Less known, but yet not less important, is the deep influence that climate change is having in the emergence and spreading of diseases, especially infectious ones [76, 77], due to the dependence of their transmission on climatic factors such as temperature, precipitations, and humidity [78]. For instance, recent studies have shown that climate change is fostering the melting of permafrost, which is in turn contributing to the release of biological agents on the Earth surface, as is happening with anthrax in the Russian Arctic

[78]. Another example is provided by diseases that are very specific to a given geographical region: as shown by Waits *et al.*, [78], there are several diseases that may spread because of climate change, as is the case of tropical diseases [79]. Higher temperatures and increased precipitations are influencing the life cycle and the distribution of ticks, their spreading, development, and reproduction. As a consequence, tick-borne diseases are also spreading. Similarly, the West Nile virus, which is transmitted by mosquitos, is also expected to proliferate based on changes in mosquito populations. The increase in temperatures is allowing the overwintering of species and expanding the range of the disease-causing vectors. Chikungunya, a virus transmitted by the *Aedes sp.* mosquito, that caught a lot of attention from the media in the last years, is also affected by climate change: once again, the increase of the global temperature could lead to a proliferation of this mosquito in southern coastal regions. Several other diseases were studied in this article, including dirofilariosis, tularemia, pumala virus, rabies, as well as airborne, food, and waterborne diseases.

It is undeniable that climate change is one of the most important scientific topics of the last year, and that the application of AI in this area is neither new nor of lesser importance [80-84]. This is also confirmed by international initiatives such as the publication of the European H2020 call “SC1-BHC-13-2019” which aims to “Mine big data for early detection of infectious disease threats driven by climate change and other factors”. In spite of this, and most surprisingly, the number of relevant scientific publications in this area is not very high, with only 13 publications being available for analysis. This suggests that this trend may change, and that the application of AI to disease spreading and climate change may soon become a major scientific field. In the following, we discuss the main efforts in the area in the last two years, organised in two main subtopics: disease forecasting and surveillance.

Disease Prediction and Forecasting

Disease prediction aims at creating models to describe or predict how diseases can be introduced in a different environment (disease transmission), or how the prevalence of a specific disease may increase based on the

measurement of different variables associated with its proliferation. This information could then be used to deploy prophylactic measures, and eventually to limit the proliferation of the disease. Climate change has been recognised as one of the key influencing elements: variations in climatic characteristics, such as temperature or humidity, are well-known factors that can modify how a disease is transmitted between different regions or environments [85]. The study of such factors and relationships is a very relevant research topic, and several research works have been done in the past based on the creation of predictive frameworks [86]. While most of those leverage on more classical statistical approaches, still more complex AI techniques relying on ML methods have yet to be considered - as it could be the case, for instance, for prediction models based on Artificial Neural Networks.

Leishmaniasis is a disease caused by leishmania parasites, and is endemic in 97 countries according to the WHO [87]. It is autochthonous of Mexico and Texas, but it has started to expand northward. According to González *et al.* [88], such expansion may be influenced by climate change, as more habitats become suitable for vector and reservoir species. The authors created a set of ecological niche models, and used a maximum entropy algorithm to predict how sand fly vector species may geographically relocate. The models were constructed over a set of strict assumptions about the ability of the species to transmit to suitable habitats. These were further applied to the extreme A2 and conservative B2 projected climate scenarios for 2020, 2050, and 2080, by using public interpolated climate data. Results for different species that can spread the disease include AUC values ranging from 0.963 to 0.984, above the conservative threshold set by the authors (0.95). An interesting observation is also that, depending on the specific analysed species, the relative importance of the predictor variables changes, thus highlighting the importance of building sophisticated models. As a conclusion, the authors suggest that, based on these results, climate change will increase the chance of leishmaniasis spreading in north America, north of México, and Texas.

Johansson *et al.* developed a framework to assess and compare dengue forecasts from different models, and to further evaluate the performance of seasonal models with and

without climate variables in forecasting dengue incidence in Mexico [89]. The dataset used consists in dengue and dengue haemorrhagic fever cases reported monthly from January 1985 to December 2012, provided by the Mexican Health Secretariat. Weather data were obtained from the National Oceanic and Atmospheric Administration North American Regional Reanalysis dataset. The analysed models were based on three techniques: linear, autoregressive, and seasonal autoregressive models. The evaluation was performed by splitting the data in three subsets: 1985-1989 for training, 1990-2007 for an 18-year model evaluation, and 2008-2012 for an out-of-sample validation after the selection of the model. Results showed that climate data were not significantly influencing the predictive power of the model and that the seasonal and short-term autocorrelations played a major role in the short- and long-term forecast precision.

