A Machine Learning-Based algorithm for the assessment of clinical metabolomic fingerprints in Zika virus disease

Enrique Peláez Jarrin
Facultad de Ingeniería en Electricidad y Computación
Escuela Superior Politécnica del Litoral (ESPOL)
Guayaquil, Ecuador
epelaez@espol.edu.ec

Fernanda Bertuccez Cordeiro
Laboratorio para Investigaciones Biomedicas
Facultad de Ciencias de la vida
Escuela Superior Politécnica del Litoral (ESPOL)
Guayaquil, Ecuador
fbertuc@espol.edu.ec

Washington Cárdenas Medranda
Laboratorio para Investigaciones Biomedicas
Facultad de Ciencias de la vida
Escuela Superior Politécnica del Litoral (ESPOL)
Guayaquil, Ecuador
wbcarden@espol.edu.ec

Michael Barrett
Centre for Integrative Parasitology, Institute of Infection, Immunity & Inflammation
University of Glasgow
Glasgow, UK
Michael.Barrett@glasgow.ac.uk

Mildred Zambrano
Servicio de Infectología e Epidemiologías
Hospital de Niños, Dr. Roberto Gilbert
Guayaquil, Ecuador
mzambranol@jbgye.org.ec

Mary Regato
Instituto Nacional de Investigación en Salud Pública (INSPI)
Guayaquil, Ecuador
mregato@inspi.gob.ec

***Abstract*—Data analysis for metabolomic studies is challenging considering the number of statistical tools and standardization processes, which provides different results and projection in a single study. In addition, generation of high complexity data is common for untargeted metabolomics, requiring careful analysis and interpretation of results. In order to propose an innovative method for the analysis of a mass spectrometry metabolomics dataset, data from a Zika virus study was used. The analysis of this dataset combined principal component analysis and supervised learning methods such as support vector machines and logistic regression, to provide a truthful prediction model for discriminating samples of individuals with Zika virus infection and healthy controls. These supervised methods were used to learn the features that configured the "fingerprint" for the viral infection, showing over 98% of accuracy in a validation set. This model could be used as a fast and reliable test for determining Zika virus infections as part of healthcare services. Furthermore, this novel method shows potential for diagnosing other arboviral diseases.**

Keywords—Machine Learning, PCA, SVM, Logistic Regression, Zika virus

I. Introduction

Applications of machine learning (ML) arise at the intersection of multivariate statistics, associating data analysis and computer science, aiming to develop efficient computing algorithms [1]. As a result, this association directs to the development of ML models from massive data sets, which can be done via supervised and unsupervised learning [2]. In practice, ML techniques are enabling the analysis of spectral data extracted from large number of samples obtained from several groups. During training and testing ML models can be continuously fed with data representing the features that will allow effective discrimination between samples and groups.

With respect to modern applications, ML have been applied to different techniques models, such as support vector machines and logistic regression, which have been used in several diagnostic applications: imaging analysis [3]; cancer diagnosis [4-6]; diagnosis of diabetes [7] and others [8]. Considering arbovirus infections, these are mostly represented by Dengue, Chikungunya and Zika virus and ML approaches have been directed to epidemiological studies [9,10].

Among the arboviruses infections, the Zika virus (ZIKV) infection is characterized for being either asymptomatic or associated with fever, headache, exanthema, joint pains and malaise [11]. The infection by ZIKV has been on the spotlight in South America since its outbreak in May of 2015, when the infection was firstly associated with newborn congenital microcephaly [12]. By 2016, more than 1 million people had been infected, and by April of 2016, ZIKV transmissions were registered in 27 countries in the Americas [13]. Additionally, ZIKV infections were globally reported by mid-2018, with more than 220.000 confirmed cases worldwide. However, the number of infections may be higher, because around 80% of infections are asymptomatic and many communities lack suitable diagnostic methods during early outbreaks [14].

