

Application of Soft Computing Techniques in the Diagnosis of Tropical Diseases: A Systematic Review

Samuel B. Oyong¹, Faith-Michael E. Uzoka^{2*}, Okure U. Obot¹, Etim Edem Ekong¹, Pius U. Ejodamen¹

¹Department of Computer Science, University of Uyo, Uyo, Nigeria; ²Department of Mathematics and Computing, Mount Royal University, Calgary, Canada

ABSTRACT

The complex nature of tropical disease variants with confusable symptoms has led to prescription errors and consequentially, many deaths. Soft-computing techniques have been proposed to handle vagueness and imprecision in the diagnosis process. This paper presents a systematic review on the use of soft-computing techniques in the diagnosis of tropical diseases. The results show that Africa, Asia-Pacific and European countries have done more research in areas that focus on the use of soft computing techniques in the diagnosis of tropical diseases, followed by the Americas. Of the twelve (12) tropical diseases commonly researched, malaria, dengue fever, skin diseases, and typhoid fever top the list. The soft-computing classifiers were evenly distributed between single and hybrid paradigms. Most of the classification engines were based on fuzzy logic (15), neural network (5), support vector machine (4), and decision tree (4). Some of the systems were not deployed and most health organizations (including WHO) have not fully embraced the use of soft-computing systems on medical diagnosis. There are many changes taking place in medical diagnosis including; resistance to drugs, high cost of manufacturing hybridized drugs to fight disease resistance and unstructured data gathering that is not suitable for traditional soft computing classifiers. To this end, it is recommended that the WHO and its allies need to diversify their policy on disease control and eradication by including soft computing techniques, especially hybridization (ensemble) methods. This inclusion will reduce morbidity and mortality rates in affected regions.

Keywords: Tropical diseases; Diagnosis; Soft computing techniques; Classifiers; Decision support; Systematic literature review

INTRODUCTION

Tropical diseases are mostly infectious and thrive in hot and humid conditions. Examples include malaria, leishmaniasis, typhoid fever, schistosomiasis, yellow fever, onchocerciasis, lymphatic filariasis, ebola, chagas disease, chicken pox, African trypanosomiasis, and dengue. These conditions are of concern to physicians and the community at large because of their morbidity and mortality rates. These diseases also prevent physicians from making an accurate early diagnosis since they often present overlapping symptoms, and thus become 'confusable' [1]. It is also noted that tropical diseases present ambiguous symptoms that are difficult to handle, especially in

situations where medical practitioners are inexperienced or are in short supply and pressured [2]. This confusable presentation of symptoms often leads to inaccurate diagnosis and therapy, and as a result, fatal complications [3]. Soft-computing techniques have been useful in handling vagueness and symptom overlaps associated with the diagnosis of tropical conditions. Diagnosis takes many forms, some of which are laboratory based; examples include: smear microscopy, radiography, molecular methods, point of care technology, amniocentesis, blood analysis, gastric fluid analysis, kidney function test, liver function test, pregnancy test, prenatal test, syphilis test, and toxicology test [4].

*Correspondence to: Faith-Michael E. Uzoka, Department of Mathematics and Computing, Mount Royal University, Calgary, Canada, Telephone: +14039191429; E-mail: fuzoka@mtroyal.ca

Received date: July 27, 2020; Accepted date: August 13, 2020; Published date: August 20, 2020

Citation: Oyong SB, Uzoka FME, Obot OU, Ekong EE, Ejodamen PU (2020) Application of Soft Computing Techniques in The Diagnosis of Tropical Diseases: A Systematic Review. J Trop Dis 8:358. doi: 10.35248/2329-891X.20.8.358.

Copyright: © 2020 Oyong SB, et al. This is an open-access article distributed under the terms of the Creative Commons Attribution License, which permits unrestricted use, distribution and reproduction in any medium, provided the original author and source are credited.

In this paper, we reviewed the soft computing techniques used in the diagnosis of tropical diseases. Soft computing techniques are used to solve many real-world problems, including those in the medical domain. The key emphases of soft-computing techniques include: management of the precision levels of problems, especially in multi-criteria decision environments; approximation of uncertainties; and provision of robust solutions. It is a concern that the use of soft-computing techniques in medical diagnosis still remain largely an academic exercise. The use of computing systems powered by soft-computing techniques in medical diagnosis is not very popular and most of the systems are mainly un-commercialized [5]. Despite this lack of adequate support for non-conventional diagnosis techniques, research on the viability of soft-computing techniques in medical diagnosis continues.

The overarching objective of this review work is to identify and clarify the extent to which research has been conducted on the use of soft-computing techniques in the diagnosis of tropical, often confusable, diseases. The remaining part of this review work is structured as follows: In Section 2, some common soft-computing techniques employed in medical diagnosis are discussed, while Section 3 presents the methodology of the review. The results and findings are presented in Section 4 and discussed in Section 5. Some conclusions are drawn in Section 6; this is followed by recommendations for future work.

SOFT COMPUTING (SC)

Soft-computing is a sub-field of artificial intelligence that utilizes approximations to find imprecise, but usable, solutions to complex problems. Soft-computing techniques are tolerant of imprecision, uncertainty, partial truth, and approximations and are characterized by their tractability, robustness, and low solution cost [6]. Some soft-computing methodologies include; Fuzzy Logic (FL), Neural Network (NN), Genetic Algorithm (GA), Particle Swarm Optimization (PSO), Support Vector Machine (SVM), K-Nearest Neighbour (K-NN), and Naïve Bayes (NB) [7,8]. These methodologies have played vital roles in medical domain research, either individually or in hybrid modes that take advantage of potential synergies between techniques. In the following sub-sections, some of the common technologies are discussed.

Neural networks

Neural Networks (NNs) are a family of statistical learning algorithms inspired by biological neurons and are used to estimate or approximate functions that can depend on a large number of inputs. They are seen as a large number of highly interconnected processing elements (nodes) that demonstrate the ability to learn and generalize from training patterns or datasets. Types of neural systems include: Multi-layer Neural Networks (MNN), Multi-Layer Perceptron (MLP), Single Layer Perceptron (SLP), and Back propagation algorithm. The Perceptron is a single layer NN whose weights and biases could be trained to produce a correct target vector when presented with a corresponding input vector, especially for simple problems.

NN has a good application potential with characteristics like fault tolerance, generalization, and the ability to learn from the environment. Some application areas include; character recognition, image processing, stock market prediction, travelling salesmen's problem, etc. A major drawback of NN is the length of time that is required to determine the connection weights needed to effectively train the neural network module. This often leads to high computational costs and constitutes a setback in real life implementation. Other problems include; the number of parameters like the optimal number of hidden nodes, the selection of relevant input variables, and the selection of optimal weights that have to be set before the training can begin. There aren't any clear rules on how to set these parameters, yet it is these parameters that determine the success of the training of neural network classifiers [9].

Genetic algorithm

Genetic Algorithm (GA) is a method used to solve both constrained and unconstrained optimization problems. It is based on the principle of natural selection, a process that drives biological evolution; and is, focused on the concept of survival of the fittest [10]. When solving optimization problems, it provides acceptable solutions within a reasonable time frame [11,12]. Genetic algorithm contains the following elements; a chromosome, a gene, a population set, fitness function, breeding, cross over, mutation, and selection. GA can be applied to areas like automotive design, engineering design, robotics, evolvable hardware, optimization telecommunication routing, etc. Some shortcomings include:

- No guarantee of finding the global maxima
- Takes time to converge
- Incomprehensible solutions

Fuzzy logic

Rather than being fixed and exact, fuzzy logic is a form of valued logic that deals with approximate reasoning. Compared to traditional binary sets (which have fixed values of TRUE or FALSE), fuzzy logic variables have truth values that range in degrees between zero and one [0, 1] [13,14]. Fuzzy logic systems are applied in many fields of study, including but not limited to communication technology, production systems, medical diagnostic systems, and data mining [3,15,16]. It is based on fuzzy sets, linguistic variables, and approximate reasoning. There are five components of the fuzzy logic system:

- The user interface: where system attributes are entered and results from the defuzzifier are put through it
- The fuzzifier: which transforms crisp attributes into fuzzy values
- The knowledge base: which is split into a database that contains static and dynamic information about patients. For instance, the fuzzy rule base finds connections between the input and output fuzzy variables
- Inference engine: which processes the decision-making output using rules contained in the rule base
- The defuzzifier: which translates the fuzzified output from the inference engine into crisp values

Fuzzy logic is applied in many areas and has various characteristics. For instance, it has a membership function and inference procedures, and also incorporates human knowledge when making decisions, such as rule base. Moreover, it is applied in systems with ambiguous elements [17,18]; it deals with approximate reasoning rather than fixed and exact reasoning, it is applied to solve problems with dynamic behaviour, it involves several variables, and it is employed in the modelling of medical diagnosis systems [19]. However, the major problem with Fuzzy Logic System (FLS) is that the determination of its fuzzy sets and fuzzy rules requires the knowledge of human experts in a particular domain [20]. The Membership Function (MF) parameters are arbitrarily chosen and hence, fixed in nature.