A more AI-focused approach was proposed by Wang *et al.* [90]. Although the paper does not explicitly mention climate change, it is focused on meteorological factors, clearly one of its major consequences. The authors tried to forecast the weekly number of infectious diarrhoea episodes in Shanghai considering meteorological factors that were likely to have a significant influence on the propagation of the disease. Nine factors were initially selected, by applying a Spearman rank correlation analysis. The considered dataset spanned from March 2005 to April 2009 and was retrieved from the National Disease Supervision Information Management System, a real-time and online system. ML models applied to the study included feed-forward back-propagation ANNs (BPNN) [91], support vector regressions (SVR) [92], and random forest regressions (RFR) [63]. Results suggested that the BPNN model with a 9-4-1 architecture provided the most accurate predictions. The sensitivity analysis also revealed that the most effective meteorological factors are the temperature-related variables (max, min, and average), whereas weekly average rainfall is the least effective one.

Disease Surveillance

According to the WHO, surveillance “is the continuous, systematic collection, analysis, and interpretation of health-related data

needed for the planning, implementation, and evaluation of public health practice”³. According to this definition, surveillance aims, among other objectives, at monitoring the spread of a disease, allowing to infer progression patterns and develop early contention measures. As in the case of the dashboard concept for the antimicrobial resistance, having reliable information about a problem is the first step towards achieving a solution. In the new era of Big Data, in which large amounts of data from different and various sources are available, disease surveillance has a unique opportunity. Data coming from sensors, smart cities, connected vehicles, smart houses, Internet of Things, and social media are providing new ways of tracking diseases. These could be complemented with the vast amount of information available at primary care centres and hospitals. However, there are new variables based on the climate evolution that ought also to be considered. Climate change is modifying the environment in such a way that the spreading of a disease can follow multiple new patterns, as has already been suggested by some studies [93]; therefore, more evidences and new types of data are required [94].

The exploitation of big data and the use of new techniques based on AI will become extremely important. As has already been pointed out, nowadays data are generated from very different sources, with volumes of the order of exabytes; yet, these have to be exploited. Along this line, Manogaran and Lopez proposed a big data surveillance system to analyse special climate big data. They performed a continuous monitoring of the correlation between climate change and the Dengue disease [95].

Leveraging on a similar approach, Traore *et al.*, proposed the use of data mining techniques for the discovery of knowledge in satellite images [96]. The study is based on the high rates of cholera which is epidemic in less developed countries. It aims at strengthening the capacity of epidemiological surveillance, by discovering risk areas through the use of remote sensing satellite data and environment, climate, and health information. In particular, satellite data were

³ https://www.who.int/topics/public_health_surveillance/en/

combined with field data to explain and deduce the causes of the disease evolution. The paper has been focused on the use of supervised classification algorithms, and the resulting models are currently deployed in the Mopti region, Mali (West Africa).

Another paper, recently published in the Proceedings of the National Academy of Sciences of the United States of America (PNAS) [97], analysed the relationship between climate change and the transmission of the dengue infectious disease. Authors stated that the disease had increasingly spread to new countries, over larger areas, and towards more temperate zones. The study was performed using dengue surveillance data for the years 2005–2015 from Guangzhou, China. The authors used the first eight years to train the models, and the remaining three for testing. An exploration of the direct and indirect effects of temperature and precipitation on dengue outbreaks was proposed based on a structural equation model. Results suggested that there were significant, direct and positive correlations between the temperature, precipitation, and the dengue incidence, as well as indirect positive effects on both variables through the vector density. Authors concluded that climate is having both direct and indirect effects on the disease incidence, the latter ones due to changes in mosquito dengue density.

Knowledge Representation

As a final interesting work within the field of AI, climate change, and diseases, we here discuss the study published by Semenza *et al.* [98]. The authors extracted all the articles published between 1998 and 2009 from PubMed and ScienceDirect that were relevant for climate change and food- and waterborne diseases. These papers included information about pathogens and their relationship with climate, being climate change summarized as a key fact. Afterwards, they stored that information in a database and tagged it with a set of predefined terms used in the field. Results highlighted that much information is still needed about the potential effects of the changes in climatic variables (changes in air or water temperature, precipitation, humidity, UV radiation, wind, cloud coverage, sunshine hours or seasonality), and on the pathogens evaluated.

