A Review on Information Technology in Obesity Epidemic: Prediction and Prevention

Lalitha Saroja Thota *
Research Scholar, Department of Computer Science,
Acharya Nagarjuna University, Guntur, AP, India

Allam Appa Rao
Director, CRRao AIMSCS,
University of Hyderabad campus, Hyderabad, AP, India

Abstract—The purpose of this review is to provide the needed understanding of the obesity problem, introduce the use of Information Technology on obesity prediction and prevention, and describe current efforts in some areas of Information Technology with a brief review of journal papers. Related journal articles on obesity prediction and prevention by previous researchers were downloaded from various databases including IEEE, Science Direct, SpringerLink, SCOPUS etc. After the research papers were downloaded, there were categorized into five different areas of information technology which are: data mining & ANN, Bioinformatics, image processing, Geographic information system and web applications, then the review process started. As a disease, obesity is complex and its solutions too are difficult. The epidemic of obesity needs to be tackled at several levels: the individual level, the community level, government level and to some extent in research & technology level. The review briefly discusses the obesity problem and risk factors for obesity. The five areas of information technology presented – data mining & ANN, bioinformatics, image processing, geographic information system and web applications are only some of active fields where research is in progress for obesity prediction and prevention. These areas each have their weaknesses and strengths. Further improvements of the techniques are necessary to fully resolve the obesity prediction problem. Some more areas of research are coming up in the field of information technology on obesity management. This review reminds the reader about co-morbidity of obesity and why it is important to prevent it. This review has shown that information technology can be used to successfully predict and prevent obesity to some extent; however interventions to tackle obesity such as improving diets and increasing physical activity can help reduce the obesity to great extent.

Keywords—epidemic, obesity, BMI, Information technology, data mining & ANN, Bioinformatics, image processing, Geographic information system, web applications

I. INTRODUCTION

There is a famous saying that Health is Wealth and Healthy mind lives in a healthy body. But today most of the people find it difficult to reduce their weight and are becoming more prone to blood pressure, cardiovascular diseases and diabetes and the core reason behind this is obesity. There are no quick-fire solutions to the looming threat of the obesity epidemic. The answers lie in innovative thinking and research efforts in different domains of science and technology to solve obesity problem. Overweight and obesity are defined as abnormal or excessive fat accumulation that presents a risk to health. A crude population measure of obesity is the body mass index (BMI), a person’s weight (in kilograms) divided by the square of his or her height (in meters). A person with a BMI of 30 or more is generally considered as obese. A person with a BMI equal to or greater than 25 is considered as overweight. Obesity can be classified as one of the dangerous illness in the world as the number of the obese person keep on increasing. Obesity is becoming increasingly common, with at least 400 million obese adults worldwide and the World Health Organization (WHO) projecting that this statistic will reach 700 million in another five years. According to a WHO report in February 2010, at least 2.6 billion people died each year because of obesity or overweight. According to the WHO, 65% of the world’s population lives in a country where overweight and obesity kills more people than underweight”. Moreover it tells 44% of diabetes, 23% of ischemic heart disease and more than 7% of certain cancers, globally, are attributable to obesity [1].

In recent years, the escalating worldwide prevalence of obesity is considered as one of the most serious issues. This is because obesity is significantly associated with diabetes, heart disease, cancer, high blood pressure, and high cholesterol [2][3]. Regarding mentioned data it is not surprising that obesity and overweight have also a significant economic impact on the health care systems. Associated medical costs may involve direct and indirect costs. Direct medical costs may include preventive, diagnostic, and treatment services related to obesity. Indirect costs relate to development of chronic diseases and also to acute health complications where drugs and medical care are important, in many cases hospitalizations are needed and also mortality is very frequent among serious health complications. Obesity severely threatens quality of life, limiting the ability to move and work, navigate public places. Both obesity and its co morbidities lead obese patients to hospitals at greater rates than patients of normal weight, and these factors, along with social pressures, motivate patients, physicians and researchers to pursue solutions for obesity, its co morbidities, or both.

