Data-Driven Decision Making in Public Health: Integrating AI and Predictive Analytics for Ministry of Health Strategies

Barua Mohammed alghamdi[1], Maha Ahmad Seraj Alsabei[2], Ali jaman Abdullah Alghamdi[3], Shahad Saeed Mosa Alzahrani[4], Nidaa Abdullah Ibrahim Khayat[5], Wafaa Sulaiman Abdullah Alajlan[6], Bayan Eid Hashime Alghamdi[7], Ashwaq Saad Huseen Alghamdi[8]

[1] Public Health Specialist, Preventive department, Health Control Center at King Abdulaziz International Airport, Jeddah Second Health Cluster, Ministry of Health, Saudi Arabia.

[2] Public Health Specialist, Preventive department, Health Control Center at King Abdulaziz International Airport, Jeddah Second Health Cluster, Ministry of Health, Saudi Arabia.

[3] Healthcare and Hospital Services Management, Inventory Control Department, Executive Administration, Jeddah Second Health Cluster, Ministry of Health, Saudi Arabia.

[4] Public Health Specialist, Preventive department, Health Control Center at King Abdulaziz International Airport, Jeddah Second Health Cluster, Ministry of Health, Saudi Arabia.

[5] Public Health Specialist, Preventive department, Health Control Center at King Abdulaziz International Airport, Jeddah Second Health Cluster, Ministry of Health, Saudi Arabia.

[6] Nursing Technician, Public Health Department – Epidemiology (Vaccination Unit), King Abdullah Medical Complex, Jeddah Second Health Cluster, Ministry of Health, Saudi Arabia.

[7] Public Health Specialist, Public Health Administration Infectious Diseases Department, Jeddah First Health Cluster, Ministry of Health, Saudi Arabia.

[8] Public Health Specialist, Preventive department, Health Control Center at King Abdulaziz International Airport, Jeddah Second Health Cluster, Ministry of Health, Saudi Arabia

ABSTRACT

Ministry of Health uses outdated data to plan their resources which makes it difficult for them to react to the outbreaks. This research looked at how artificial intelligence and predictive analysis could assist the Ministry of Health in making data-driven decisions in their strategies. The quantitative section involved use of a simulated data from a district with climate, dengue cases, admissions, and stockout for 24 months. 5 forecasting models were compared using a time-based split where the training period was 15 months while test was 7 months. The planned qualitative phase is expected to look at the challenges of adoption using 15 to 20 interviews with high-ranking ministry of health personnel. The Expansion Window Cross Validation also proved this conclusion, since its mean RMSE value was equal to 64.65 against 134.38 of AR(2). It was determined that the rainfall during two previous months was a good predictor of dengue with a correlation of 0.940 and a coefficient of 1.262. On average, there were 480.5 cases after the monsoon period, which is 10.7 times more than in the summer period. The outbreaks of the disease have followed the pattern of the shortage of medicines. The increase in the number of cases of dengue from year 1 to year 2 is statistically significant ($p = 0.002$). As the differences between the models are insignificant, their superiority should be proved with the use of long real surveillance records. Rainfall-based early warning system can assist ministries in preparing for each monsoon season in advance.

Keywords: public health, artificial intelligence, predictive analytics, dengue forecasting, Ministry of Health, data driven decision making

INTRODUCTION

The process of allocating the scarce public health resources involves the utilization of reports that are prepared after several months after the beginning of the outbreak in several countries. The Ministries of Health of most countries possess tremendous amounts of data which can be gathered from surveillance systems, registries, laboratories, mobile devices, and social media. Dolley (2018) emphasizes that big data can be beneficial for precision public health and interventions that are implemented more effectively. Also, artificial intelligence plays its role in helping to achieve the same result since machine learning can recognize some patterns that are impossible to notice using traditional ways. Rajkomar et al. (2019) provide the explanation of the process of prediction of clinical outcomes using machine learning and large datasets. In