4 Conclusions and Outlook

The increasing interest in AI witnessed in the last years is undoubtedly composed in part by hype. However, the interest also clearly reflects the high expectations that it generates, as AI enables radical new solutions to complex unsolved problems. Not surprisingly, AI and ML are gaining momentum within the public health field. Yet, many barriers have still to be overcome. Examples of these barriers include the access to the data, the development or improvement of methods to be converted in white-box approaches allowing to have a rational explanation of the results of AI and ML techniques, or the exploitation of complementary sources, like, for example, social networks, as this has been done in the past [99, 100] for similar objectives in public health problems.

This overview focused on the main research works that have been developed (or at least published) in the last two years in the addressed topics. The authors did an analysis that had some limitations: other sources could be used, other keywords and topics could be considered, and a more rigorous process could be carried out. However, the aim of the work was more focused on providing the reader with a broad perspective about these two important public health challenges and the main initiatives that have been done in terms of applying artificial intelligence methods. Even at this high level of granularity, we detected an important lack of papers matching the criteria for being included in this analysis.

The analysis has been focused on two seemingly different topics, i.e., antimicrobial resistance and climate change in epidemiology, which nevertheless share many common characteristics. First of all, they have attracted the attention of both the scientific and general communities, due to the potentially hefty impact they will have in the upcoming decades. Secondly, both problems can also benefit from an increasing availability of real data, not necessarily just coming from traditional sources, but even from social networks and media [101, 102]. Yet, analysing those data is not a simple endeavour: it requires combining knowledge and techniques coming from fields far apart, and overcoming biases and bad practices [36, 37, 39, 41, 42]. Finally, and possibly because of this complexity, a surprising low number of research studies have been published in the

last two years on AI applied to antimicrobial resistance and climate change in epidemiology; and this in spite of many opinions in favour of such integration [19,94,103], and even with the availability of specialised research funds. Factoring the need for new solutions, the promise of relevant results, and the availability of funds, we believe this state of affairs will change, and that AI will become not only the centre of attention, but more importantly, the source of innovative solutions for public health.

Acknowledgments

This paper is supported by project “DISNET (Creation and analysis of disease networks for drug repurposing from heterogeneous data sources applied to rare diseases)”, that is being developed under grant “RTI2018-094576-A-I00” from the Spanish Ministerio de Ciencia, Innovación y Universidades.

References

1. Russell S, Norvig P. Artificial Intelligence: A Modern Approach. 3rd ed. Prentice Hall; 2009.
2. Patel VL, Shortliffe EH, Stefanelli M, Szolovits P, Berthold MR, Bellazzi R, et al. The coming of age of artificial intelligence in medicine. *Artif Intell Med* 2009 May 1;46(1):5–17.
3. Ramesh AN, Kambhampati C, Monson JRT, Drew PJ. Artificial intelligence in medicine. *Ann R Coll Surg Engl* 2004 Sep;86(5):334–8.
4. Artificial Intelligence in Medicine | NEJM [Internet]. *N Engl J Med* [cited 2018 Dec 2]. Available from: <https://www.nejm.org/doi/pdf/10.1056/NEJM198703123161109>.
5. Horvitz EJ, Breese JS, Henrion M. Decision theory in expert systems and artificial intelligence. *Int J Approx Reason* 1988 Jul 1;2(3):247–302.
6. Miller PL. The evaluation of artificial intelligence systems in medicine. *Comput Methods Programs Biomed* 1986;22(1):3–11.
7. Bini SA. Artificial Intelligence, Machine Learning, Deep Learning, and Cognitive Computing: What Do These Terms Mean and How Will They Impact Health Care? *J Arthroplasty* 2018 Aug;33(8):2358–61.
8. Du X-L, Li W-B, Hu B-J. Application of artificial intelligence in ophthalmology. *Int J Ophthalmol* 2018;11(9):1555–61.
9. Hashimoto DA, Rosman G, Rus D, Meireles OR. Artificial Intelligence in Surgery: Promises and Perils. *Ann Surg* 2018 Jul;268(1):70–6.
10. Jiang F, Jiang Y, Zhi H, Dong Y, Li H, Ma S, et al. Artificial intelligence in healthcare: past, present and future. *Stroke Vasc Neurol* 2017 Dec;2(4):230–43.
11. Krittanawong C, Zhang H, Wang Z, Aydar M, Kitai T. Artificial Intelligence in Precision Cardiovascular Medicine. *J Am Coll Cardiol* 2017 May