Information technology IT is the study, design, development, application, implementation, support or management of computer-based information systems. It is used in various fields such as Exploring Natural Resources, Weather
II. RISK FACTORS FOR OBESITY

A People have too many daily habits to detect all of them, such as brushing of teeth, walking a dog, eating spaghetti, and so on. Therefore, we focused on habits that are considered as the risk factors of overweight for medical reasons. Some risk factors for overweight and obesity have been previously reported [4]. Eating quickly, gorging, and binge eating have been associated with total energy intake, satiety, and insulin resistance, which links to body weight [5]. Examinations for verifying the association between quickly eating and overweight are also reported [6].

Nutritionists and nutrition texts recommend breakfast as an important part of healthy eating habits, and breakfast is associated with overweight. This is because breakfast is associated with maintenance of a constant blood glucose concentration, prevention of hunger, and subsequent overeating later in the day [7]. Energy-dense, nutrient-poor foods and snacks are increasingly consumed by adolescents, and this phenomenon has been linked to higher energy intakes and inadequate intake of certain micronutrients, which may increase the risk of developing obesity or other chronic diseases such as cancer and cardiovascular disease [8][9]. It is reported that circadian timing of food intake contributes to weight gain because of clock gene [10]. Particularly evening and night eating are associated with obesity [11][12]. Thus, bedtime snacks and night eating should be avoided. Moreover, you should not eat too much at evening, and should have enough intervals between dinner and getting to bed. Sleeping is associated with obesity, and some researchers suggest that short sleep duration is associated with obesity. Chaput et al. reported that short sleep duration is associated with reduced leptin levels, which is related to appetite and metabolism, and increased adiposity [13]. Furthermore, examinations for verifying the association between short sleep duration and obesity had been conducted [4].

Alcohol intake also associated with obesity because of following reasons. The first reason is related to appetite. Alcohol intake improves appetite, and it causes energy imbalance. In addition, the fact that snacks to go with alcohol have high calorie content causes excess energy intake. On the other reason, excessive alcohol intake increases neutral fat. Alcohol is decomposed in a liver to some extent, but excessive intake makes aliphatic acid synthesize neutral fat in the liver. It causes an increase in fat mass. Chaput et al. also verified the association between alcohol intake and obesity [4].

Nonparticipation in high-intensity physical exercise is also associated with obesity, as it is reported [4][14]. WHO suggests that people should engage in adequate levels of physical activity throughout their lives for preventing obesity and some other diseases [1]. The reason why nonparticipation in physical exercise is associated with weight-gain is because physical exercise improves energy expenditure, and also it is related to basal metabolism. Weight gain is basically caused by energy intake in excess of energy expenditure [6]. Thus, calorie balance is important part of weight-gain. Thus obesity is associated with various habits of our daily life, e.g. short sleep duration, low dietary calcium intake, high alcohol intake, and nonparticipation in high-intensity physical exercise [4], and changing these habits in order to reduce disease burden is recognized as an effective solution [15][16].

III. METHODOLOGY

Several steps have been done in producing this paper. Related journal articles on obesity prediction and prevention by previous researchers were downloaded from various databases including IEEE, Science Direct, Springer Link, SCOPUS etc. We had downloaded about 30 research papers involving information technologies and obesity disease, there were categorized into five different areas of information technology which are: data mining & ANN, Bioinformatics, image processing, Geographic information system and web applications, then the review process started.

IV. DISCUSSION

A Information technology, IT, is now frequently used for health care [17], since it is useful for monitoring and logging health information with database, DB, technology [18]. In this section we describe different areas of information technology and present a brief review of efforts made in obesity prediction and prevention.

4.1 Data mining and ANN

Although data mining is a relatively new term, the technology is not. Companies have used powerful computers to sift through volumes of supermarket scanner data and analyze market research reports for years. However, continuous innovations in computer processing power, disk storage, and statistical software are dramatically increasing the accuracy of analysis while driving down the cost. Generally, data mining (sometimes called data or knowledge discovery) is the process of analyzing data from different perspectives and summarizing it into useful information - information that can be used to increase revenue, cuts costs, or both. Thus data mining refers to extracting or mining knowledge from large amounts of data. Summarization, classification and prediction, regression, association, clustering etc are some of the data mining functionalities used to extract hidden knowledge.