Current diagnosis for ZIKV infections is challenging due to several factors, such as overlapping symptoms with other arboviral diseases, and limitations in sensitivity of genetic and serological diagnostics [15-17]. Therefore, development of an effective diagnostic tool could contribute for both identification of the infection agent and patient health status, specially in ZIKV infection that can progress to Guillain–Barré syndrome and affect fetus development by causing congenital microcephalia [18,19].

Mass spectrometry-based techniques have been used for diagnostic purposes in a wide spectrum of biological samples,

978-1-7281-5666-8/19/$31.00 ©2019 IEEE

which produce a rich variety of datasets for statistical analysis [20]. Together with recent computational capabilities that use sophisticated data processing algorithms, these high-throughput techniques are capable of correctly differentiate conflicting conditions [21].

Generating rich datasets is the starting point for current data analysis techniques, which learn from "positive" and "negative" real examples; such as, biological samples from infected and healthy individuals. Learning from this data will allow us to discover which are the specific spectral signatures of a particular biological condition based on their intrinsic differences, even if they are imperceptible to humans [22]. Such differences would precisely allow us to learn the particular features, such as the "fingerprint" that makes a group behavior; features that further will compose the model that will be used to recognize the occurrence of the learned patterns in new unseen data.

In this research, ML modeling was applied to a metabolomics dataset that compared patients with ZIKV infection and their health control counterparts. Prior to learning the classification model, we performed a feature importance analysis to extract and to isolate the most important components, which will be used to identify the presence (or not) of endogenous signatures from either ZIKV infected or healthy patient's metabolic fingerprinting. This paper describes the dataset, including the experimental design for the LC-MS metabolomics. We also discuss the machine learning models proposed for classifying the samples and its mathematical formulation.

II. The Dataset

A. Ethics Statement

The present study has approval from the Ethics Committee of the Hospital General Luis Vernaza, from Guayaquil, Ecuador. A written informed consent was obtained from all participants.

B. Study Design

This is a prospective study that included 20 patients between 18 and 40 years old. All patients were clinically evaluated to determine their health status, considering that excluded cases refer to patients with confirmed diagnosis for overlapping infections, such as Influenza, Dengue fever, Chikungunya fever, and other viruses' infections. Molecular diagnosis was used to confirm infection by Dengue and Chikungunya fever. Eligible patients were divided into 2 groups: The ZIKV group (n=10), with patients who presented clinical suspicion and positive molecular diagnosis of ZIKV infection; and, the Healthy Control group (HC; n=10), which included healthy volunteers that did not present fever or symptoms of other viral infection for at least 3 months.

C. Metabolomics

The samples were submitted to metabolites extraction based on the protocol stablished by the Glasgow Polyomics (University of Glasgow, United Kingdom). For liquid chromatography–mass spectrometry (LC-MS) matabolomics, each sample was randomly run on a ZIC-HILIC (hydrophilic interaction chromatography) column (SeQuant) coupled to the Orbitrap Q Exactive mass spectrometer (Thermo Scientific), as previously described [23].

III. Machine Learning Models

The proposed machine learning models for ZIKV infection detection have the following phases:

1. *Pre-processing of data*: Before performing the classification, data from the Zika group (positive examples or with ZIKV) and from the Control group (negative examples or patients without ZIKV) are standardized, normalized, merge and randomly divided into two subsets: Training and Test.

   - 70% of the patients will be assigned for training; this set will be called: The train set. This portion will be further divided into 2 sub-sets: 80% for training and the other 20% for cross-validating the process of determining most discriminant features for Zika detection, as well as for controlling overfitting.
   - The rest, 30% of the dataset will be reserved untouched for blind testing the models; the set will be called: The test set. This approach will allow us to evaluate the classifier and check for under or overfitting the available data.

   For possible variations during the training and validation process, we repeat this process 10 times, referred as epochs, and report the average performance and standard deviation for the validation set.