- 30;69(21):2657–64.
12. Miller DD, Brown EW. Artificial Intelligence in Medical Practice: The Question to the Answer? *Am J Med* 2018 Feb;131(2):129–33.
 13. Sniecinski I, Seghatchian J. Artificial intelligence: A joint narrative on potential use in pediatric stem and immune cell therapies and regenerative medicine. *Transfus Apher Sci Off J World Apher Assoc Off J Eur Soc Haemapheresis* 2018 Jun;57(3):422–4.
 14. Stewart J, Sprivilis P, Dwivedi G. Artificial intelligence and machine learning in emergency medicine. *Emerg Med Australas EMA* 2018 Dec;30(6):870–4.
 15. Chen JH, Asch SM. Machine Learning and Prediction in Medicine — Beyond the Peak of Inflated Expectations. *N Engl J Med* 2017 Jun 29;376(26):2507–9.
 16. Fogel AL, Kvedar JC. Artificial intelligence powers digital medicine. *Npj Digit Med* 2018 Mar 14;1(1):5.
 17. Naughton J. Magical thinking about machine learning won't bring the reality of AI any closer | John Naughton. *The Guardian* [Internet] 2018 Aug 5 [cited 2018 Nov 21]; Available from: <https://www.theguardian.com/commentisfree/2018/aug/05/magical-thinking-about-machine-learning-will-not-bring-artificial-intelligence-any-closer>.
 18. Schwartz O. “The discourse is unhinged”: how the media gets AI alarmingly wrong. *The Guardian* [Internet] 2018 Jul 25 [cited 2018 Nov 21]; Available from: <https://www.theguardian.com/technology/2018/jul/25/ai-artificial-intelligence-social-media-bots-wrong>.
 19. Lynch SM, Moore JH. A call for biological data mining approaches in epidemiology. *BioData Min* [Internet] 2016 Jan 4 [cited 2018 Nov 21];9. Available from: <https://www.ncbi.nlm.nih.gov/pmc/articles/PMC4700596/>.
 20. Jones N. How machine learning could help to improve climate forecasts. *Nat News* 2017 Aug 24;548(7668):379.
 21. McGovern A, Elmore KL, Gagne DJ, Haupt SE, Karstens CD, Lagerquist R, et al. Using Artificial Intelligence to Improve Real-Time Decision-Making for High-Impact Weather. *Bull Am Meteorol Soc* 2017 Mar 24;98(10):2073–90.
 22. Flahault A, Bar-Hen A, Paragios N. Public Health and Epidemiology Informatics. *Yearb Med Inform* 2016 Nov 10;(1):240–6.
 23. Thiébaud R, Thiéssard F. Public Health and Epidemiology Informatics. *Yearb Med Inform* 2017;26(1):248–51.
 24. Levy SB. Microbial resistance to antibiotics. An evolving and persistent problem. *Lancet Lond Engl* 1982 Jul 10;2(8289):83–8.
 25. Barber M, Rozwadowska-Dowzenko M. Infection by penicillin-resistant staphylococci. *Lancet Lond Engl* 1948 Oct 23;2(6530):641–4.
 26. WHO | Antimicrobial resistance: global report on surveillance 2014 [Internet]. WHO. 2014 [cited 2018 Nov 26]. Available from: <http://www.who.int/drugresistance/documents/surveillance-report/en/>.
 27. Zhanel GG, Wiebe R, Dilay L, Thomson K, Rubinstein E, Hoban DJ, et al. Comparative review of the carbapenems. *Drugs* 2007;67(7):1027–52.
 28. Cassini A, Högberg LD, Plachouras D, Quattrocchi A, Hoxha A, Simonsen GS, et al. Attributable deaths and disability-adjusted life-years caused by infections with antibiotic-resistant bacteria in the EU and the European Economic Area in 2015: a population-level modelling analysis. *Lancet Infect Dis* 2019 Jan;19(1):56–66.