The information and knowledge gained can be used for applications ranging from market analysis, fraud detection and customer retention, to production control and science exploration. Several types of analytical software are available: statistical, machine learning, and neural networks. Generally, any of four types of relationships are sought: Classes, Clusters, Associations and Sequential patterns. The methods of artificial intelligence particular artificial neural networks (ANN) could handle this problem efficiently because of their speed and ability to handle large amount of input
data. Today mostly used forms of the ANNs are fully connected feedforward supervised ANNs with the error back propagation learning method (BP). Such networks are called back propagation networks (BPANNs). The learning method is the gradient descent method.

Childhood obesity is disease sue generis and is background for development of different illness in further life. If obesity could be predicted it could be prevented in most cases and degenerative and chronic diseases caused by it avoided. But such prediction is very complex medical and sociological problem. A new generation of artificial neural network is applied to handle this difficult problem successfully by Bojan Novac and Martin Bige [19]. There are many data mining techniques like Artificial Neural Network (ANN), Naïve Bayes classification, and decision tree that have been applied to predict childhood obesity by Adnan et al [20]. Muhammad Hariz Muhammad Adnan et al [21] propose a framework for childhood obesity classification and prediction using NBtree. The framework uses hybrid Naïve Bayes and decision tree technique NBtree, Naïve Bayes and decision tree classifier techniques are good classification and prediction techniques individually. By combining both techniques, the results of prediction are more accurate than previous prediction techniques.

GWAS studies have been successful in finding genetic determinants of obesity. To translate discovered genetic variants into new therapies or prevention strategies, molecular or physiological mechanisms need to be discovered. One strategy is to perform data mining of data sets with detailed phenotypic data, such as those present in dbGAP (database of Genotypes and Phenotypes) for hypothesis generation. Christine W. Duarte et al [22] propose a novel technique that combines the power and computational efficiency of existing Bayesian Network (BN) learning algorithms with the statistical rigor of Structural Equation Modeling (SEM) to produce an overall system that searches the space of potential networks and evaluates promising candidates using standard SEM model selection criteria.

Identification of gene-gene and gene-environment interactions that contribute in the onset of a multi-factorial disease supports the prevention of diseases like the Cardiovascular Disease (CVD). Body Mass Index (BMI), a measure of human obesity, is an independent risk factor of CVD. Furthermore, it is known that a subject's BMI is affected both by his/her lifestyle, e.g. nutrition, and genetic profile. Ioannis K. Valavanis et al [23], aim to predict a subject's onset of obesity using lifestyle and genetic information. The prediction system is performed by a computational intelligence base system using a Parameter Decreasing Method (PDM) combined with an Artificial Neural Network (ANN) aiming to identify factors related to the onset of obesity, as an example of CVD risk factor, using persons’ information regarding sex, nutrition habits and genetic variations.

4.2 Bioinformatics

Computers and information technology have become indispensable tools for most of us. This is particularly true in biological research, where scientists increasingly apply information technology to biological problems — a science called bioinformatics. Bioinformatics is the application of computer technology to the management of biological information. Computers are used to gather, store, analyze and integrate biological and genetic information which can then be applied to gene-based drug discovery and development. Bioinformatics provides central, globally accessible databases that enable scientists to submit search and analyze information. It offers analysis software for data studies and comparisons and provides tools for modeling, visualizing, exploring and interpreting data. Scientists have developed many bioinformatics technologies because of the need to understand DNA, the code of life, and other components of living organisms. Bioinformatics provides novel methods to store analyze and visualize this information — creating new knowledge to enhance our standard of life. Bioinformatics helps us to visualize invisible structures such as proteins and to learn more about their work and function. This leads to insight into central questions of life: How do organisms work? How did life develop? How can new treatments against diseases such as obesity be developed? The science of Bioinformatics, which is the melding of molecular biology with computer science, is essential to the use of genomic information in understanding human diseases like obesity and in the identification of new molecular targets for drug discovery for obesity. In recognition of this, many universities, government institutions and pharmaceutical firms have formed bioinformatics groups, consisting of computational biologists and bioinformatics computer scientists. Such groups will be key to unraveling the mass of information generated by large scale sequencing efforts underway in laboratories around the world.