Particle swarm optimization

Particle Swarm Optimization (PSO) is a population-based optimization approach, where individual particles are grouped into a swarm. Each particle in the swarm represents a candidate solution to the optimization problem. The practice in PSO is that each candidate or particle “flows” through a multi-dimensional search space, adjusting its position according to its experience and the experiences of its neighbouring particles.

The performance of each particle is measured according to a predefined fitness function. The objective of PSO is to model simple behaviours and local interactions with the environment, such that each individual move towards its closest neighbour and moves back to the state that the individual has experienced to be the best for itself. The consequence of modelling this social behaviour is that the search process involves particles that stochastically return towards the previously successful regions in the search space. Applications of PSO include [21]; function approximation, clustering, optimization of mechanical structures (such as neural networks), solving system equations, game learning, power systems, design, scheduling, data mining, bioinformatics, signal processing, machine learning, and adaptation control.

The advantages of PSO include:

- It is based on intelligence
- It has no overlapping and mutation calculation
- The calculation is very simple
- PSO adopts the real number code and is decided directly by the solution.

The disadvantages of PSO are:

- The method easily suffers from partial optimism
- It cannot work out the problem of scattering and optimization [22]
- The method cannot workout the problem of non-coordinate systems

Support Vector Machine (SVM)

A Support Vector Machine (SVM) is a machine learning algorithm that analyses data for classification and regression analysis. SVM is a supervised learning method that looks at data and sorts it into one of two categories. An SVM produces a map of the sorted data with margins between the two being as far

apart as possible. SVMs are used in text categorization, image classification, handwriting recognition, and in the sciences. SVM relies on finding the hyperplane that would separate the classes in the best way possible. The term “support vectors” refers to the points closest to the hyperplane that would change its position if they were removed. The distance between the support vectors and hyperplane is referred to as the margin.

The SVM contains a series of data that classifies into two categories. SVM algorithm determines which category a new data point belongs to. This makes SVM a non-binary linear classifier. An SVM algorithm not only places objects into categories, it also ensures that the margins between them on a graph are as wide as possible. Some applications of SVM include: text and hypertext classification, image classification, recognizing hand written characters, biological sciences, and protein classification.

SVMs generally have good accuracy, especially when working with clean data sets, high dimensional data sets, or when the number of dimensions is higher than the number of samples. It can be deemed more effective with large data sets, overlapping classes, and in areas with noise. With larger datasets, the training time increases. Despite the advantages, an SVM is complex, hard to interpret, and expensive in terms of computer resource utilization (CPU and memory).

Naïve Bayes (NB)

In machine learning, the Naïve Bayes classifiers are a family of simple “probabilistic classifiers” that apply Bayes’ theorem with strong (naïve) independent assumptions between the features. This machine learning classifier relies on Bayes theorem in its classification and prediction. It can be used for both binary and multiclass problems. The Naïve Bayes method evaluates the probability of each feature independently, regardless of any correlation. The algorithm is based on the concept of class probabilities and conditional probabilities. A class probability is the probability of a class in the dataset. In other words, if a random item is selected from a data set, the probability of it belonging to a certain class can be determined. Given the class, conditional probability is the probability of the feature value.

This method is not only easy to understand, it also performs well with noisy or irrelevant data features. Its main drawback is the fact that each feature is treated independently, which in practice cannot be true [23]. The advantages of Naïve Bayes include:

- Simplicity of implementation and speed
- If the NB conditional independence assumption holds, it converges quicker than discriminative models like logistic regression
- Even if the NB assumption doesn’t hold, it works great in practice
- It needs less training data
- It is highly scalable; it scales linearly with the number of predictors and data points
- It can be used for both binary and multi-class classification problems and can make probabilistic predictions
- It handles continuous and discrete data and

- It is not sensitive to irrelevant features

K-Nearest Neighbor (KNN)

The k-Nearest Neighbor (K-NN) is a non-parametric and lazy classification model that predicts a class based on the features of known observations that are close to it. A k-nearest-neighbor algorithm is an approach to data classification that estimates how likely a data point is to a member of one group or the other, depending on its nearest data points. That is, it attempts to determine what group a data point is in by looking at the data points around it.

The k-nearest-neighbor is an example of a "lazy learner" algorithm because it does not generate a dataset model beforehand. The only calculation it involves in is when it is asked to poll the data points of neighbors. This makes the K-NN very easy to implement for data mining. The term non-parametric refers to the fact that the K-NN does not make assumptions about the distribution of the data it is modelling. For example, with linear regression, the model assumes the data is normally distributed. If the data follows a much different distribution, like a bimodal, a linear regression no longer works. Here, not having to worry about distribution is a big advantage because it means that the K-NN can be applied to a lot of data sets. A K-NN is said to be lazy because it takes little to no training time. Normally, when building a model, data is separated into a 'training set' and a 'testing set.' The training set is a group of data observations used to develop the model. The testing set uses the rest of the data to evaluate how well that model works with data it has never encountered. The training phase refers to the time spent using the training set to create the model. The K-NN has no training phase because it essentially uses all of the training data whenever it tests. The last part of this definition is based on the features of known observations that are close to it.

Essentially, the K-NN predicts the classification of some data points based on known classifications of observations that are most likely that data point. There are different instant measurement methods used to find the closest neighbours. These methods include: the hamming distance, Manhattan distance, and Murkowski distance. The drawback of the K-NN is its poor performance when using unevenly distributed datasets. K-NN is a great classification algorithm because it involves looking at observations with similar features to classify other observations and is hence, easy to explain.

METHODOLOGY

This systematic literature review is based on the procedures and guidelines proposed by Kitchenham B, which are used in collaboration with ideas that constitute an effective literature review by Webster and Watson [24, 25]. For this review, the following subsections were considered: protocol development, research questions/search strings, search process (using digital libraries/search engines and grey libraries), search for relevant primary literature, identification of the inclusion/exclusion criteria, audit review, data extraction from the accepted articles, and data analysis.

Protocol development

The systematic literature review protocol was not only developed, it was also registered with PROSPERO (with registration number CRD42017069633). The purpose of registering with PROSPERO was to ascertain that the research topic was not already registered or being worked on by any person or group. Figure 1 illustrates the taxonomy of the review protocol.

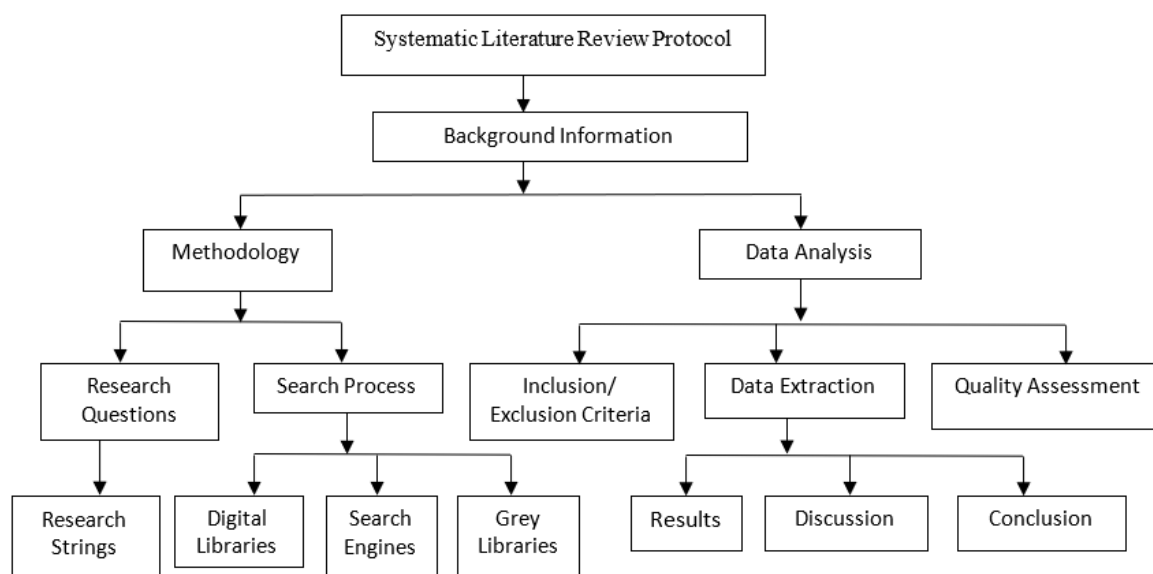


Figure 1: Taxonomy of systematic literature review protocol.