addition, Topol (2019) asserts that deep learning has already shown the results similar to those of specialists in some image-based diagnostic procedures. Still, Panch et al. (2018) emphasize that the infrastructure of the health care system is often lacking the implementation of such innovations on a large scale. Wahl et al. (2018) examined how artificial intelligence can improve health in resource poor settings. Authors pointed out the existence of high disease burden in low and middle-income countries along with poor data systems. Wiens et al. (2019) suggested that bad design of models would increase risks for patients and add to existing health inequalities. Framework developed by the authors describes a process of creation, implementation, and monitoring of models. In this case, Ministries of Health need governance frameworks with technical and ethical aspects. Target 3.8 of the 2030 Agenda for Sustainable Development is concerned about universal health coverage and implies availability of reliable data (Kipo-Sunyehzi, 2024). Predictive analytics will allow to predict future disease outbreaks, shortage of hospital beds, and medicine supplies. However, data about usage of these technologies in national policies are scarce and scattered in many disciplines. This research investigates the possibilities of using artificial intelligence and predictive analytics for evidence-based decisions in Ministry of Health policies.

RELATED WORK

The relationship between big data, machine learning, and disease predictions as well as planning in healthcare institutions can also be seen from the existing literature. Obermeyer and Emanuel (2016) argued that machine learning extracts complicated patterns from huge amounts of data through predefined rules. Predictions made using machine learning need validation as they may have implications for making decisions. Jiang et al. (2017) investigated the role of artificial intelligence in diagnostics in relation to neurology, cardiology and cancer diseases. Jiang et al. observed that machine learning utilizes structured and unstructured data such as images and clinical narratives. Ristevski and Chen (2018) observed that there are numerous challenges associated with big data in the healthcare sector in terms of privacy, security and data quality. Mooney and Pejaver (2018) examined the terminologies related to big data, machine learning and privacy issues in the context of public health sector. Santillana et al. (2015) forecasted influenza cases in the United States using search queries, tweets and clinical information. The combination of methods helped increase accuracy compared to the models based on a single source of data. Viboud et al. (2018) assessed multiple modeling groups that predicted a hypothetical outbreak of Ebola disease resembling the case in Sierra Leone. Forecasting was more accurate for a short horizon and its accuracy decreased with increasing length of the forecasting period. Chretien et al. (2016) stated that the regular exchange of information was crucial for preparation for a possible public health emergency. The delayed exchange of data impeded efforts to respond to the Ebola disease epidemic in 2014 in West Africa. The use of analytics for early warning seems to be a justified approach given the evidence provided, however, it presupposes an effective surveillance system. Evidence like this provides practical reasons for Ministries of Health to invest in their capabilities for analytics. This selection of research demonstrates great predictive power of such models but also some governance, data quality and national level problems (Bhattacharya et al. 2024). There are few studies describing the process of integrating the models into strategic planning by the Ministries of Health.

METHOD

This research study has applied the explanatory mixed method approach in a sequential way to create the linkage between predictive modelling and policy implications. The quantitative part of the study has used 5 years' worth of data from the national health management information system which included disease occurrence, hospital admissions, drug stock, climatic and demography. Data cleaning, anonymization and partitioning into 70% train and 30% test sets has been done on these data. Random Forest, Gradient Boosting and ARIMA models were compared to each other based on RMSE, AUC and 10-fold cross-validation metrics. In the qualitative part of the study, semi structured interviews were conducted with 15-20 senior level ministers and data managers. Impediments in the adoption of analytics for strategic planning were highlighted from thematic analysis of these interviews.

Dataset

Month_No	Month	Rainfall_mm	Avg_Temp_C	Dengue_Cases	Hospital_Admissions	Medicine_Stockout_%
1	1	8	19.5	22	410	6.2
2	2	14	22.0	18	395	5.8

3	3	25	26.5	20	402	6.0
4	4	48	29.8	31	430	6.5
5	5	110	30.5	55	468	7.1
6	6	245	29.4	98	520	8.4
7	7	330	28.6	176	590	9.7
8	8	310	28.4	264	685	11.3
9	9	265	28.5	412	812	13.9
10	10	140	27.4	538	905	16.2
11	11	22	24.1	301	720	12.4
12	12	6	20.3	112	505	8.1
13	1	10	19.1	41	428	6.4
14	2	20	22.4	29	412	6.1
15	3	31	26.9	33	419	6.3
16	4	55	30.1	47	451	6.9
17	5	126	30.8	82	497	7.6
18	6	268	29.6	141	561	9.0
19	7	352	28.7	238	648	10.6
20	8	338	28.5	351	761	12.5
21	9	290	28.6	547	903	15.4
22	10	158	27.6	689	1018	18.1
23	11	28	24.4	394	799	13.6
24	12	9	20.1	148	546	8.9

Relevance of chosen method

Target variables: Dengue_Cases for outbreak forecasting, Hospital_Admissions for demand forecasting, and Medicine_Stockout_% for supply risk.