Bioinformatics has been in the focus since recent years for unraveling the structure and function of complex biological mechanisms. The analysis of primary gene products has further been considered as diagnostic and screening tool for disease recognition. Such strategies aim at investigating all gene products simultaneously in order to get a better overview about disease mechanisms and to find suitable therapeutic targets. Recently Gerken et al [24] performed bioinformatic analysis and reported that the variants in the fat mass and obesity associated gene are associated with increased body mass index in humans. Barcelo-Batllo et al [25] utilizes the DIGE and Bioinformatic analysis for identification of potential drug targets of tungstate, DIGE analysis identified 20 proteins as tungstate obesity-direct targets, involved in: Krebs cycle, glycolysis, lipolysis and fatty acid oxidation, electron transport and redox.

Most common complex traits such as obesity, hypertension, diabetes, and cancers are known to be associated with multiple genes, environmental factors, and epistasis. Recently, the development of advanced genotyping technologies allows us to perform the genome-wide association studies (GWAS). For detecting the effects of multiple genes on complex traits, many approaches have been proposed for GWAS. Multifactor dimensionality reduction (MDR) proposed by Ritchie et al. [26] is one of the powerful methods for detecting epistasis, which detects high order interactions among genes. However, MDR is computationally intensive due to its requirement of testing all possible n-way combinations. Thus, it is not practically feasible to apply MDR directly to analyzing large-scale GWAS data.
Sungyoung Lee et al [27] propose an efficient strategy to perform MDR analysis for GWAS data. Genome-wide association studies (GWAS) provide a new and powerful approach to investigate the effect of inherited genetic variation on risks of complex diseases like obesity. With recent advances in genotyping technology, genome-wide association studies are now becoming a reality.

Kanchana Narayanan and Jing Li [28] implemented a web application tool named MAVEN for Management, Analysis, Visualization and results sharing of GWA data using cutting edge technologies. The human body plays host to thousands of bacterial species in a variety of ecosystems. Until recently, microbial communities have been impossible to investigate thoroughly, as the vast majority of bacteria cannot be cultured through laboratory techniques. Recent studies in this context have revealed that human obesity has a microbial component: obese gut microbiomes are distinct from the lean population. This result indicates potential therapeutic approaches to treating obesity by manipulating gut microflora. However, limited knowledge of the microbial interactions in the gut hinders our ability to design future experiments or effective treatments. New technologies like bioinformatics (e.g. high-throughput sequencing, 16S rRNA surveys) allow us to deeply sample the genetic content of a microbial environment in order to estimate its overall composition and functional capacity. James Robert White and Mihai Pop [29] use 16S rRNA time-series data sequence from obese individuals on a one-year diet and employ mathematical modeling to study microbial population dynamics in the human gut. The model indicates several interspecific interactions in this microbial community and the impact of prebiotic and probiotic therapies for obesity through simulation.

4.3 Image Processing

Image processing is a method to convert an image into digital form and perform some operations on it, in order to get an enhanced image or to extract some useful information from it. It is a type of signal dispensation in which input is image, like video frame or photograph and output may be image or characteristics associated with that image. Image processing includes: Image display and printing, Image editing and manipulation, Image enhancement, Feature detection and Image compression. The purpose of image processing is divided into five groups. They are: Visualization - Observe the objects that are not visible. Image sharpening and restoration - To create a better image. Image retrieval - Seek for the image of interest. Measurement of pattern – Measures various objects in an image. Image Recognition – Distinguish the objects in an image. Image processing basically includes the following three steps. Importing the image with optical scanner or by digital photography. Analyzing and manipulating the image which includes data compression and image enhancement and spotting patterns that are not to human eyes like satellite photographs. Output is the last stage in which result can be altered image or report that is based on image analysis. It is among quickly growing technologies today, with its applications in various aspects of a business as well as Biology, Astronomy, Medicine, Security, Biometrics, Satellite Imagery etc. Image Processing forms core research area within engineering and computer science disciplines too.