The review protocol specifies the elements that are used in a specific systematic literature review. Figure 1 includes the following elements and additional planning information [26]:

- The background, which addresses the rationale for the systematic literature review
- The research questions that the review intends to answer
- The strategy that will be used to search for primary papers. This includes search strings and resources to be searched: digital libraries, manual search, conference papers, and grey papers
- The study selection criteria. This is used to determine whether the papers will be accepted (included) in the review or rejected (excluded) from the review
- The study selection procedure. This includes the number of assessors that will review the primary papers as well as the method to be used to arrive at a compromise during disagreement(s)
- The quality assessment. This includes a check-list of procedures to ensure that the individual papers meet the necessary conditions; and that the researchers have fulfilled such conditions
- The data collection strategy. This describes how the gathered data can be categorized; for instance: the authors and their affiliations, the country of residence, the journals published in with dates
- The data collected is then discussed, analysed, and conclusions are drawn

Research questions

For many countries in the developing world, tropical diseases represent a significant health burden. Neglected Tropical Diseases (NTDs) alone affect an estimated 1.4 billion people. NTDs therefore, constitute a significant burden on the already strained healthcare systems and economies of many developing countries [27]. To complicate matters, there is not much interest in developing new diagnostic tools for the diseases. Even while areas like Europe, Australia, and Florida have reduced the disease burden, there are still considerable diagnostic challenges. Given the fact that some of these diseases produce confusable symptoms, it is evident that better techniques and strategies for diagnosis need to be developed [28].

While laboratory techniques have not been much help, soft computing techniques are now being explored to solve these problems. It is in light of the above that the following research questions have been formulated. The answers to these will provide meaningful information for decision making [29]:

RQ1: How much research has been conducted in diagnosing tropical diseases using soft computing tools from 2008 to 2017, inclusive?

RQ2: What tropical diseases are diagnosed using soft computing techniques?

RQ3: Which institutions and countries most actively use soft computing techniques for the diagnosis of tropical diseases?

RQ4: Which soft computing tools are frequently used in the diagnosis of tropical diseases?

RQ5: What are the limitations of current research on tropical diseases diagnosis?

Search process

In this review, we used PubMed, Google-Scholar, Science Direct, and Directory of Open Access to gather literature. All the searches were based on a title, keywords, and abstract; and all of the sourced articles were downloaded and printed for review. A total of one hundred and forty-seven (147) articles were sourced, out of which, fifty-eight (58) primary articles were accepted, and eighty-nine (89) articles were rejected. Boolean operators were used to bind the search strings. For instance, AND was used between strings, while OR was used within strings. Some of the search strings that were used are as follows:

- ('Soft computing techniques') AND ('Diagnosis of Tropical Diseases')
- ('Soft computing' OR 'Machine learning' OR 'Computational Intelligence') AND ('Diagnosis of Tropical Diseases')
- ('Machine learning classifiers' OR 'Artificial Neural Network' OR 'Support Vector Machine' OR 'Fuzzy logic' OR 'Naïve Bayes' OR 'K-Nearest Neighbour' OR 'Decision Trees' OR 'Genetic Algorithm' OR 'Particle Swarm Optimization') AND ('Diagnosis of Tropical Diseases').

Inclusion and exclusion criteria

All the researchers went through the titles of the primary journals in order to determine their relevance to the systematic literature review. We discovered that some journals were not relevant; others discussed tropical diseases; while some only discussed soft computing techniques. These were all rejected.

The journals accepted for inclusion were the ones that discussed the diagnosis of tropical diseases using soft computing techniques. Some articles discussed the diagnosis of individual tropical diseases using single soft computing classifiers, while others discussed multiple tropical diseases using multiple classifiers. All of these variations were accepted if the disease in question was a tropical disease and the technique used for diagnosis was a soft computing technique. However, it was observed that abstracts, titles, and keywords were not always directly related to the topic or what the article was about. In these situations, we read the introduction, methodology, and conclusion of such articles to determine their relevance.

Quality assessment

Every aspect of the protocol or review was evaluated using the York University Centre for Reviews and Dissemination (CRD) and the Database of Abstracts, Reviews, and Effects (DARE) Criteria [30]. Although the DARE criteria suggested the use of linguistic variables such as YES, PARTIAL and NO, which are then translated to an interval scale of [0,1] [31], the researchers decided to use a simple evaluation scale of 1-10 to evaluate the literature review against the DARE criteria (questions). Using a

Delphi technique, the researchers arrived at the average scores and reasons for the ratings (Table 1).

Table 1: Research quality evaluation.

Question	Average score (/10)	Reason
Is the review's inclusion and exclusion criteria described and appropriate?	7.2	Members of the team were mainly computer scientists, some of whom had good experience in medical information systems research.
Does the literature search likely cover all relevant studies?	4	The researchers made extensive use of search engines like Google scholar and open access online journals for the primary papers used in this research. Access to a number of digital libraries was limited due to subscription challenges.
Did the reviewers assess the quality/validity of the included studies?	7.7	The data to be extracted from the primary papers were agreed upon by the team in a systematic manner.
Was the basic dataset adequately described?	8	Information presented within the data was illustrated using tables, charts, and line diagrams
Is there a clear statement about the aims of the research?	8.7	The aim of the study was properly established and discussed over several team meetings. The study emanated from an identified literature gap on the use of soft-computing techniques for diagnosis of tropical diseases.

As evident from the highest score, researchers seemed to be very confident in the aims of this research. This confidence was also demonstrated by the average score of 8 for data description. This data was shown through tables, charts, and line diagrams. The digital search areas used were scored at a low 4 because it was difficult to obtain and learn from the primary papers within these digital libraries; this may have resulted from the search strings used.

RESULTS OF FINDINGS

We identified one hundred and forty-seven (147) primary sources, out of which, eighty-nine (89) did not meet the inclusion criteria and hence, were rejected. Fifty-eight (58) articles were "hit" and all of them carried out diagnostic analysis of tropical diseases using soft computing techniques. The revolution of large data in Electronic Health Records (EHR) and the glossing over of symptoms, which has made them confusable in many diseases, has led researchers to adopt classifier hybridization when diagnosing tropical diseases. For instance, to diagnose hepatitis, which has confusable symptoms with malaria and typhoid, Chen et al. proposed a hybrid method based on local fisher discriminant analysis and Support Vector Machine (SVM) [32]. Faisal et al. developed a non-invasive intelligent technique for predicting the risk of dengue fever and its variants for patients [33].

As evidenced from the sixteen hybridized research papers reviewed in this work (Table 2) and given the drug resistance of current disease variants, hybridization appears to be taking centre stage.

Research questions

RQ1: How much research has been conducted in diagnosing

tropical diseases using soft computing tools from 2008 to 2017, inclusive?: There were fifty-eight (58) articles that met the inclusion criteria. This number represents 39% of the total articles sourced (one hundred and forty-seven). There were eighty-nine (89) articles rejected, which accounts for 61% of the total articles sourced. Table 2 shows the distribution of articles that met the inclusion criteria, per year.

Table 2: Number of publications per year, and their percentages.

Year	No. of publications	Percentage
2008	1	1.72
2009	6	10.35
2010	2	3.44
2011	6	10.35
2012	5	8.62
2013	6	10.35
2014	8	13.8
2015	9	15.51
2016	10	17.24
2017	5	8.62
Total	58	100

Figure 2 is a graphical representation of academic awareness and research efforts made by third world countries that are most affected by these tropical diseases. As shown, the number of publications gradually increased through the years under analysis, except for 2010, 2012 and 2017 that seemed to be outliers. In fact, 2010 had two (2) publications, while the years 2012 and 2017 each had five (5) publications. This number then climaxed to ten (10) in 2016.

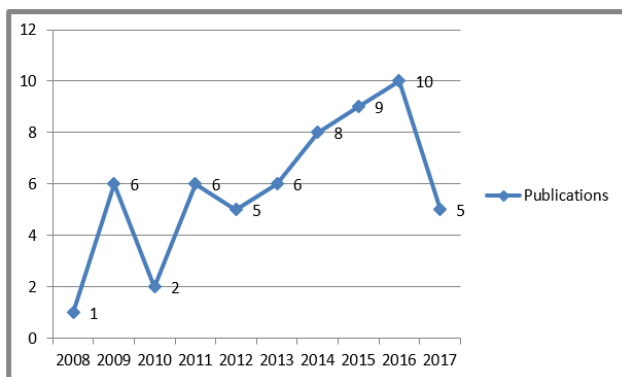


Figure 2: Publications trend.

It is also observed that the research community was spread around the world with Africa having twenty-eight (28) publications on the diagnosis of tropical diseases from 2008 to 2017; Asia-Pacific had twenty-two (22) publications spread across seven (7) countries: India (11), Malaysia (3), Singapore (3), China (2), while Taiwan, Pakistan and Thailand had one each (1). In Africa, Nigeria had twenty (20) publications, others had one publication each. The Middle East had one (1) publication from Turkey. The Americas had seven (7) publications from four (4) countries: Canada had three (3) publications, Brazil had two (2), while USA and Mexico had one (1) publication each. Europe had eleven (11) publications from seven countries: Indonesia had five (5) publications while the others had one (1) publication each.