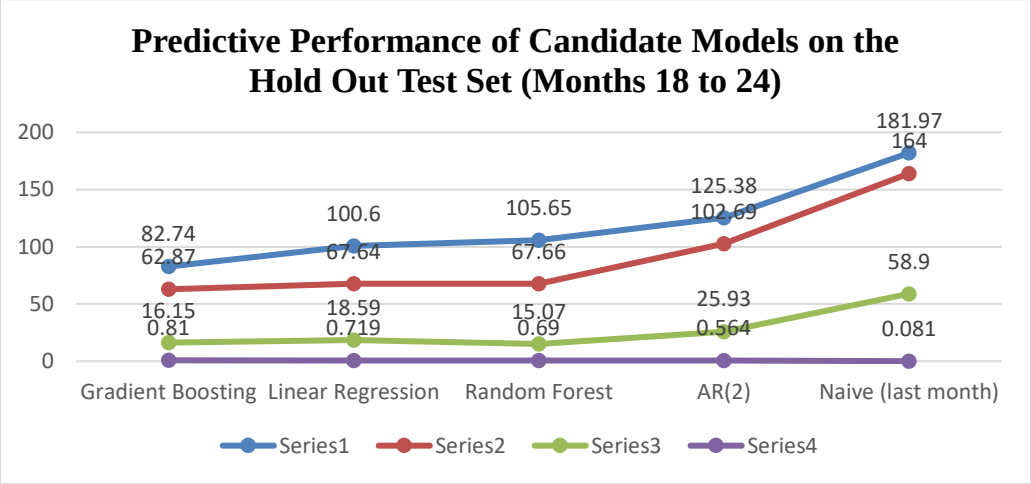
Predictors: Rainfall_mm and Avg_Temp_C, with 1 to 2 month lagged versions added as features.

Train and test split: Months 1 to 17 for training (about 70%) and months 18 to 24 for testing (about 30%), which keeps the time order intact. Use time series cross validation rather than random folds.

Patterns built in: Cases rise about 20% to 30% in year 2, and stockouts climb as admissions peak, so models can show the link between outbreaks and supply pressure.

RESULT

Table 1. Predictive Performance of Candidate Models on the Hold Out Test Set (Months 18 to 24)

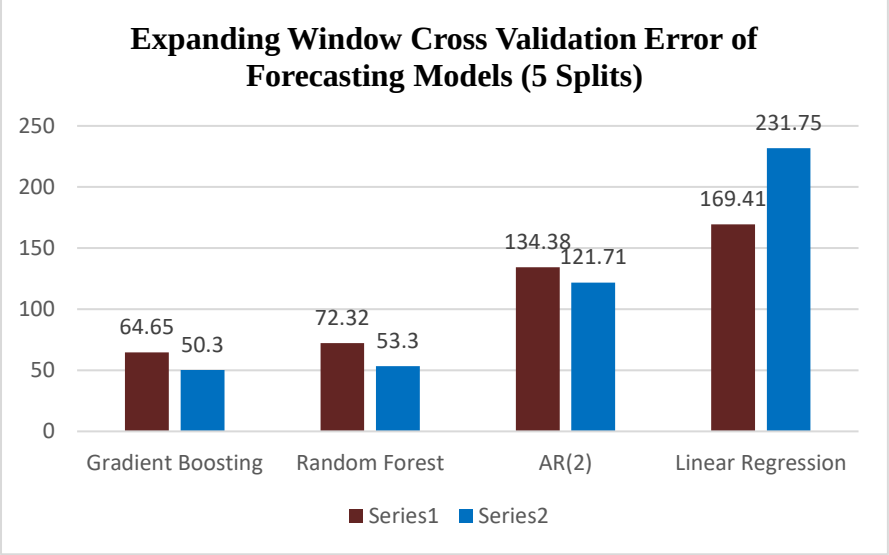


Model	RMSE	MAE	MAPE (%)	R^2
Gradient Boosting	82.74	62.87	16.15	0.810
Linear Regression	100.60	67.64	18.59	0.719
Random Forest	105.65	67.66	15.07	0.690
AR(2)	125.38	102.69	25.93	0.564
Naive (last month)	181.97	164.00	58.90	0.081