   In this phase we have also assigned the output labels to each sample of each group:

$$Labels = \{l_i\},\ \ l_i = [0, 1] \tag{1}$$

   - 1 represents the Zika group, corresponding to positive examples or with ZIKV, and
   - 0 represents the Control group, the negative examples or patients without ZIKV.

   We need a well-defined range of feature values for feeding the ML models; hence, we normalized the combined relative intensity with the corresponding spectral mass of each sample. For normalizing the merged vectors of the samples, we use the relative mass spectral-intensity of each vector and divide each vector element by their maximum value, as defined in the following equation:

$$nv_{i,j} = \frac{mv_{ij}}{max_{(mv_{i,j=1:k})}},\ \ i \in [1 \dots N]; j = \in [1 \dots k] \tag{2}$$

   Where N is the total number of samples in the dataset (Control and Zika groups); k is the total number of m/z mass measurements; nv is the normalized value of measurement j of the correspondent patient i; and, mv is the merged mass spectral-intensity value.

   Reduction of number of features through Principal Component Analysis:

   Our dataset has thousands of features, which makes the modeling challenging; the inspection made to the dataset showed that there were values missing, which could make our model less skillful. Since it is hard to know which features of the dataset are relevant and which are not, we will perform a dimensionality reduction using principal

978-1-7281-5666-8/19/$31.00 ©2019 IEEE

component analysis (PCA) in the process of training our models.

Once the projection on the principal components is calculated, we can apply this transformation to the train and test sets, which can be projected into a subspace with the principal components or dimensions. For extracting the principal components, we have standardized the data before using the PCA method available in the Scikit-learn library for Python. After reducing the dimensions, the explained_variance_ratio_ function from Scikit-learn, reported that more than 97% of information is kept into the 2 principal components, hence, we will use them for performing classification: `array([0.92461621, 0.05301557])`

2. *Determination of the classifying boundaries*: After dimensionality reduction, we use the 2 principal components to perform the classification; we will use the following two classifiers and compare their results: Support Vector Machines (SVM) and Logistic Regression (LR).

   2.1 *Support Vector Machines*: An SVM is a flexible nonparametric machine learning algorithm and it will be used in this work for binary classification to determine if a sample corresponds to an infected or not-infected group. The goal will be to find the values of the hyperplane's parameters that best separates the 2 classes. To discover these parameters, we will use the Sequential Minimal Optimization approach. In this method a random training sample is selected in each iteration and used to update the parameters.

   For learning the hyperplane that best separates the 2 groups, we will use two different update procedures, depending on the output value. If the output value is > 1, it suggests that the training sample was not a support vector, hence the instance was not directly involved in calculating the output, in which case the parameter is slightly decreased, using formula (3):

$$w = \left(1 - \frac{1}{t}\right) \times w \tag{3}$$

   Where: w is the parameter that is being learned, t is the current iteration (e.g. 1 for the first update, 2 for the second and so on).

   If the output is < 1 then it is assumed that the training instance is a support vector and must be updated, to better explain the data using the following formula:

$$w = \left(1 - \frac{1}{t}\right) \times w + \frac{1}{\lambda \times t} \times (y \times x) \tag{4}$$

   Where: w is the parameter that is being updated, t is the current iteration and λ is the regularization parameter, which was set to 0.01 in our experiments. This regularization parameter serves as a degree of importance given to miss-classifications. Since SVM looks for maximizing the margin between both classes and minimizing the amount of miss-classifications, and as we can see in Fig. 1, there are a couple of non-separable samples in each group, in order to find a solution, the miss-classification constraint has been relaxed, and this is done by setting the mentioned "regularization" parameter.