RQ2: What tropical diseases are diagnosed using soft computing techniques?: Although the World Health Organization (WHO) identified seventeen (17) neglected tropical diseases for control and eradication, researchers focused on the diagnosis of twelve (12) different tropical diseases using soft computing techniques. Malaria was the most researched disease, with seventeen (17) publications; dengue fever had fourteen (14) publications referencing it, while skin diseases and hepatitis had nine (9) publications referencing each of them. In contrast, Tuberculosis had four (4) publications referencing it. Soil transmitted parasites, epilepsy, and chikungunya had two (2) publications referencing each of them; while the Ebola virus, chagas parasite, and peripheral neuropathy were mentioned in one article each. Figure 3 illustrates the distribution of diseases considered in the publications reviewed.

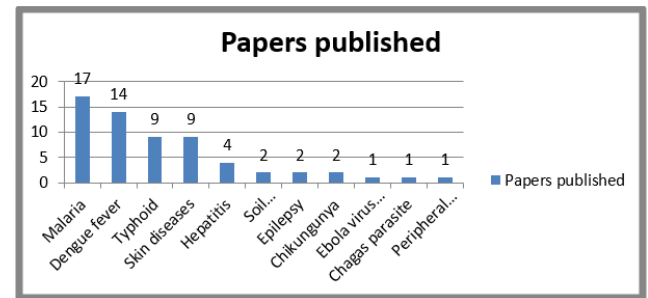


Figure 3: Distribution of diagnosed diseases.

The number of publications per disease seems to be dependent on each disease's extent of devastation and the corresponding effort to contain it. As illustrated in Figure 3, the graph shows the number of papers published per disease. The researchers have also proposed systems for the diagnosis and treatment of many diseases. Some of the systems include; Mobile Based Fuzzy Expert System (MFES) for the diagnosis of malaria [34], and Ebola fuzzy informatics system for the diagnosis and therapy of the Ebola Virus Disease (EVD) using Fuzzy-Neural System (FNS) [35]. Unfortunately, it appears that the soft computing techniques and systems used in the diagnoses of these diseases are still waiting to be commercialized from academic archives.

RQ3: Which institutions and countries most actively use soft computing techniques for the diagnosis of diseases?: Most research efforts appear to be carried out by universities. In fact, there were forty-four (44) universities that participated in this research; seven (7) specialist hospitals, five (5) polytechnics, one (1) research institute, one (1) college of education, and one (1) school of nursing. As illustrated in Table 3, there were seven (7) collaborations between the universities and specialist hospitals; five (5) collaborations between the universities and polytechnics; while there was only one (1) collaboration each, between the universities and colleges of education, research institutes, and schools of nursing.

Table 3: Collaborations between the universities and other institutions.

S/N	Institutions	Collaborations
1	Total number of universities that contributed to the research process	44
2	University and specialist Hospitals	7
3	University and polytechnics	5
4	Universities and colleges of education	1
5	Universities and research institutions and hospitals	1
6	University and research centers	1
7	Universities and school of nursing	1

Unfortunately, as shown in Table 3, since there are very few institutions of learning and research centers recorded, there appears to be a “research divide” between universities and other institutions of learning. Just as diagnostic techniques are hybridized to yield better results, positive collaboration between these institutions could produce diverse ideas and hence, better research results. Moreover, these collaborations could persuade respective bodies (governments, the WHO, and other stakeholders) to view these research activities as economic, social, and commercial ventures worthy of commercialization.

Twenty-eight (28) countries were represented in this research. Nigeria had the highest literary contribution, with twenty (20) research papers; India had eleven (11) papers, Indonesia had five (5) papers; Singapore, Malaysia, and Canada had three (3) papers each; while Brazil and China published two (2) papers each, and other countries published one (1) paper each. Figure 4 illustrates publications by country of residence.

From a regional perspective, nine (9) African countries contributed twenty-eight (28) publications; Asia-Pacific had seven (7) countries that published twenty-two (22) papers.

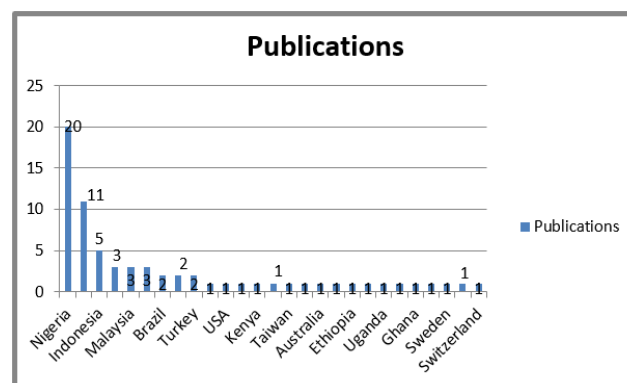


Figure 4: Publications by country of residence.

Europe had seven (7) countries that participated in the research with eleven (11) publications. Four (4) countries from the Americas contributed seven (7) publications to the fight against tropical diseases. The Middle East had one (1) publication from a country. Figure 5 illustrates the regions and their contributions to the fight against tropical diseases.

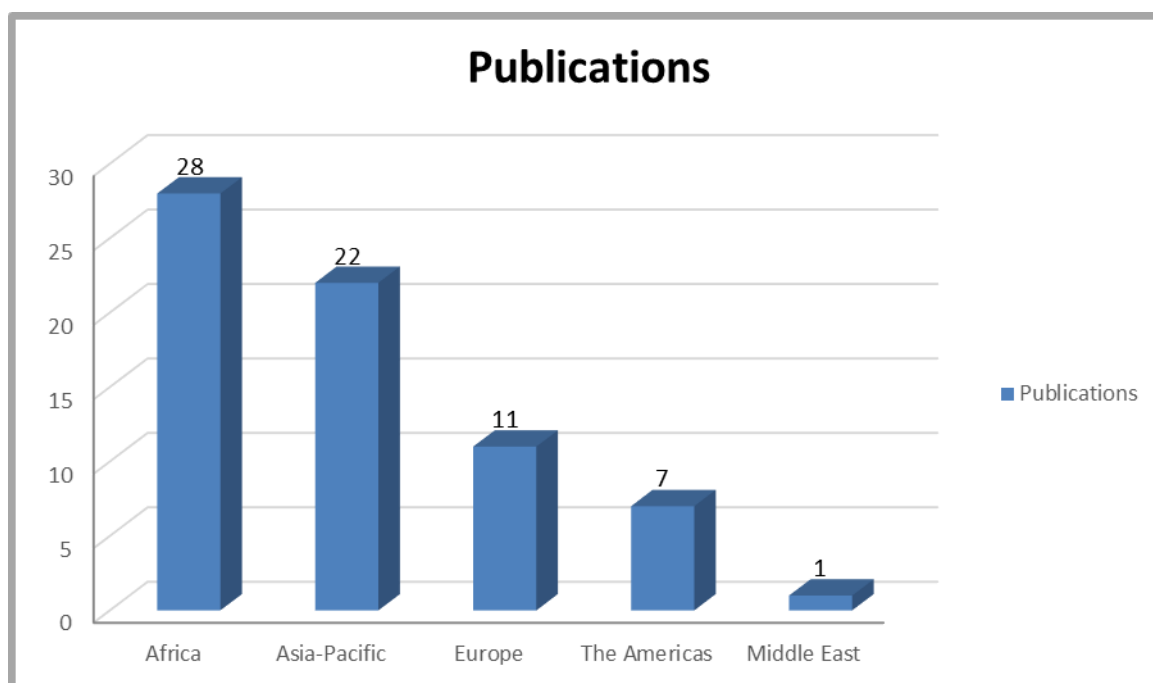


Figure 5: Regional contribution to the fight against tropical diseases.

Table 4 shows the countries within these regions and their contributions to the fight against tropical diseases using soft computing techniques.

Table 4: Contributions by countries within the regions.

S/N	Region	Publications
Africa		28
1	Nigeria	20

2	Egypt	1
3	Kenya	1
4	Botswana	1
5	Ghana	1
6	Uganda	1
7	Sudan	1

8	Tanzania	1
9	Ethiopia	1
	Asia-Pacific	22
10	India	11
11	Malaysia	3
12	Singapore	3
13	China	2
14	Taiwan	1
15	Thailand	1
16	Pakistan	1
	Middle East	1
17	Turkey	1
	The Americas	7
18	Canada	3
19	Brazil	2
20	USA	1
21	Mexico	1
	Europe	11
22	Indonesia	5
23	Ireland	1
24	United Kingdom (UK)	1
25	Australia	1
26	Finland	1
27	Sweden	1
28	Switzerland	1
	Total	69

Table 5: Hybridized Classifiers used to diagnose tropical diseases.