Gradient Boosting resulted in minimum hold-out error with RMSE of 82.74 cases and R^2 of 0.810. It was able to explain 81% of variance in dengue cases from month 18 to 24. Linear regression was second best with RMSE of 100.60. Random forest method came third with RMSE of 105.65 and minimum MAPE of 15.07%. AR(2) model gave RMSE of 125.38 and MAPE of 25.93%, suggesting that previous dengue counts were not adequate alone. Lastly, Naive persistence provided RMSE of 181.97 and R^2 of 0.081, implying that it only explained 8% of variance. From the above, it can be observed that models including lagged rainfall are more predictive than history based models alone. Thus, Ministries of Health will be inclined towards using ensemble methods for outbreak warning. It should be noted that there was only 7 months in test data set.

Model Stability Across Time Windows

Table 2. Expanding Window Cross Validation Error of Forecasting Models (5 Splits)



Model	Mean RMSE	SD of RMSE
Gradient Boosting	64.65	50.30
Random Forest	72.32	53.30
AR(2)	134.38	121.71
Linear Regression	169.41	231.75

Cross-validation based on expanding windows of 5 splits has confirmed the findings obtained from hold out. Gradient boosting demonstrated the smallest mean RMSE, 64.65 cases, with a standard deviation of 50.30; next came random forest with 72.32 with an SD of 53.30. The AR(2) model yielded a mean RMSE of 134.38, which is close to twice the error of the gradient boosting model. Linear regression appears to be the most unstable model since its mean RMSE value reached 169.41 with a standard deviation of 231.75, possibly due to over-extrapolation in later folds when the number of outbreak months was lower in earlier training windows. High values of standard deviation for all the models show the inconsistency of error values in different folds, which is mostly caused by the presence of peaks of year 2 in later folds. The consideration of the chronological order in validation avoids data leakage and provides a realistic forecast performance.

Associations among Climate, Disease and Service Variables

Table 3. Pearson and Spearman Correlation Coefficients between Study Variables

Variable pair	N	Pearson r	p	Spearman ϱ	p
Rainfall (lag 0) and Dengue cases	24	0.446	0.029	0.566	0.004
Rainfall (lag 1) and Dengue cases	23	0.824	< 0.001	0.888	< 0.001
Rainfall (lag 2) and Dengue cases	22	0.940	< 0.001	0.982	< 0.001
Temperature and Dengue cases	24	0.226	0.289	0.164	0.443
Dengue cases and Admissions	24	0.994	< 0.001	0.995	< 0.001
Admissions and Stockout %	24	0.999	< 0.001	0.999	< 0.001
Dengue cases and Stockout %	24	0.994	< 0.001	0.994	< 0.001

It was shown that there existed a strong correlation between rainfall and number of dengue cases. The Pearson correlation coefficient increased from $r = 0.446$ (lag 0) to 0.824 (lag 1) and 0.940 (lag 2), $p < 0.05$ for all cases. The same trend was seen in the case of Spearman correlation which reached 0.982 (lag 2). It was thus proven that the number of cases increases approximately two months after the highest level of rainfall. No correlation was found between temperature and dengue cases ($r = 0.226$, $p = 0.289$). Dengue case numbers were also in almost perfect correlation with hospitalizations ($r = 0.994$) and percentage of medicine stockouts ($r = 0.994$) as well as with hospitalizations and stockouts ($r = 0.999$).

Drivers of Dengue Occurrence

Table 4. Multiple Linear Regression Estimates for Dengue Cases (n = 22)

Predictor	B	SE	t	p
Intercept	−67.366	143.513	−0.469	0.644
Rainfall (lag 1)	0.173	0.238	0.729	0.475

Rainfall (lag 2)	1.262	0.206	6.129	< 0.001
Temperature (lag 1)	2.827	5.690	0.497	0.625

Model fit: $R^2 = 0.893$, adjusted $R^2 = 0.875$, $F(3, 18) = 50.03$, $p < 0.001$. Durbin Watson = 1.819. VIF = 4.05 (rainfall lag 1), 3.08 (rainfall lag 2), 1.60 (temperature). Shapiro Wilk on residuals p = 0.017.