   2.2 *Logistic Regression*: The LR model takes real-valued inputs, as the intensity with the corresponding spectral mass in our dataset and makes a prediction as to the probability of the input belonging to the class 1 (Sika group). This learning algorithm also has the objective of discovering the best parameters of the boundary line based on the training data. In this case, we will estimate the values of the parameters using Stochastic Gradient Descent, and for updating the parameter values we use equation (5).

$$w = w + \alpha \times (y - prediction) \times prediction \times (1 - prediction) \times x \tag{5}$$

   Where: w is the parameter we are updating and $prediction$ is the output of making a prediction using this model; α is the learning rate stablished at the beginning of the training. We are using an α value of 0.01.

   Logistic regression assumes no error in the output variable (y), hence we have tested the outputs removing the outliers from the training data, the results show that indeed the accuracy improves, as it is shown in Fig. 1 and Fig. 2.

3. *The classifier*: Once the principal components were discovered, we trained the SVM and LR classifiers using the selected components only, which allows us to simplify the process with a subset of the projected features. As seen in Fig. 3, a line perfectly separates the 2 classes, hence we used a linear kernel for the SVM model. To train the classifiers we used the data from the training set; later tested with the blind-test data from the testing set; and, then report the result for Zika detection. During training we have used cross-validation as an approach to estimate the performance of the learning algorithm with less variance than a single train-test set split. We have set 5 folds. After running cross-validation we end up with k different performance scores that we summarized using a mean and a standard deviation, which are reported in Table 1.

IV. Results and Discussion

Metabolomics has expanded considerably over the last two decades and the amount of data being generated by these experiments keeps increasing. In addition, different tools have been developed for the analysis of such amount of data, which comprehends development of workflow, data processing, machine learning modeling, and interpretation of results [21]. In the present study, PCA allowed us to visualize the high-dimensional dataset (1162 *m/z* measurements per patient) by projecting the two principal components onto a two-dimensional space, in which we can see that it is possible to learn a hyperplane that cleanly divides the samples in 2 classes, class 0 as the non-infected class and class 1 as the Sika virus group, before pre-processing the data as shown in Fig. 1.

Regarding clinical metabolomics for the study of ZIKV infection, current literature demonstrates group separation in a similar study, considering data analysis was conducted by

978-1-7281-5666-8/19/$31.00 ©2019 IEEE

orthogonal partial least square-discriminant analysis (OPLS-DA) [22].

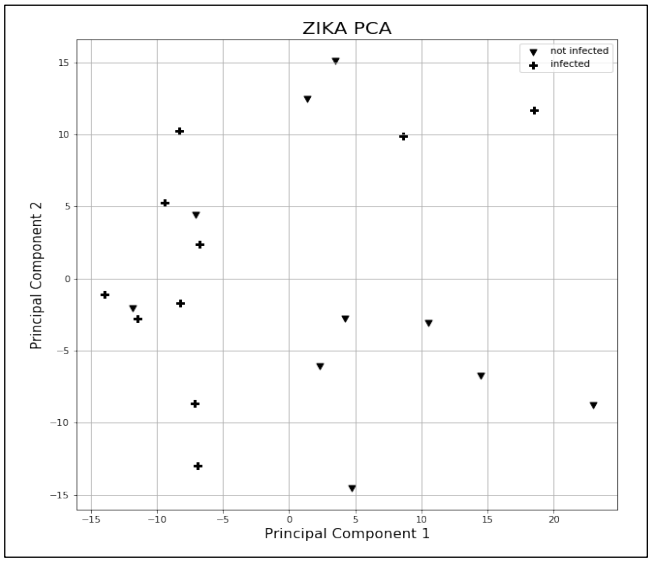

Fig. 1. Two Principal Components extracted from the train set, before pre-processing. Squares indicate healthy controls, whereas triangles indicate ZIKV infection

PCA indeed was used to evaluate neuroinflammation driven by ZIKV, showing that metabolite profiles were different according with time of macroglia cells infection [24]. Therefore, our results indicate new insights for the development of statistical tools that assist translation medicine. Fig. 2, shows the principal components after removing the outliers from the training data.