S/N	Hybridized classifiers	Code	Research papers
1	Self-organizing map and multilayer feed forward neural network	SOM-MFNN	Faisal et al. [33]
2	Local fisher discriminant analysis and support vector machine	LFDA-SVM	Chen et al. [32]
3	Adaptive neuro-fuzzy inference system	ANFIS	Faisal et al.; Panford et al. [33,40]

There were twenty-eight countries that participated in this exercise. It is worthy to note that in some papers, two or more authors came from different countries to the one under analysis. The trend shows that tropical countries are no longer waiting for developed countries to fight the diseases for them, nor are developed countries undermining these diseases' global effects. Countries in Africa, Asia-Pacific, and the Middle East seem to be taking the lead in finding such solutions.

Developed countries seem to be apathetic to this problem because it seems to be an issue mainly affecting developing countries. However, tropical diseases have now become a global issue. For instance, the Ebola Virus Disease (EVD) was first reported in 1976 but the largest outbreak started from West Africa in 2014 and was transmitted to Europe and America relatively quick [36-39]. This incident helped to establish a more global understanding of the disease. Therefore, health workers, especially those in developed countries, cannot afford to neglect the fact that diseases can simultaneously impact more than one region. Hence, to ensure the health of everyone, disease prevention and control should be addressed through a global perspective [36]. In fact, tropical diseases such as polio, dengue fever, Ebola, etc are resurfacing from one geographical region to another. Such transition is accomplished through the following means: transportation (import and export), immigration, business ventures, and tourism. Also, countries with internal conflicts, such as Syria, Liberia, Cameroon, and those experiencing famine, such as Ethiopia, usually have citizens fleeing to other countries for a better quality of life.

In addition to tropical diseases, most developing countries are also facing wars and famine. Such countries include Libya, Syria, Liberia, Cameroon; and Ethiopia suffering from draught and famine. Consequently, western countries have to accommodate refugees trying to seek asylum from these problems in their countries. If developed countries assisted developing countries in solving these problems, such pressure for both parties would have been greatly minimized. Hence, global collaborative efforts provide the most effective and beneficial solution to these issues.

RQ4: Which soft computing tools are frequently used in the diagnosis of tropical diseases?: Thirty-two (32) classifiers were uncovered in this review. Sixteen (16) of them were hybridized while the remaining sixteen (16) were single unique classifiers. Table 5 lists the hybridized soft computing techniques used in the diagnosis of tropical diseases, along with their codes and corresponding research papers.

4	Fuzzy cognitive map	FCM	Djam and Wajiga [41]
5	Genetic algorithm and vertex chain code	GA-VCC	Mansour [42]
6	Support vector machine and simulated annealing	SVM-SA	Sartakhti et al. [43]
7	Support vector machine and bayesian network	SVM-BNet	Rusdah [44]
8	Neuro-case base reasoning	N-CBR	Obot and Uzoka [45]
9	Particle swarm optimization and C4.5 decision tree	PSO-DT	Chen et al. [46]
10	Scatter search and support vector machine	SS-SVM	Afif et al. [47]
11	Naïve bayes and J48 decision tree	NB-J48	Shaukat et al. [48]
12	Fuzzy min-max neural network- classification and regression tree-random forest	FMM-CART-RF	Seera and Lim [49]
13	Genetic adaptive neuro-fuzzy inference system	GANFIS	Asogbon et al. [50]
14	Fuzzy-neural system	FNS	Abiyev and Abizade [51]
15	Decision tree, rough set, naïve bayes	DT-RS-NB	Singh et al. [52]
16	Artificial neural network and support vector machine and linear discriminant analysis	ANN-SVM-LDA	Alanis-Reyes et al. [53]

Table 6 lists the other sixteen (16) single classifiers mentioned earlier, along with the corresponding research paper(s).

Table 6: Unique single classifiers used in the diagnosis of tropical diseases.

S/N	Name of classifier	Code	Research paper
1	Fuzzy logic	FL	15
2	Artificial neural network	ANN	5
3	Support vector machine	SVM	4
4	Decision tree	C4.5/J48	4
5	Extreme learning machine	ELM	3
6	Adaboost	A	3
7	Naïve bayes	NB	3
8	Bayesian net	BNet	3
9	Rough set theory	RST	1
10	Random forest	RF	1
11	Case-based-reasoning	CBR	1
12	Sequential minimal optimization	SMO	1

13	Relevance vector machine	RVM	1
14	Deep learning	DL	1
15	Image processing	IP	1
16	Analytic hierarchy process	AHP	1

As confirmed by the “no free lunch” theory [54], any single classifier is not universal to all cases. Therefore, hybridization has become a method of choice. This is the result of the polymorphic and metamorphic nature of tropical diseases with similar and confusable symptoms, their resistance to drugs, and the weaknesses of most classifiers in one form or another. Essentially, combining classifiers will improve their ability to diagnose diseases.

When reviewing the fifty-eight (58) primary articles, it was observed that some papers researched with the same classifier or the same combination of classifiers. For instance, fifteen (15) primary papers researched the use of fuzzy logic in the diagnosis of tropical diseases. More so, two (2) papers used Adaptive Neuro-Fuzzy Inference System (ANFIS) in their models. Another two (2) primary papers referenced or used ABM1-NB-BNet-SMO-RF. Similarly, the C4.5 decision tree and ANN were used in the diagnosis of tropical diseases by three (3) primary papers, while CBR appeared in two (2) primary papers. The rest of the papers each used one (1) classifier.

RQ5: What are the limitations of current research on tropical disease diagnosis?: The fundamental principle of soft computing involves bearing inconsistencies, fuzziness, and partial truth. It also ensures an estimate to attain tractability, strength, and low

solution cost. Over the years, soft computing techniques have used symptoms to identify diseases. Symptoms may incorporate clinical parameters like blood pressure, blood glucose, and scanning reports, or it could involve linguistic expressions like nausea, weakness, etc. These attributes can be taken as symptoms and input into the training classifier for analysis. Despite the huge successes recorded on disease diagnosis using soft computing, several short comings still exist:

- Vague symptomatology still exists in the present diagnosis system. Some diseases have not been characterized
- The system is lacking statistical tools in the processing of training data. This makes the symptoms too difficult to be understood by non-expert physicians
- Patients' symptoms vary across regions due to a variety of factors including diet. As such, the system lacks the ability to advise the right prescription for confusable diseases in a generalized framework
- Given the confusable symptoms, a unique classifier may not be able to effectively diagnose a disease. For instance, yellow fever diagnosis is not adaptive and soft computing techniques have not been successful at classifying or clustering it
- There is still a problem with the response time in these techniques, especially the neural network technique

The reasons above emphasize why hybridization can no longer be ignored. This is indicated within this research work, where up to sixteen (16) hybridized techniques were used in the diagnosis of tropical diseases. This addresses the key problem within the classifiers' comparative advantages.

The interplay between imprecision and vague features often influences physicians to make decisions based on their experiences rather than relying on a scientific model [55]. The ability of a model to reveal its reasoning in dealing with input data and justify its prediction would give the physician confidence in using the result. This is lacking in constituent classifiers [56,57]. Without understanding how the model achieved the result, physicians will not be confident in the model [56]. An example of the single classifier's weakness in this direction is given in the use of neural network; all neural networks are efficient in solving various problems, but they lack the ability to explain their answers and present gathered knowledge in a comprehensive way [57]. In fact, SVM and ANN are viewed as black boxes with interest only in their inputs and outputs [57].

To win the physician's confidence in intelligent support systems, hybridization of classifiers is employed. Seera and Lim demonstrated this by combining Fuzzy Min-Max Neural Network (FMM-NN), Classification and Regression Tree (CART) and Random Forest (RF) [58]. Similarly, Yedjour et al. combined traditional NNs with Genetic Algorithm (GA) to extract simple, intelligible, and useful rules from trained NN [57]. The combination of these classifiers made use of their advantages and avoided their shortcomings. For instance, while CART has the advantage of rule extraction, it is less flexible in incremental learning from sample data. On the other hand, FMM-NN is good in incremental learning but lacks the capacity to explain its predictions (black box); RF has the benefit of

combining (ensemble) CART to produce the best tree with high prediction accuracy [58].

Researchers observed that soft computing techniques have been developed to support many decision-making tasks. For instance, classifiers are used for the diagnosis, prognosis, and screening of diseases. Several neuro-fuzzy models have been used as classifiers for malaria, dengue fever, etc. because they are capable of learning from the patients' records (data samples) and can also generalize beyond the training samples [59].

DISCUSSION

The awareness about the threat of tropical disease has steadily grown from the years under review. Current medical diagnosis is based on human abilities, uncertainty factors, ambiguous symptoms, high accuracy requirements, and bulk medical records. Therefore, using soft computing techniques would allow for more effective diagnosis because they can deal with imprecision, uncertainty, and partial truth [60]. In fact, diagnosis through computer aided design is affordable and can efficiently serve many patients. This is especially useful in developing countries that often lack the necessary medical facilities.

With proper training, testing and application, soft computing techniques display tacit knowledge of experts in respective fields (domains) and can reduce the spread of tropical diseases. They could aid physicians (especially inexperienced ones) to make appropriate decisions regarding patients' confusable symptoms and determine corresponding therapy. Soft computing approaches have been used for many diseases and various aspects of research and development; yet, the data produced by such approaches have not been widely accessible to others or integrated into current health routines [61].