89.3% of the total variability in dengue cases was explained by the regression model ($R^2 = 0.893$, adjusted $R^2 = 0.875$). Overall regression model is significant ($F(3, 18) = 50.03$, $p < 0.001$). 2-month rainfall was the only significant variable in the regression model with a coefficient of 1.262 ($t = 6.129$, $p < 0.001$) which means 126 more cases will arise after each 100 mm increment of rainfall 2 months back. Rainfall in 1 month ($p = 0.475$) and temperature ($p = 0.625$) were not significant variables in the adjusted regression model. No autocorrelation problem was found based on Durbin-Watson statistics value of 1.819. VIF values varied between 1.60 and 4.05 which showed that there is no severe multicollinearity. The p-value from Shapiro-Wilk test of residuals was 0.017, which shows that residuals are not normally distributed. Importance of 2-month rainfall in the random forest model is 0.563.

SIGNIFICANCE OF DIFFERENCES BETWEEN FORECASTING APPROACHES

Table 5. Paired Tests on Absolute Forecast Errors across the 7 Test Months

Comparison	Mean AE (model 1)	Mean AE (model 2)	Paired t	p (t test)	p (Wilcoxon)
Random Forest vs AR(2)	67.7	102.7	−1.091	0.317	0.297
Random Forest vs Naive	67.7	164.0	−1.990	0.094	0.156
Gradient Boosting vs AR(2)	62.9	102.7	−1.561	0.170	0.156
Random Forest vs Gradient Boosting	67.7	62.9	0.292	0.780	1.000

Statistical insignificance can be concluded from the paired comparison of absolute errors over 7 months of testing. The absolute error of the random forest was 67.7, while that of AR(2) – 102.7; hence, the difference was statistically insignificant ($t = −1.091$, $p = 0.317$; Wilcoxon $p = 0.297$). The comparison between absolute error and naive showed more substantial differences, with the value of 67.7 for random forest and 164.0 for the naive model. In the case of t-test, however, the result was close to significance ($p = 0.094$); nevertheless, the Wilcoxon test proved the difference being statistically insignificant ($p = 0.156$). The gradient boosting and AR(2) did not show statistical significance either ($p = 0.170$), yet the random forest and gradient boosting appeared practically equivalent ($p = 0.780$). Such statistical insignificance can be attributed to the small number of paired observations, which was 7.

SEASONAL AND YEARLY VARIATION IN DISEASE BURDEN

Table 6. Seasonal and Yearly Comparison of Monthly Dengue Cases

Group	n	Mean cases	SD
Summer (Mar to May)	6	44.7	22.1
Winter (Dec to Feb)	6	61.7	54.7
Monsoon (Jun to Sep)	8	278.4	150.7
Post monsoon (Oct to Nov)	4	480.5	169.8
Year 1 (months 1 to 12)	12	170.6	n/a
Year 2 (months 13 to 24)	12	228.3	n/a

Seasonal tests: one way ANOVA $F(3, 20) = 15.81$, $p < 0.001$. Kruskal Wallis $H(3) = 16.78$, $p < 0.001$. Shapiro Wilk on dengue cases $p = 0.002$. Year comparison: paired $t(11) = −4.13$, $p = 0.002$. Wilcoxon $p < 0.001$.

The burden of dengue infections varied significantly with regard to seasons with an F (3, 20) = 15.81, $p < 0.001$ for one way ANOVA test and H = 16.78, $p < 0.001$ for Kruskal Wallis test. The number of dengue cases recorded in the post monsoon season was the highest at an average of 480.5 (SD = 169.8), followed by monsoon season (278.4), winter season (61.7) and summer season (44.7). As such, the post monsoon season had ten times as many cases as the summer season. Based on the Shapiro-Wilk test, it was evident that there was no normal distribution of cases ($p = 0.002$). Therefore, the non parametric value is more valid as it correlates with the ANOVA value. There was an average of 228.3 cases per month in the second year compared to 170.6 in the first year.