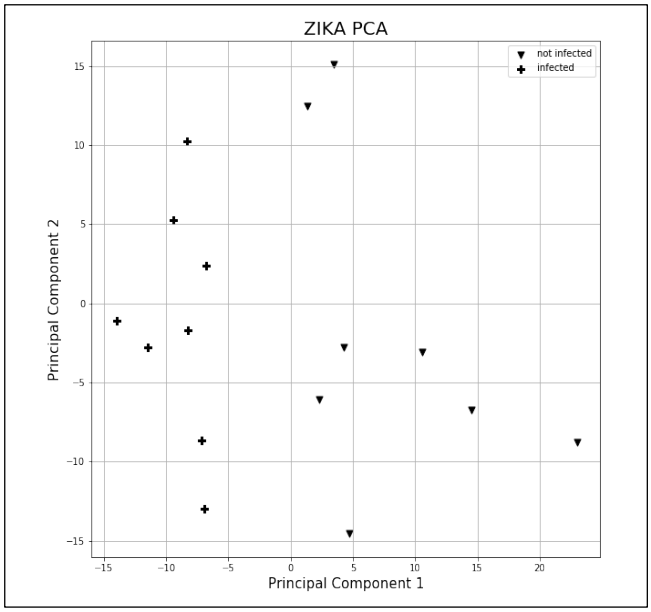

Fig. 2. Principal Components after removing the outliers from the training data. Squares indicate healthy controls, whereas triangles indicate ZIKV infection

Since the 2 principal components, out of the 1162-dimensions, kept more than 97% of the information, the SVM classifier can easily separate the samples as infected and not infected, or ZIKV and HC groups, as shown in the plot of the principal components in Fig. 3. The chart indicates a proper sample classification for both groups.

For the LR classifier case, as shown in Fig. 4, there is also a perfect separation of the classes using the two principal components of the training set.

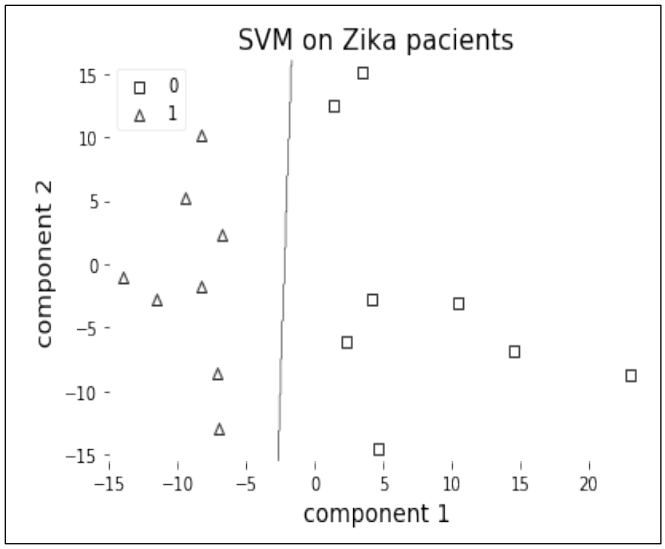

Fig. 3. Decision boundary between Zika infected and not infected ccontrol groups samples using an SVM classifier. Squares indicate healthy controls, whereas triangles indicate ZIKV infection

A. Comparing the two Classifiers

We have used a mixture of one linear classifier, such as the Logistic Regression, and one nonlinear algorithm, such as the Support Vector Machine with a linear kernel. Before each run, we reset a random number seed to ensure that the evaluation of each algorithm was performed using exactly the same data splits and it ensures the results are directly comparable. Table 1 shows that accuracy for LR is 0.967 or 96.7% and for the SVM is 0.98 or 98%. The SVM classifier performs best on the selected main features, thus justifying its use.