Unfortunately, there isn't much governmental support for this type of diagnostic approach. More so, there is a lack of economic incentive or market to incentivize these products. There is also a relatively nonchalant global and national response to tropical diseases. Instead of being neglected, these diseases should be prioritized within the health agendas of every affected country [62]. Though soft computing techniques have proved to be successful, the WHO and allies don't seem to acknowledge them. Instead, the WHO is concerned with funding other laboratory diagnostic techniques [63]. It is possible that the WHO may be unaware of soft computing diagnostic processes, which is why she does not show an interest in soft computing techniques. The academic communities, already constrained by research funds, are yet to sell their proposals, models and frameworks to the public, be it by way of advertisements or collaborations. Research results are mostly published in academic journals as part of job and promotion requirements. There has been no adequate effort made by governments or the WHO to advance these systems and frameworks for commercial purposes. Therefore, researchers in soft computing processes should take advantage of social media to not only spread public awareness of soft computing processes, but to also produce a market for them. In 2010, the WHO, Bill and Melinda Gates Foundation, and thirteen (13) leading pharmaceutical companies resolved to sustain and expand programs that ensure

the necessary drug supply and interventions for the eradication of guinea worm disease, Lymphatic Filariasis (LF), leprosy, Human African Trypanosomiasis (HAT) etc., by 2020 [64]. Unfortunately, there was no mention or plan for soft computing techniques; which could be resting in academic archives.

The impact of soft computing in medical diagnosis cannot be overemphasised. The first attempt at creating decision support system for medical diagnosis began with the application of statistical methods in the 1950s. However, by early 1970s, it was discovered that this technique could not handle complex clinical problems. This encouraged the search for solutions using soft computing techniques. Pattern recognition methods became the focus of soft computing in medical diagnosis until 1974 when rule-based approach for analysis and treatment (therapy) of infectious diseases was used [65]. Rule-based programs used the “IF-THEN” rules, which became difficult to use with complex diagnostic problems.

Given the shortcomings of rule-based techniques, emphasis was shifted to representation of unstructured, imprecise and dynamic knowledge, which characterized the sources of information fed into the soft computing tools [65]. Following the development of disease variants and their resistance to drugs, researchers resorted to hybridizing soft computing techniques to fight the menace of tropical diseases. This increased the effectiveness and efficiency of the tools in the management of imprecision [66]. The use of Fuzzy Logic (FL), Neural Network (NN) and Support Vector Machine (SVM) in medical diagnosis became common place; Fuzzy Cognitive Map (FCM), Genetic Adaptive Neuro-Fuzzy Inference System (GANFIS) among others, are hybridized tools used in the diagnosis of tropical diseases. Given such a long history, one wonders why the WHO and her allies are giving soft computing diagnostic techniques little or no attention at all.

Soft computing techniques not only tackle current disease symptoms, they also address new and obfuscated variants that have become resistant to drugs [67]. Moreover, these techniques reduce mortality rates by storing treated patients’ records in distributed databases for dynamic Intrusion Detection Systems (IDS), Intrusion Prevention Systems (IPS), and Real-time Intrusion Response Systems (IRS) [68]. Hybridizing IDS and IRS that use anomaly (dynamic) methods promotes the effectiveness and efficiency of intrusion detection services [68]. Huge datasets are being used as of recent, and the hybridization of soft computing and data mining would go a long way to differentiate the confusable symptoms generated from these huge datasets. In fact, data mining is a powerful tool used to locate unknown patterns and useful information from huge datasets [69] this technique extracts hidden predictive information from a huge dataset [70].

Over the years, the fundamental principles of soft computing have involved inaccuracies, fuzziness, and low solution cost. Several techniques have been proposed to apply the use of soft computing to medical related fields [71]. These techniques use symptoms to identify diseases. Such symptoms could involve clinical parameters like blood pressure, blood glucose, sputum, and scanning reports. The use of clinical parameters by laboratory diagnostic techniques like smear microscopy,

radiography, molecular methods, etc., have not yielded satisfactory results because of the time it takes to complete test analysis and their inability to identify confusable symptoms. These shortcomings encourage researchers to turn their attention to using soft computing techniques in tropical disease diagnosis. Laboratory and statistical techniques have not been able to deal with complex clinical problems, which paves the way for the exploration of soft computing techniques in medical diagnosis. Soft computing utilizes unstructured, imprecise, and dynamic knowledge, which characterizes the sources of information available to decision support systems.

Even within the funded drug industry, the complete discovery process can take many years. Within these trial years, negative results are often not published. In order to promote patenting and trade secrets, these negative results are shrouded in secrecy [72]. This negative result is also beneficial to molecular modelling and computational chemistry, but are hidden in research centres because of market rivalry. These approaches, to attain a competitive advantage over other companies in the trade, slow down the drug production process. As such, it is a waste of effort to duplicate research laboratories that are working towards drug production for tropical diseases [61]. Instead, such negative data needs to be collated and curated for effective mining efforts.

Tropical diseases constitute a large fraction of the world’s disease burden; yet, they receive only a small fraction of global research and development (R and D) spending [73]. This occurs because private and public purchasers in the developing world are not rich enough to pay for treatment and bulk drugs. The cost and uncertainty of innovation is increased by Intellectual Property (IP) issues like patent, cost, complexity, and breadth. The for-profit pharmaceuticals might acknowledge the above statement, while also pointing out that the cause lies in a lack of incentives to enable changes in this area. Another disincentive for wider data sharing is the difficulty in securing competitive research funds, which are tied to strong preliminary data acquisition. Research bodies fear that sharing these preliminary data could provide information for competing laboratories. This lack of sharing of clinically relevant knowledge and scientific progress results in the duplication of models, laboratories, and research efforts. This leads to high cost and a delay in the progress of tropical disease research and development.

“Open sourcing” could provide an opportunity for improvement. Open sourcing involves sharing data, experiences, and resources to increase collaboration, transparency, and cumulative public knowledge [74]. Although there are legal issues with respect to data sharing, there are ways to circumvent them. For instance, such data could be released via an online database where users can enter into a “click-wrap agreement” that governs the use of data. This could also be done through an open content licensing model, which allows researchers to access and reuse licensed data [75]. These strategies not only help to minimize duplication efforts, they also create a database of knowledge from which future innovation can grow. Luckily, there are new collaboration models evolving in “open science”. For example, in the malaria drug discovery field, there is an “open source” drug discovery campaign [74]. In the field of soft

computing, obfuscating disease variants and their resistance to drugs is being analysed through a combination of classifiers (hybrid) that take advantage of their comparative strengths. An example of this is the use of a genetic algorithm to enhance the adaptive neuro-fuzzy inference system's (GANFIS) performance when diagnosing typhoid fever [76-78].

CONCLUSION

Soft computing techniques deal with approximate models that provide solutions to real-life problems that are often complex, and characterized by uncertainty, imprecision, and approximations. These techniques are becoming a major focus of research and have been successfully utilized in various fields such as medicine, commerce and industry. In this paper, we conducted a systematic literature review on the use of soft computing techniques in the diagnosis of tropical diseases. This research covered a ten-year span, from 2008 to 2017. Our study focused on filtering the following information: year of publication, authors, institutions of affiliation, country of residence, and the number of articles published per year. In addition, we recorded the diseases diagnosed and the corresponding soft computing techniques used in the diagnosis of the diseases.

This study displayed an encouraging trend. The research community had a steady increase in publication on various topics using varied classifiers, including hybridized types. Hybridization, indeed, is gaining grounds in the analysis, partly because of tropical disease resistance to drugs and their exhibition of confusable symptoms. It also pointed to a growing awareness on the variability of soft computing techniques in the diagnosis of tropical diseases. Nigeria, a third world country, took the lead in the diagnosis of tropical diseases using soft computing techniques. Overall, there was a balanced show of interest in the fight against tropical diseases. Globally, the five zones of Africa, the Middle East, Asia-Pacific, Europe, and the Americas contributed to the research using soft computing techniques. The diseases researched on were twelve (12) compared to the WHO's target of seventeen (17) neglected tropical diseases. That was a mark of awareness on the part of researchers, especially those from the most hit regions like Africa, with malaria still dominating in the research process. Fuzzy logic classifier was mostly used either in combination with other techniques or as a single classifier in the diagnosis of tropical diseases. Since no single classifier performs excellently in all cases, as confirmed by the no-free-lunch theory, there has been a transition from mono-system to hybridized system.

The spread of tropical diseases no longer segregates the rich and poor nations. In fact, the Ebola "busts" in 2014 took the world by storm and provided an eye opener for developed countries. Tackling health related issues through global cooperation is a panacea to the well-being of all. More so, promoting world peace will assist in the containment of these diseases, especially those transmitted by people fleeing their war-torn countries. Here, industrialized countries should not fuel the political instability of such countries like Southern Sudan, Nigeria, Libya, Iraq, Syria, etc. While selling arms and ammunitions can be a

financial advantage, such support can endanger the lives of many people.