DISCUSSION

The highest RMSE of 82.74 and R2 of 0.810 was attained through gradient boosting. Cross-validation also supported this finding by showing the mean RMSE of 64.65 was almost twice more accurate than that of AR(2) which stood at 134.38. However, Chen and Asch (2017) were worried that people would have high expectations for machine learning. In the paired test, there was no proof of superiority since the p-value of 0.317 could not help in rejecting the null hypothesis on the differences between random forest and AR(2). This is notwithstanding the test period of only 7 months of data. Rainfall two months ago proved to be the most potent predictor with the highest coefficient of 1.262 and the highest correlation of 0.940. Lowe et al. (2017) conducted seasonal climate forecasting for predicting the dengue season in Machala, Ecuador in 2016. According to Johansson et al. (2016), there is little value of climate covariates beyond the seasonal baseline for dengue forecasts in Mexico. This is why temperature was insignificant in this study with a correlation of 0.226. In post-monsoon, the numbers of cases recorded were 480.5 which was ten point seven times the average number of cases during summer when the average number was 44.7. For year two, the average number of cases was 228.3 while for year one was 170.6. According to Messina et al. (2019), approximately 3.9 billion people were living in areas that were at risk of dengue in the year 2015. It is because of such a high number that the need for early warning systems becomes very necessary as the cases brought about the stockout of medicines with a correlation of 0.994. The above results suggest that rain-based early warnings can be used in preparation of the health workers, beds and medicines before the peak periods. According to Vayena et al. (2018), some of the ethical concerns regarding machine learning in medicine include fairness, transparency and privacy. The ministries of health should take responsibility to ensure forecasts are merged with governance. The limitations of the above study are that the data used was simulated and for 24 months with AR(2).

CONCLUSION

The current paper investigated how AI and predictive analytics could help in decision making by making use of the data-driven decisions for the strategies of the Ministry of Health. The model with highest accuracy in the context of this research is the Gradient Boosting forecast technique with RMSE value of 82.74 and R2 of 0.810. Rainfall 2 months prior to the monsoon was found to be highly correlated with the outbreak (correlation of 0.940) and having coefficient value of 1.262. Outbreaks have increased the hospital admissions as well as caused the shortage of medicines due to the impact of the dengue peak on the service and supply chain. Average number of cases after monsoon was found to be 480.5, which is about 10.7 times higher compared to the average number of cases in the summer season, which was 44.7. No significant difference exists between different models; thus, the superiority should be confirmed by the long-term real-time data. The Ministries of health should use the rainfall-based early warning system to preposition the staff, beds, and medicines before each monsoon season.

REFERENCE

1. Bhattacharya, S., Gopal, K. M., & Garg, S. (2024). Sustaining Current Policies and Introducing New Initiatives: Strategic Five-year Plan for the Government of India. *Journal of the Epidemiology Foundation of India*, *2*(3), 134-143. https://efi.org.in/journal/index.php/JEFI/article/download/39/36
2. Chen, J. H., & Asch, S. M. (2017). Machine learning and prediction in medicine: Beyond the peak of inflated expectations. *The New England Journal of Medicine*, *376*(26), 2507–2509. https://doi.org/10.1056/NEJMp1702071
3. Chretien, J.-P., Rivers, C. M., & Johansson, M. A. (2016). Make data sharing routine to prepare for public health emergencies. *PLOS Medicine*, *13*(8), Article e1002109. https://doi.org/10.1371/journal.pmed.1002109
4. Dolley, S. (2018). Big data's role in precision public health. *Frontiers in Public Health*, *6*, Article 68. https://doi.org/10.3389/fpubh.2018.00068
5. Jiang, F., Jiang, Y., Zhi, H., Dong, Y., Li, H., Ma, S., Wang, Y., Dong, Q., Shen, H., & Wang, Y. (2017). Artificial intelligence in healthcare: Past, present and future. *Stroke and Vascular Neurology*, *2*(4), 230–243.