TABLE I. PERFORMANCE ACCURACY COMPARISON OF THE TWO CLASSIFIERS USING MACHINE LEARNING

Accuracy:	Mean accuracy	STD
LR	0.966667	0.020825
SVM	0.981667	0.015000

The algorithms were trained and evaluated multiple times on different data and cross-validation provided reliable results of the performance of the algorithm on new data. The choice of k, the number of folds, allowed the size of each test partition had the necessary sample size for the training, whilst allowing enough repetitions of the train-test evaluation of the algorithm to provide a fair estimate of the algorithm's performance on unseen data.

978-1-7281-5666-8/19/$31.00 ©2019 IEEE

TABLE II. BREAKDOWN OF EACH CLASS BY PRECISION, RECALL, F1-SCORE AND SUPPORT VALUES FOR THE SVM MODEL

Group	Precision	Recall	f1-score	support
0 (HC)	1	0.95	0.97	3
1 (SIKA)	0.93	1	0.96	3
Accuracy			0.98	6
macro avg	0.97	0.98	0.97	6
weighted avg	0.9	0.97	0.96	6

The SVM classifier performed better than the LR. As seen in table 2, which reports the performance metrics for the SVM model, it was able to identify almost all cases, from the test set, labeled as not infected (recall 0.95), and from these cases identified as not infected, all of them indeed belong to the HC group (precision 1.0). The model also has the ability to recall all cases (recall 1.0) which were labeled as infected, and from these cases identified as infected, 93% (precision 0.93) of them indeed belong to the Zika group. Finally, the f1 score tells us that the SVM classifier is 97% accurate in identifying not infected cases as compared to all other cases; and, 96% accurate in identifying the Zika cases as compared to all other cases, as reported in Table 2 above.

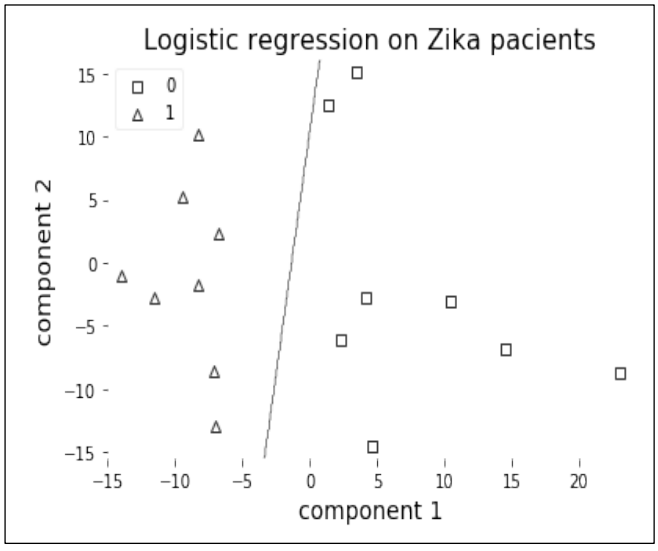

Fig. 4. Separation between Zika and Control group samples using an LR classifier. Squares indicate healthy controls, whereas triangles indicate ZIKV infection

B. Computing Performance Metrics

All experiments were performed using a Macbook Pro, 2.8 GHz Intel Core i7, 16 GB 2133 MHz LPDDR3 memory, 1 TB HD. Programs were written in Python 3.7 and we used the Scikit-learn library for pre-processing data, as well as for PCA, and classification with LR and SVM. Pre-processing the data did take less than a minute after we prepared the algorithms for decomposing the principal components and visualized the trends of the data. Training the ML algorithms using cross-validation and 5 folds, as well as its visualization took about 32 minutes for the whole batch (considering the one thousand one hundred sixty-two different measurements per patient). Once the model and its parameters were learned, the time to analyze a new feature vector, not seen before of a patient at prediction time was less than a second.

V. CONCLUSIONS

This preliminary study, based upon Zika metabolomic data, has shown outstanding results. Although the population of samples was limited, with this proposed methodology it is possible to envision a breakthrough technique in disease diagnosis tests, with potential to be used in further studies for biodiscovery.