Our study also found that most research breakthroughs in the use of soft-computing technologies are within the confines of academia, with most of the models not implemented. Researchers need to partner with health institutions to test and ultimately deploy these systems. Although soft computing techniques are very effective, it is important to note that they also have the following limitations: vague symptomatology, lack of statistical tools to process the training dataset, and regional diets that produce different symptoms for similar diseases. These confusable symptoms make it difficult for less experienced physicians to effectively and efficiently diagnose these diseases. Above all, soft computing techniques lack the ability to explain their predictions. This forces physician to make decisions based on their experiences, rather than the suggestions made by computer assisted systems.

This study is significant because it represents a major effort in articulating existing literature on the application of soft-computing technologies in the diagnosis of tropical diseases. Out of the one hundred and forty-five (145) publications that covered the diagnosis of tropical diseases using soft computing techniques, not a single systematic literature review was found. It is possible that a systematic literature review on this topic had not been conducted before. This observation was confirmed by PROSPERO, an international prospective register of systematic reviews. The group sent a protocol to PROSPERO to confirm that the mentioned title had not been registered with the database and to determine if any person or group of persons was currently working on it.

The results of this study have implications for research and for medical practices. First, there is a growing awareness of the danger of tropical diseases to the global community. For medical practice, computer assisted systems have greatly aided nascent medical practitioners in overcoming the confusable symptoms of related diseases and reduced morbidity and mortality rates. It is also observed that the number of diseases researched on is progressively reaching the WHO's target. Money, time, and resources need to be pooled together for the diagnosis of tropical diseases using soft computing techniques, by governments, the WHO and other stakeholders. Since traditional diagnostic techniques known to the WHO could not curb the menace of tropical diseases, it is high-time soft computing techniques-which are cheaper, varied, and can handle fuzzy and confusable problems - should be given the publicity and commercialization they deserve.

There is significant "research divide" between the Universities and other research centres and institutions of higher learning; this is sending the wrong signals to governments, the WHO, and other stakeholders. Research results are buried in the archives of Universities with little or no publicity to the larger community. Hence, these research works are seen as mere University academic exercises, rather than research efforts worth commercializing. The research divide is obvious as shown in Table 3: only seven (7) collaborations were recorded during the research period between Universities and specialist hospitals; five (5) collaborations with Polytechnics; and one (1) each with

other research centres and institutions of higher learning compared to forty-four (44) universities. One of the reasons for this “research divide” is the political and academic superiority enjoyed by university graduates in Nigeria; for instance: the disparity between university graduates and polytechnic graduates on employment positions and salary, without minding the productivity of polytechnic graduates in the field. Similarly, medical doctors see themselves as very superior to other paramedic graduates in the health sector. Therefore, collaboration between such institutions with universities is looked down upon by university lecturers and/or researchers in Nigeria. This divide is extended to homes and parents, who ignore the six-three-three-four (6-3-3-4) education systems in Nigeria. Given the obvious disparity, many parents would want their children or wards to be educated in universities only. Even within the university system, lecturers see themselves as superior to technologists, who are more practical inclined in these research areas; as such, collaboration is difficult.

There is more to gain in the collaboration of ideas and publicity of efforts, especially during these days of free social media usage. Since the fight against tropical diseases is global, there is a possibility that third world countries can sort for and get collaboration for commercialization of these results, methods, and processes from richer countries. Universities in such countries can take the initiative; they only need to make a move, and the time is now.

Malaria, typhoid, dengue fever, and hepatitis, among others, now exhibit resistance to drugs and confusable symptoms. The implication is that the hybridization of soft computing techniques can no longer be ignored, as their applications are numerous and diverse. More so, given the diverse eating habits of people in different regions of the world, producing different symptoms for similar diseases, personalized therapy is becoming apparent to overcome the problem of vague symptomatology. Again, since soft computing techniques lack the ability to explain their predictions to the understanding of medical practitioners, further research needs to be carried out on how to build the tacit knowledge of experts in soft computing techniques.

This study is not without limitations, most of which, are related to the study procedure. Due to resource constraints, the researchers were only able to access a few research databases such as PubMed, Google-scholar, and other open access sources. Secondly, the search string used may have missed some articles that had different key words. Not finding articles from grey literature may have been as a result of publication bias, which occurs when papers with negative results or those that have not been defended and certified, are not published. Third, our quality assessment criteria were very subjective because we utilized the DARE criteria. However, the quality assessment criteria were independently evaluated by three researchers, which reduced the likelihood of inaccurate results.

REFERENCES

1. Uzoka FME, Nwokoro C, Debele F, Akinnuwesi B, Olaniyan M. AHP model for diagnosis of tropical confusable diseases: Proceedings of the 2017 International Conference on

- Computational Science and Computational Intelligence (CSCI). IEEE Computer Society. 2017;306: 54450888.
2. Driver C. Malaria and its avoidance. *Pract Nurse*. 2009;37:19-28.
3. Djam X, Wajiga G, Kimbi Y, Blamah N. A fuzzy expert System for the management of malaria. *Int J Pure App Sci Tech*. 2011;5:84-108.
4. Christenson JC, Korgenski EK. Principles and practice of pediatric infectious disease. 4th Edition. 2012:1-1744.
5. World Health Organization. WHO research and development blue print: Evaluation of ideas for potential platforms to support development and production of health technologies priority infectious diseases with epidemic potential? 2016.
6. Begg RK, Lai D, Palaniswami M. Computational Intelligence in biomedical engineering. Taylor and Francis Books Inc. CRC Press. 1st Edition. 2008;1-392.
7. Chaturvedi DK. Soft Computing techniques and its applications in electrical engineering Studies in Computational Intelligence (SCI). Springer Sci Busin Media. 2008;103:363-380.
8. Yardimci A. Application of soft computing to medical problems: 9th International conference on intelligent systems design and application. IEEE. 2009;168:614-619.
9. Zadeh LA. The calculus of fuzzy IF/THEN rules. *AI Expert*. 1992;7:1-27.
10. Saramourtsis A, Damousis J, Bakirtzis A, Dokopoulos P. Genetic Algorithm solution to the economic dispatch problem-application to the electrical power grid of Crete Island, Aristotle University of Thessaloniki. *Depart Elect Comp Engg Power Sys Lab*. 1999;540:1-36.
11. Bakare GA, Aliyu UO, Venayagamoorthy GK, Shu-Aibu YK. Genetic algorithms based economic dispatch with application to coordinate of Nigerian thermal power plants. *IEEE Transact Power Sys*. 2005;148:9725.
12. Shahriar A, Kainfar J. A Hybrid neuro-genetic approach to short term traffic volume prediction. *Int J Civil Engg*. 2009;7:41-48.
13. Zadeh LA. Fuzzy Sets. *Info Control*. 1965;8:338-353.
14. Dubois D, Prade H. An introduction to fuzzy systems. *Chin Chim Acta* 1998;270:1-29.
15. Jimoh RG, Awotunde JB, Babatunde AO, Ameen AO, Fatai OW, James RT. Simulation of medical diagnosis system for malaria using fuzzy logic. *Comput Info System Devel Allied Res J*. 2014;5:1-18.
16. Sharma P, Singh DBV, Bandil MK, Mashre N. Decision Support System for Malaria and Dengue Disease Diagnosis (DSSMD). *Int J Info Comput Tech*. 2013;3:633-640.
17. Zadeh LA, Nikravesh M, Loiz V. Fuzzy logic and the internet. *Stud Fuzz Soft Comput*. 2004;137:1-25.
18. Arabacioglu BC. Using fuzzy inference system for architectural Space analysis. *Applied Soft Comput*. 2010;10:926-937.
19. Obot OU, Uzoka FME. Fuzzy rule-based framework for the management of tropical diseases. *Int J Med Engg Info*. 2008;1:7-17.
20. Samuel OW, Omisore MO, Ojokoh BA. A web-based decision support system, driven by fuzzy logic for the diagnosis of typhoid fever. Elsevier. 2013;40:4164-4171.
21. Jianchao Z, Jing J, Zhihua C. Particle swarm optimization algorithm. *Sci Pub Com Beijing*. 2004;4:294-301.
22. Yonggang C, Fengjie Y, Jigui S. A new particle swarm optimization algorithm. *J Jilin Univ*. 2006;24:181-183.
23. Bishop CM. Pattern recognition and machine learning. *Info Sci Stat*. 2006;6:6-758.
24. Kitchenham B. Procedures for understanding systematic reviews: Joint Technical Report. Computer Science Department, Keele