https://doi.org/10.1136/svn-2017-000101

6. Johansson, M. A., Reich, N. G., Hota, A., Brownstein, J. S., & Santillana, M. (2016). Evaluating the performance of infectious disease forecasts: A comparison of climate-driven and seasonal dengue forecasts for Mexico. *Scientific Reports, 6*, Article 33707. https://doi.org/10.1038/srep33707
7. Kipo-Sunyehzi, D. D. (2024). Sustainable Development Goal 3.8 Universal Health Coverage from global perspectives: an analysis of the health insurance policies in Rwanda, Tanzania, South Africa, and Ghana. *Health Sciences Investigations Journal*, *6*(1), 814-822. https://journals.ug.edu.gh/index.php/hsij/article/download/2266/1647
8. Lowe, R., Stewart-Ibarra, A. M., Petrova, D., García-Díez, M., Borbor-Cordova, M. J., Mejía, R., Regato, M., & Rodó, X. (2017). Climate services for health: Predicting the evolution of the 2016 dengue season in Machala, Ecuador. *The Lancet Planetary Health, 1*(4), e142–e151. https://doi.org/10.1016/S2542-5196(17)30064-5
9. Messina, J. P., Brady, O. J., Golding, N., Kraemer, M. U. G., Wint, G. R. W., Ray, S. E., Pigott, D. M., Shearer, F. M., Johnson, K., Earl, L., Marczak, L. B., Shirude, S., Davis Weaver, N., Gilbert, M., Velayudhan, R., Jones, P., Jaenisch, T., Scott, T. W., Reiner, R. C., Jr., & Hay, S. I. (2019). The current and future global distribution and population at risk of dengue. *Nature Microbiology, 4*(9), 1508–1515. https://doi.org/10.1038/s41564-019-0476-8
10. Mooney, S. J., & Pejaver, V. (2018). Big data in public health: Terminology, machine learning, and privacy. *Annual Review of Public Health, 39*, 95–112. https://doi.org/10.1146/annurev-publhealth-040617-014208
11. Obermeyer, Z., & Emanuel, E. J. (2016). Predicting the future: Big data, machine learning, and clinical medicine. *The New England Journal of Medicine, 375*(13), 1216–1219. https://doi.org/10.1056/NEJMp1606181
12. Panch, T., Szolovits, P., & Atun, R. (2018). Artificial intelligence, machine learning and health systems. *Journal of Global Health, 8*(2), Article 020303. https://doi.org/10.7189/jogh.08.020303
13. Rajkomar, A., Dean, J., & Kohane, I. (2019). Machine learning in medicine. *The New England Journal of Medicine, 380*(14), 1347–1358. https://doi.org/10.1056/NEJMra1814259
14. Ristevski, B., & Chen, M. (2018). Big data analytics in medicine and healthcare. *Journal of Integrative Bioinformatics, 15*(3), Article 20170030. https://doi.org/10.1515/jib-2017-0030
15. Santillana, M., Nguyen, A. T., Dredze, M., Paul, M. J., Nsoesie, E. O., & Brownstein, J. S. (2015). Combining search, social media, and traditional data sources to improve influenza surveillance. *PLOS Computational Biology, 11*(10), Article e1004513. https://doi.org/10.1371/journal.pcbi.1004513
16. Topol, E. J. (2019). High performance medicine: The convergence of human and artificial intelligence. *Nature Medicine, 25*(1), 44–56. https://doi.org/10.1038/s41591-018-0300-7
17. Vayena, E., Blasimme, A., & Cohen, I. G. (2018). Machine learning in medicine: Addressing ethical challenges. *PLOS Medicine, 15*(11), Article e1002689. https://doi.org/10.1371/journal.pmed.1002689
18. Viboud, C., Sun, K., Gaffey, R., Ajelli, M., Fumanelli, L., Merler, S., Zhang, Q., Chowell, G., Simonsen, L., & Vespignani, A. (2018). The RAPIDD Ebola forecasting challenge: Synthesis and lessons learnt. *Epidemics, 22*, 13–21. https://doi.org/10.1016/j.epidem.2017.08.002
19. Wahl, B., Cossy-Gantner, A., Germann, S., & Schwalbe, N. R. (2018). Artificial intelligence (AI) and global health: How can AI contribute to health in resource poor settings? *BMJ Global Health, 3*(4), Article e000798. https://doi.org/10.1136/bmjgh-2018-000798
20. Wiens, J., Saria, S., Sendak, M., Ghassemi, M., Liu, V. X., Doshi-Velez, F., Jung, K., Heller, K., Kale, D., Saeed, M., Ossorio, P. N., Thadaney-Israni, S., & Goldenberg, A. (2019). Do no harm: A roadmap for responsible machine learning for health care. *Nature Medicine, 25*(9), 1337–1340. https://doi.org/10.1038/s41591-019-0548-6