 29. Stemming the Superbug Tide [Internet]. [cited 2019 Apr 10]. Available from: https://www.oecd-ilibrary.org/social-issues-migration-health/stemming-the-superbug-tide_9789264307599-en
 30. Levy SB. Balancing the drug-resistance equation. *Trends Microbiol* 1994 Oct;2(10):341–2.
 31. Wormser GP, Bergman MM. The Antibiotic Paradox: How the Misuse of Antibiotics Destroys Their Curative Powers, 2nd Edition By Stuart B. Levy Cambridge, Massachusetts: Perseus Publishing; 2002. 376 pp., illustrated. \$17.50 (paper). *Clin Infect Dis* 2003 Jan 15;36(2):238.
 32. Milkman R. Gene Transfer in the Environment. Stuart B. Levy and Robert V. Miller. McGraw-Hill, New York, 1989. x, 434 pp., illus. \$54.95. *Environmental Biotechnology. Science* 1990 Jan 19;247(4940):350–1.
 33. Levy SB, Marshall B. Antibacterial resistance worldwide: causes, challenges and responses. *Nat Med* 2004 Dec;10(12 Suppl):S122–129.
 34. Giuliani A, Rinaldi AC, editors. *Antimicrobial Peptides: Methods and Protocols* [Internet]. Humana Press; 2010 [cited 2018 Nov 26]. (Methods in Molecular Biology). Available from: <http://www.springer.com/la/book/9781607615934>.
 35. Fjell CD, Hancock REW, Cherkasov A. AMPper: a database and an automated discovery tool for antimicrobial peptides. *Bioinforma Oxf Engl* 2007 May 1;23(9):1148–55.
 36. Lee EY, Lee MW, Fulan BM, Ferguson AL, Wong GCL. What can machine learning do for antimicrobial peptides, and what can antimicrobial peptides do for machine learning? *Interface Focus* 2017 Dec 6;7(6):20160153.
 37. Mousavizadegan M, Mohabatkari H. Computational prediction of antifungal peptides via Chou's PseAAC and SVM. *J Bioinform Comput Biol* 2018 May 29;1850016.
 38. Cristianini N, Shawe-Taylor J. An Introduction to Support Vector Machines and Other Kernel-based Learning Methods by Nello Cristianini [Internet]. Cambridge Core 2000 [cited 2019 Apr 10]. Available from: [/core/books/an-introduction-to-support-vector-machines-and-other-kernel-based-learning-methods/A6A6F4084056A4B23F88648DDBFDD6FC](http://core/books/an-introduction-to-support-vector-machines-and-other-kernel-based-learning-methods/A6A6F4084056A4B23F88648DDBFDD6FC).
 39. Vishnepolsky B, Gabrielian A, Rosenthal A, Hurt DE, Tartakovsky M, Managadze G, et al. Predictive Model of Linear Antimicrobial Peptides Active against Gram-Negative Bacteria. *J Chem Inf Model* 2018 May 29;58(5):1141–51.
 40. Hochreiter S, Schmidhuber J. Long Short-Term Memory. *Neural Comput* 1997 Nov;9(8):1735–80.
 41. Youmans M, Spainhour C, Qiu P. Long short-term memory recurrent neural networks for antibacterial peptide identification. In: 2017 IEEE International Conference on Bioinformatics and Biomedicine (BIBM); 2017. p. 498–502.
 42. Porto WF, Pires AS, Franco OL. Antimicrobial activity predictors benchmarking analysis using shuffled and designed synthetic peptides. *J Theor Biol* 2017 Aug 7;426:96–103.