ML algorithms take care of extracting discriminative fingerprints for the condition of interest. The objective is that for any given patient with an unknown disease, we can use new samples to multiple classifiers simultaneously, with a fast and reliable response to potential diagnostics. Additionally, after PCA, SVM shows high accuracy for classifying (98%), which reinforces its statistical potential for a real-time embedded SVM-based diagnosis system. This approach is clinically based on patients' response to ZIKV infection instead of virus detection and can be used in primary care for early diagnosis and prognosis.

REFERENCES

[1] Hastie, T., Tibshirani, R., Friedman, J. The Elements of Statistical Learning. New York, NY: Springer Science & Business Media; 2009.

[2] Deo RC. Machine Learning in Medicine. Circulation. 2015 Nov 17;132(20):1920-30. doi: 10.1161/CIRCULATIONAHA.115.001593.

[3] Marleen de Bruijne (2016). Machine learning approaches in medical image analysis: From detection to diagnosis. Medical Image Analysis 33, 94–97

[4] Zheng, B., Yoon, S. W., and Lam, S. S. (2014). Breast cancer diagnosis based on feature extraction using a hybrid of K-means and support vector machine algorithms. Expert Syst. Appl. 41, 1476–1482.

[5] Murata T, Yanagisawa T, Kurihara T, Kaneko M, Ota S, Enomoto A, Tomita M, Sugimoto M, Sunamura M, Hayashida T, Kitagawa Y, Jinno H. Salivary metabolomics with alternative decision tree-based machine learning methods for breast cancer discrimination. Breast Cancer Res Treat. 2019 Jul 8. doi: 10.1007/s10549-019-05330-9.

[6] Gao Q, Su X, Annabi MH, Schreiter BR, Prince T, Ackerman A, Morgas S, Mata V, Williams H, Lee WY. Application of Urinary Volatile Organic Compounds (VOCs) for the Diagnosis of Prostate Cancer. Clin Genitourin Cancer. 2019 Jun;17(3):183-190. doi: 10.1016/j.clgc.2019.02.003.

[7] Iyer, A., Jeyalatha, S. and Sumbaly, R. (2015) Diagnosis of Diabetes Using Classification Mining Techniques. International Journal of Data Mining & Knowledge Management Process (IJDKP), 5, 1-14.

[8] Wang J, Yan D, Zhao A, Hou X, Zheng X, Chen P, Bao Y, Jia W, Hu C, Zhang ZL, Jia W. Discovery of potential biomarkers for osteoporosis using LC-MS/MS metabolomic methods. Osteoporos Int. 2019 Jul;30(7):1491-1499. doi:10.1007/s00198-019-04892-0.

[9] Han BA, Majumdar S, Calmon FP, Glicksberg BS, Horesh R, Kumar A, Perer A, von Marschall EB, Wei D, Mojsilović A, Varshney KR. Confronting data sparsity to identify potential sources of Zika virus spillover infection among primates. Epidemics. 2019 Jun;27:59-65. doi:10.1016/j.epidem.2019.01.005.

[10] Guo P, Liu T, Zhang Q, Wang L, Xiao J, Zhang Q, Luo G, Li Z, He J, Zhang Y, Ma W. Developing a dengue forecast model using machine learning: A case study in China. PLoS Negl Trop Dis. 2017 Oct 16;11(10):e0005973. doi:10.1371/journal.pntd.0005973.

[11] Rather IA, Lone JB, Bajpai VK, Park Y-H. Zika Virus Infection during Pregnancy and Congenital Abnormalities. Frontiers in Microbiology. 2017; 8:581.