- University (TR/SE-0401) and National ICT Australia. *J Soft Engg Appl*. 2004;25:353-401.
25. Webster J, Watson R. Webster and Watson literature review. *MIS Quarterly*. 2002;26:1-69.
26. Evidence Based Software Engineering (EBSE) technical report. Guide lines for performing systematic literature reviews in Software Engineering. 2007.
27. Hotez PJ, Alvarado M, Basanez MG, Bolliger I, Bourne R, Boussinesq M, et al. The global burden of disease study 2010: Interpretations and implications for the neglected tropical diseases. *PLOS Neg Trop Dis*. 2014;8:e2865.
28. Murray NEA, Quam MB, Wilder-Smith A. Epidemiology of dengue: Past, present and future prospects. *Clin Epidemiol*. 2013;5:299-309.
29. Center for Review and Development (CDR). Systematic reviews: CRD's guide for undertaking reviews in health care. 2009.
30. Centre for Reviews and Dissemination (CRD). What are the criteria for the inclusion of reviews on DARE? 2007.
31. Kitchenham B, Brereton P, Turner M, Niazi M, Pretorius P, Budgen D. The impact of search procedures for systematic literature reviews- a participant-Observer case study: Proceeding of symposium on Empirical software Engineering and Metrics (ESEM). *Ass Comp Mach*. 2009;9:336-345.
32. Chen HL, Liu DY, Yang B, Liu J, Wang G. A new hybrid method based on local fisher discriminant analysis and support Vector Machine for Hepatitis disease diagnosis. Elsevier. 2011;38:11796-11803.
33. Faisal T, Ibrahim F, Faib MN. A non-invasive intelligent approach for predicting the risk in dengue patients. Elsevier. 2009;37:2175-2181.
34. Owoseni AT, Ogundahunsi IO. Mobile based fuzzy expert system for diagnosing malaria (MFES). *Int J Info Engg Electr Busin*. 2016;2:14-22.
35. Oluwagbemi O, Oluwagbemi F, Abimbola O. Ebola fuzzy informatics systems on the diagnosis, prediction and recommendation of appropriate treatment of Ebola Virus Disease (EVD). Elsevier. 2016;2:12-37.
36. Atatah PE, Kisavi-Atatah CW. Globalization 2: Revisiting neglected tropical diseases such as polio, dengue fever, and particularly Ebola. *Open J Soc Sci*. 2016;4:1-13.
37. World Health Organization. Department of Communicable Disease Surveillance. 2013.
38. Byrne JP. Encyclopedia of pestilence, pandemics, and plagues. ABC-CLIO. 2008;2:1-920.
39. McNeil Jr. DG. Fewer Ebola cases go unreported than thought, study finds. *New York Times*. 2015.
40. Panford JK, Hayfron-Acquoh BB, Riverson K. A neuro-fuzzy expert system for malaria diagnosis-The machine learning approach. *Res J Comput Sci*. 2015;2:1-39.
41. Djam XV, Wajiga GM. A novel diagnostic framework: Application of soft computing technology. *Pac J Sci Tech*. 2012;3:1-12.
42. Mansour RF. Using Genetic Algorithm for identification of diabetic retinal exudates in digital color images. *J Intell Learn Sys App*. 2012;4:188-198.
43. Sartakhti JS, Afrabandpey H, Saraee M. Simulated Annealing Least Squares Twin Support Vector Machine (SA-LSTSVM) for pattern classification. *Soft Comput Springer*. 2016;21:4361-4373.
44. Rusdah EW. Review on data mining methods for tuberculosis diagnosis. *Inform Sys Int Conf*. 2013;1:563-568.
45. Obot, OU, Uzoka FME. A framework for application of neuro-case-base-rule base hybridization in medical diagnosis. *App Soft Comput*. 2009;9:245-253.
46. Chen KH, Wang KJ, Tsai ML, Wang KM, Adrian AM, Cheng WC, et al. Gene selection for cancer identification: a decision tree model empowered by particle swarm optimization algorithm. *BMC Bioinform*. 2014;15:1-49.
47. Afif MH, Heder AR, Abdel-Haid TH, Mahdi YB. SS-SVM (3SVM): A new classification method for hepatitis disease diagnosis. *Int J Adv Comp Sci App*. 2013;4:1-15.
48. Shaukat K, Masood N, Mehreen S, Azmeen U. Dengue fever prediction: A data mining problem. *J Data Mining Genom Proteom*. 2015;6:1-5.
49. Seera M, Lim CP. A hybrid intelligent system for medical data classification. Elsevier. 2014;41:2239-2249.
50. Asogbon MG, Samuel OW, Omisore MO, Awonusi O. Enhanced neuro-fuzzy system based on genetic algorithm for medical diagnosis. *J Med Diagnos Methods*. 2016;5:1-23.
51. Abiyev RH, Abiyade S. Diagnosing Parkinson's Disease using fuzzy neural system. *Int J Sci Res*. 2016;3:1-9.
52. Singh S, Singh S, Samson, Singh M. Recommender system for detection of dengue fever using fuzzy logic. *Int J Comput Engg Tech*. 2016;7:44- 52.
53. Alanis-Reyes EA, Hernández-Cruz JL, Cepeda JS, Castro C, Terashima-Marín H, Conant-Pablos SE. Analysis of machine learning techniques applied to the classification of masses and microcalcification clusters in breast cancer computer-aided detection. *J Cancer Ther*. 2012;3:1020-1028.
54. Wolpert DH. The supervised learning "No free lunch theorem". NASA Ames Research Center, Moffett Field. 2001.
55. Srivastava P, Sharma N. A spectrum of soft computing model for medical diagnosis. *App Math Info Sci*. 2014;8:1225-1230.
56. Bologna G, Hayashi Y. A comparison study on rule extraction from neural network ensembles, boosted shallow trees and SVM. *App Comput Intellig Soft Comput*. 2018;18:1-20.
57. Yedjour D, Yedjour H, Bonyettou A. Explaining results of artificial neural networks. *J App Sci*. 2011;11:2855-2860.
58. Seera M, Lim CP. A hybrid intelligent system for medical data classification. *Expert Sys App*. 2013;9:1-22.
59. Seera M, Lim CP. Condition monitoring of induction motors: A review and an application of an ensemble of hybrid intelligent models. *Exp Sys App*. 2014;41:4891-4903.
60. Rajasekaran S, Vijayalakshmi GA. Neural network, fuzzy logic, genetic algorithm-synthesis and application. PHI Learn. 5th Edition. 2011;1439.
61. Ponder EL, Freundlich JS, Sarker M, Ekins S. Computational models for neglected diseases: Gaps and opportunities. *Pharma Res*. 2014;31:271-277.
62. Suleiman MM. Neglected tropical diseases in Nigeria, situation analysis. *Adv Soc Sci Res J*. 2016;3:1-10.
63. World Health Organization. Department of communicable disease surveillance and response. 2013.
64. Uniting to combat neglected tropical diseases. The London Declaration. 2010.
65. Uzoka FME, Akinnuwesi B, Amoo T, Aladi F, Fashoto S, Olaniyan M, et al. A framework for early differential diagnosis of tropical confusable diseases using the fuzzy cognitive map engine. *Conf Paper*. 2016;16:345-78.
66. Uzoka FME, Obot OU, Barker K, Osuji J. An experimental comparison of fuzzy logic and analytic hierarchy process for medical decision support systems. *Comp Method Prog Biomed*. 2011;103:10-27.
67. Lin R, Cuang C. A Hybrid diagnosis model for determining the types of the liver disease. *Comput Biol Med*. 2010;40:665-670.

68. Inayat Z, Gani A, Anuar NB, Khan MK, Anwar S. Intrusion response systems: Foundations, design, and challenges. *J Net Comput App*. 2016;62:53-74.
69. Hand D, Manila H, Smyth P. Principles of data mining. MIT Press. 1st Edition. 2001;1-578.
70. Taneja GSA. Study of classifiers in data mining. *Int J Comput Sci Mob Compu*. 2015;3:263-269.
71. Lin R. An intelligent model for liver disease. *Artific Intell Med*. 2009;47:53-62.
72. Dutton G. The rise of precompetitive collaboration. *Life Sci Lead Magazine*. 2012;12:40-41.
73. Hotez PJ, Pecoul B. Manifesto for advancing the control and elimination of neglected tropical diseases. *PLOS Neglect Trop Dis*. 2010;4:e718.
74. Gamo FS, Sanz LM, Vidal J, de Cozar C, Alvarez E, Lavandera JL, et al. Thousands of chemicals starting points for ant malaria lead identification. *Nature*. 2010;465:305-310.
75. Fitzgerald A, Pappalardo K, Austin A. Building the infrastructure for data access and reuse in collaborative rework: An analysis of the legal context. *OAK law Project*. 2007;7:1-290.
76. Asogbon MG, Samuel OW, Omisore MO, Awonusi O. Enhanced neuro-fuzzy system based on Genetic Algorithm for medical diagnosis. *J Med Diagnos Methods*. 2016;5:37-49.
77. Ibrahim D. An overview of soft computing. *Procedia Comp Sci*. 2016;102:34-38.
78. PROSPERO. Guidance notes for registering a systematic review protocol with PROSPERO. *International Prospective register of systematic reviews*. 2013.