 43. Müller AT, Hiss JA, Schneider G. Recurrent Neural Network Model for Constructive Peptide Design. *J Chem Inf Model* 2018 Feb 26;58(2):472–9.
 44. Schneider P, Müller AT, Gabernet G, Button AL, Posselt G, Wessler S, et al. Hybrid Network Model for “Deep Learning” of Chemical Data: Application to Antimicrobial Peptides. *Mol Inform* 2017 Jan 1;36(1–2):1600011.
 45. Yoshida M, Hinkley T, Tsuda S, Abul-Haija YM, McBurney RT, Kulikov V, et al. Using Evolutionary Algorithms and Machine Learning to Explore Sequence Space for the Discovery of Antimicrobial Peptides. *Chem* 2018 Mar 8;4(3):533–43.
 46. Veltri D, Kamath U, Shehu A. Improving Recognition of Antimicrobial Peptides and Target Selectivity through Machine Learning and Genetic Programming. *IEEE/ACM Trans Comput Biol Bioinform* 2017 Mar;14(2):300–13.
 47. Wernli D, Jørgensen PS, Harbarth S, Carroll SP, Laxminarayan R, Levrat N, et al. Antimicrobial resistance: The complex challenge of measurement to inform policy and the public. *PLoS Med* 2017 Aug;14(8):e1002378.
 48. Safdari R, Ghazi-Saeedi M, Masoumi-Asl H, Rezaei-Hachesu P, Mirnia K, Samad-Soltani T. Knowledge discovery and visualization in antimicrobial resistance surveillance systems: a scoping review. *Artif Intell Rev* [Internet] 2018 Sep 25 [cited 2018 Nov 26]; Available from: <https://doi.org/10.1007/s10462-018-9659-6>.
 49. Rezaei-Hachesu P, Samad-Soltani T, Yaghoobi S, Ghazi-Saeedi M, Mirnia K, Masoumi-Asl H, et al. The design and evaluation of an antimicrobial resistance surveillance system for neonatal intensive care units in Iran. *Int J Med Inf* 2018 Jul;115:24–34.
 50. Schouten J, De Waele J. *Antimicrobial Stewardship in ICU*. Academic Press; 2017.
 51. Barlow G. Clinical challenges in antimicrobial resistance. *Nat Microbiol* 2018 Mar;3(3):258.
 52. Kim S-W. Antimicrobial Stewardship with Intravenous to Oral Conversion and Future Directions of Antimicrobial Stewardship. *Infect Chemother*. 2017 Mar;49(1):87–9.
 53. Beaudoin M, Kabanza F, Nault V, Valiquette L. Evaluation of a machine learning capability for a clinical decision support system to enhance antimicrobial stewardship programs. *Artif Intell Med* 2016 Mar;68:29–36.
 54. Cánovas-Segura B, Campos M, Morales A, Juarez JM, Palacios F. Development of a clinical decision support system for antibiotic management in a hospital environment. *Prog Artif Intell* 2016 Aug 1;5(3):181–97.
 55. Evans RS, Pestotnik SL, Classen DC, Clemmer TP, Weaver LK, Orme JF, et al. A computer-assisted management program for antibiotics and other antiinfective agents. *N Engl J Med* 1998 Jan 22;338(4):232–8.
 56. Leonard H, Colodner R, Halachmi S, Segal E. Recent Advances in the Race to Design a Rapid Diagnostic Test for Antimicrobial Resistance. *ACS Sens* 2018 Nov 26;3(11):2202–17.
 57. Yu H, Jing W, Iriya R, Yang Y, Syal K, Mo M, et al. Phenotypic Antimicrobial Susceptibility Testing with Deep Learning Video Microscopy. *Anal Chem* 2018 May 15;90(10):6314–22.
 58. Wu T-F, Chen Y-C, Wang W-C, Fang Y-C, Fukuoka S, Pride DT, et al. A Rapid and Low-Cost Pathogen