978-1-7281-5666-8/19/$31.00 ©2019 IEEE

[12] Hazin AN, Poretti A, Turchi Martelli CM, Huisman TA; Microcephaly Epidemic Research Group., Di Cavalcanti Souza Cruz D, Tenorio M, van der Linden A, Pena LJ, Brito C, Gil LH, de Barros Miranda-Filho D, Marques ET, Alves JG. Computed Tomographic Findings in Microcephaly Associated with Zika Virus. N Engl J Med. 2016 Jun 2;374(22):2193-5

[13] Paixão E. S., Barreto F., Teixeira M. d. G., Costa M. d. C., Rodrigues L. C. (2016). History, epidemiology, and clinical manifestations of Zika: a systematic review. Am. J. Public Health 106, 606–612

[14] Organization PAH (2018). Zika Cumulative Cases. Available online at: https://www.paho.org/hq/index.php?option=com_content&view=article&id=12390:Zika-cumulative-cases&Itemid=42090&lang=en

[15] Edwards T, Del Carmen Castillo Signor L, Williams C, Larcher C, Espinel M, Theaker J, Donis E, Cuevas LE, Adams ER. Analytical and clinical performance of a Chikungunya qRT-PCR for Central and South America. Diagn Microbiol Infect Dis. 2017 89:35-39.

[16] Schuler-Faccini, L., Ribeiro, E. M., Feitosa, I. M. L., Horovitz, D. D. G., Cavalcanti, D. P., Pessoa, A., et al. (2016). Possible association between Zika virus infection and microcephaly – Brazil, 2015. MMWR Morb. Mortal. Wkly. Rep. 65, 59–62.

[17] Pardee, K., Green, A. A., Takahashi, M. K., Braff, D., Lambert, G., Lee, J. W., et al. (2016). Rapid, low-cost detection of Zika virus using programmable biomolecular components

[18] Cao-Lormeau V.-M., Blake A., Mons S., Lastère S., Roche C., Vanhomwegen J., Dub T., Baudouin L., Teissier A., Larre P., Vial A.-L., Decam C., Choumet V., Halstead S.K., Willison H.J., Musset L., Manuguerra J.-C., Despres P., Fournier E., Mallet H.-P., Musso D., Fontanet A., Neil J., Ghawché F. Guillain-Barré syndrome outbreak associated with Zika virus infection in French Polynesia: a case-control study. Lancet. 2016.

[19] Hennessey M, Fischer M, Staples JE. Zika virus spreads to new areas—region of the Americas, May 2015–January 2016. MMWR Morb Mortal Wkly Rep. 2016; 65:55–8.

[20] Takayama, T., Tsutsui, H., Shimizu, I., Toyama, T., Yoshimoto, N., Endo, Y., et al. (2016). Diagnostic approach to breast cancer patients based on target metabolomics in saliva by liquid chromatography with tandem mass spectrometry. Clin. Chim. Acta 452, 18–26.

[21] Beirnaert C, Peeters L, Meysman P, Bittremieux W, Foubert K, Custers D, Van der Auwera A, Cuykx M, Pieters L, Covaci A, Laukens K. Using Expert Driven Machine Learning to Enhance Dynamic Metabolomics Data Analysis. Metabolites. 2019 Mar 20;9(3). pii: E54. doi:10.3390/metabo9030054.

[22] Melo, C. F. O. R., Delafiori, J., Oliveira, D. N., Guerreiro, T. M., Esteves, C. Z., Lima, E. O., et al. (2017). Serum metabolic alterations upon Zika infection. Front. Microbiol. 8:1954.

[23] Vincent IM, Creek DJ, Burgess K, Woods DJ, Burchmore RJ, Barrett MP. Untargeted metabolomics reveals a lack of synergy between nifurtimox and eflornithine against Trypanosoma brucei. PLoS Negl Trop Dis. 2012;6(5):e1618.

[24] Diop F, Vial T, Ferraris P, Wichit S, Bengue M, Hamel R, Talignani L, Liegeois F, Pompon J, Yssel H, Marti G, Missé D. Zika virus infection modulates the metabolomic profile of microglial cells. PLoS One. 2018 Oct 25;13(10):e0206093. doi: 10.1371/journal.pone.0206093.

978-1-7281-5666-8/19/$31.00 ©2019 IEEE