- Detection Platform by Using a Molecular Agglutination Assay. *ACS Cent Sci* [Internet] 2018 Nov 5 [cited 2018 Nov 26]; Available from: <https://doi.org/10.1021/acscentsci.8b00447>.
59. Athamanolap P, Hsieh K, Chen L, Yang S, Wang T-H. Integrated Bacterial Identification and Antimicrobial Susceptibility Testing Using PCR and High-Resolution Melt. *Anal Chem* 2017 Nov 7;89(21):11529–36.
 60. Hernandez B, Herrero P, Rawson TM, Moore LSP, Evans B, Toumazou C, et al. Supervised learning for infection risk inference using pathology data. *BMC Med Inform Decis Mak* 2017 Dec 8;17(1):168.
 61. Zhao J. Study on the application of classification tree model in building the risk model for nosocomial infection in critically ill patients of emergency department. *Chin J Postgrad Med* 2017 Jan 1;40(9):791–4.
 62. Quinlan JR. Induction of decision trees. *Mach Learn* 1986 Mar 1;1(1):81–106.
 63. Breiman L. Random Forests. *Mach Learn* 2001 Oct 1;45(1):5–32.
 64. Hartvigsen T, Sen C, Brownell S, Teeple E, Kong X, Rundensteiner E. Early Prediction of MRSA Infections using Electronic Health Records. In 2018 [cited 2018 Nov 26]. p. 156–67. Available from: <http://www.scitepress.org/PublicationsDetail.aspx?ID=wFHke4fgXY4=&t=1>.
 65. Ling CX, Huang J, Zhang H. AUC: A Statistically Consistent and More Discriminating Measure Than Accuracy. In: Proceedings of the 18th International Joint Conference on Artificial Intelligence [Internet]. San Francisco, CA, USA: Morgan Kaufmann Publishers Inc.; 2003 [cited 2019 Apr 10]. p. 519–24. (IJCAI'03). Available from: <http://dl.acm.org/citation.cfm?id=1630659.1630736>.
 66. Estrada F, Botzen WJW, Tol RSJ. A global economic assessment of city policies to reduce climate change impacts. *Nat Clim Change* 2017 Jun;7(6):403–6.
 67. Hsiang S, Kopp R, Jina A, Rising J, Delgado M, Mohan S, et al. Estimating economic damage from climate change in the United States. *Science* 2017 Jun 30;356(6345):1362–9.
 68. Tol RSJ. The Economic Impacts of Climate Change. *Rev Environ Econ Policy* 2018 Feb 1;12(1):4–25.
 69. Zhang P, Zhang J, Chen M. Economic impacts of climate change on agriculture: The importance of additional climatic variables other than temperature and precipitation. *J Environ Econ Manag* 2017 May 1;83:8–31.
 70. Pecl GT, Araújo MB, Bell JD, Blanchard J, Bonebrake TC, Chen I-C, et al. Biodiversity redistribution under climate change: Impacts on ecosystems and human well-being. *Science* 2017 Mar 31;355(6332):eaai9214.
 71. Glaciers and climate change | National Snow and Ice Data Center [Internet]. [cited 2018 Nov 25]. Available from: <https://nsidc.org/cryosphere/glaciers/questions/climate.html>.
 72. Seneviratne SI, Donat MG, Mueller B, Alexander LV. No pause in the increase of hot temperature extremes. *Nat Clim Change* 2014 Feb 26;4:161–3.
 73. Ghini R, Bettli W, Hamada E. Diseases in tropical and plantation crops as affected by climate changes: current knowledge and perspectives. *Plant Pathol* 2011 Feb 1;60(1):122–32.
 74. Hoegh-Guldberg O, Bruno JF. The Impact of Climate Change on the World's Marine Ecosystems. *Science* 2010 Jun 18;328(5985):1523–8.
 75. Climate Change | Turning Up the Heat [Internet]. Taylor & Francis. [cited 2018 Nov 25]. Available from: <https://www.taylorfrancis.com/books/9781134035588>.
 76. Lafferty KD. The ecology of climate change and infectious diseases. *Ecology* 2009 Apr 1;90(4):888–900.
 77. Patz JA, Epstein PR, Burke TA, Balbus JM. Global Climate Change and Emerging Infectious Diseases. *JAMA* 1996 Jan 17;275(3):217–23.
 78. Waits A, Emelyanova A, Oksanen A, Abass K, Rautio A. Human infectious diseases and the changing climate in the Arctic. *Environ Int* 2018 Dec 1;121:703–13.
 79. Booth M. Chapter Three - Climate Change and the Neglected Tropical Diseases. In: Rollinson D, Stothard JR, editors. *Advances in Parasitology* [Internet]. Academic Press; 2018 [cited 2018 Nov 25]. p. 39–126. Available from: <http://www.sciencedirect.com/science/article/pii/S0065308X18300046>.
 80. Manogaran G, Lopez D. Spatial cumulative sum algorithm with big data analytics for climate change detection. *Comput Electr Eng* 2018 Jan 1;65:207–21.
 81. Faghmous JH, Kumar V. A Big Data Guide to Understanding Climate Change: The Case for Theory-Guided Data Science. *Big Data* 2014 Aug 25;2(3):155–63.
 82. Jang SM, Hart PS. Polarized frames on “climate change” and “global warming” across countries and states: Evidence from Twitter big data. *Glob Environ Change* 2015 May 1;32:11–7.
 83. Hampton SE, Strasser CA, Tewksbury JJ, Gram WK, Budden AE, Batcheller AL, et al. Big data and the future of ecology. *Front Ecol Environ* 2013 Apr 1;11(3):156–62.
 84. O’Gorman PA, Dwyer JG. Using Machine Learning to Parameterize Moist Convection: Potential for Modeling of Climate, Climate Change, and Extreme Events. *J Adv Model Earth Syst* 2018 Oct 1;10(10):2548–63.
 85. McMichael AJ, Woodruff RE, Hales S. Climate change and human health: present and future risks. *The Lancet* 2006 Mar 11;367(9513):859–69.
 86. Altizer S, Ostfeld RS, Johnson PTJ, Kutz S, Harvell CD. Climate Change and Infectious Diseases: From Evidence to a Predictive Framework. *Science* 2013 Aug 2;341(6145):514–9.
 87. WHO | Leishmaniasis [Internet]. WHO. [cited 2018 Nov 27]. Available from: http://www.who.int/gho/neglected_diseases/leishmaniasis/en/.
 88. González C, Wang O, Strutz SE, González-Salazar C, Sánchez-Cordero V, Sarkar S. Climate Change and Risk of Leishmaniasis in North America: Predictions from Ecological Niche Models of Vector and Reservoir Species. *PLoS Negl Trop Dis* 2010 Jan 19;4(1):e585.
 89. Johansson MA, Reich NG, Hota A, Brownstein JS, Santillana M. Evaluating the performance of infectious disease forecasts: A comparison of climate-driven and seasonal dengue forecasts for Mexico. *Sci Rep* 2016 Sep 26;6:33707.
 90. Wang Y, Li J, Gu J, Zhou Z, Wang Z. Artificial neural networks for infectious diarrhea prediction using meteorological factors in Shanghai (China). *Appl Soft Comput* 2015 Oct 1;35:280–90.
 91. Haykin S. *Neural Networks: A Comprehensive Foundation*. 2nd ed. Prentice Hall; 1998.
 92. Smola AJ, Schölkopf B. A tutorial on support vector regression. *Stat Comput* 2004 Aug 1;14(3):199–222.
 93. Jones KE, Patel NG, Levy MA, Storeygard A, Balk D, Gittleman JL, et al. Global trends in emerging infectious diseases. *Nature* 2008 Feb;451(7181):990–3.
 94. Kovats RS, Campbell-Lendrum DH, McMichel AJ, Woodward A, Cox JSH. Early effects of climate change: do they include changes in vector-borne disease? *Philos Trans R Soc Lond B Biol Sci* 2001 Jul 29;356(1411):1057–68.
 95. Manogaran G, Lopez D. Disease Surveillance System for Big Climate Data Processing and Dengue Transmission. *Clim Change Environ Concerns Breakth Res Pract* 2018;427–46.
 96. Traore BB, Kamsu-Foguem B, Tangara F. Data mining techniques on satellite images for discovery of risk areas. *Expert Syst Appl* 2017 Apr 15;72:443–56.
 97. Xu L, Stige LC, Chan K-S, Zhou J, Yang J, Sang S, et al. Climate variation drives dengue dynamics. *Proc Natl Acad Sci* 2017 Jan 3;114(1):113–8.
 98. Semenza JC, Hoenser C, Herbst S, Rechenburg A, Suk JE, Frechen T, et al. Knowledge Mapping for Climate Change and Food- and Waterborne Diseases. *Crit Rev Environ Sci Technol* 2012 Feb 15;42(4):378–411.
 99. Aramaki E, Maskawa S, Morita M. Twitter Catches the Flu: Detecting Influenza Epidemics Using Twitter. In: Proceedings of the Conference on Empirical Methods in Natural Language Processing [Internet]. Stroudsburg, PA, USA: Association for Computational Linguistics; 2011 [cited 2019 Mar 30]. p. 1568–76. (EMNLP '11). Available from: <http://dl.acm.org/citation.cfm?id=2145432.2145600>.
 100. Charles-Smith LE, Reynolds TL, Cameron MA, Conway M, Lau EHY, Olsen JM, et al. Using Social Media for Actionable Disease Surveillance and Outbreak Management: A Systematic Literature Review. *PLoS One* 2015 Oct 5;10(10):e0139701.
 101. Paul MJ, Dredze M, Broniatowski D. Twitter Improves Influenza Forecasting. *PLoS Curr* [Internet] 2014 Oct 28 [cited 2018 Nov 28];6. Available from: <https://www.ncbi.nlm.nih.gov/pmc/articles/PMC4234396/>.
 102. Signorini A, Segre AM, Polgreen PM. The Use of Twitter to Track Levels of Disease Activity and Public Concern in the U.S. during the Influenza A H1N1 Pandemic. *PLoS One* 2011 May 4;6(5):e19467.
 103. Macesic N, Polubriaginof F, Tatonetti NP. Machine learning: novel bioinformatics approaches for combating antimicrobial resistance. *Curr Opin Infect Dis* 2017 Dec;30(6):511–7.

Correspondence to:

Alejandro Rodríguez González
 Centro de Tecnología Biomédica
 Universidad Politécnica de Madrid
 Campus de Montegancedo, Pozuelo de Alarcón
 28223 Madrid, Spain
 E-mail: alejandro.rog@upm.es