TM Kriegel et al [30] carried experiments for identification of diabetes- and obesity associated proteomic changes in spermatozoa by difference gel electrophoresis. In their study Image acquisition and computer-assisted analysis are used for the identification of identical spots in different gels by employing the fully automated computer-assisted alignment module (Batch Processor) while the Biological Variation Analysis (BVA) program was applied to manually revise matches and to provide statistical data on protein expression levels based on Student’s t-test.

4.4 Geographic information system

A geographic information system (GIS) integrates hardware, software, and data for capturing, managing, analyzing, and displaying all forms of geographically referenced information. GIS allows us to view, understand, question, interpret, and visualize data in many ways that reveal relationships, patterns, and trends in the form of maps, globes, reports, and charts. GIS gives us a new way to look at the world around us. With GIS you can: Map Where Things Are, Map Quantities, Map Densities, Find what’s inside, Find what’s nearby and Map Change. The benefits of GIS generally fall into five basic categories: The benefits of GIS generally fall into five basic categories: Cost Savings and Increased Efficiency, Better Decision Making, Improved Communication, Better Recordkeeping and Managing Geographically.

Lauren Rosenshein and Nigel Waters [31] use a GIS based cluster analysis of Los Angeles County school level data on percentages of overweight 5th graders to expose clusters of obese children in the area. In their study, each school is associated with a school polygon, which represents the geographical area that the school serves. These school polygons are used in the cluster analysis to locate areas with low accessibility to healthy foods. Several indicators will be used to establish an accessibility value, including distance to closest supermarket and the number of supermarkets within walking distance. The relationship between clusters of overweight 5th graders and low supermarket accessibility are examined. The accessibility of healthy food in certain areas of the county is leading to unhealthy diets cause obesity in certain communities, and thus negative health outcomes for adolescents.

4.5 Web applications

Over the past decade or so, the web has been embraced by millions of businesses as an inexpensive channel to communicate and exchange information with prospects and transactions with customers. From a technical view-point, the web is a highly programmable environment that allows mass customization through the immediate deployment of a large and diverse range of applications, to millions of global users. Two important components of a modern website are flexible web browsers and web applications; both available to all and sundry at no expense. Web applications are, therefore, computer programs allowing website visitors to submit and retrieve data to/from a database over the Internet.
using their preferred web browser. The data is then presented to the user within their browser as information is generated dynamically (in a specific format, e.g., in HTML using CSS) by the web application through a web server. Websites and related web applications must be available 24 hours a day, 7 days a week, to provide the required service to customers and other stakeholders. Web applications are, therefore, fundamental to businesses for leveraging their online presence thus creating long-lasting and profitable relationships with prospects and customers. No wonder web applications have become such a ubiquitous phenomenon using BMI index in dealing obesity identification. However, due to their highly technical and complex nature, web applications are a widely unknown and a grossly misunderstood fixture in our everyday cyber-life. Obesity is associated with various habits of our daily life, e.g., short sleep duration, high alcohol intake, and nonparticipation in physical exercise. Therefore, changing these habits in order to reduce body weight is recognized as an effective solution. There are some web applications in effort to treat obesity, which verify the usefulness of web applications for health-care support. Yoshitaka kato et al [32] had developed an obesity prevention system that helps a user to change his lifestyle with a web application and health care devises. The web application of this system detects whether the user has the habits considered as risk factors of obesity. After the detection, this application analyzes first-order and second order factors of weight gain as the higher-order factor analysis. It displays the results of analysis in forms of graph and list, in order to encourage the user to change their lifestyle by him.

V. CONCLUSIONS

As a disease, obesity is complex and its solutions too are difficult. The epidemic of obesity needs to be tackled at several levels: the individual level, the community level, government level and to some extent in research & technology level. Obesity needs to be addressed at the socioeconomic level as well as in the biological and the behavioral arenas. The review briefly discusses the obesity problem and risk factors for obesity. The five areas of information technology presented – data mining & ANN, bioinformatics, image processing, geographic information system and web applications are only some of active fields where research is in progress for obesity prediction and prevention. These areas each have their weaknesses and strengths. Further improvements of the techniques are necessary to fully resolve the obesity prediction problem. Some more areas of research are coming up in the field of information technology on obesity management. This review reminds the reader about co-morbidity of obesity and why it is important to prevent it. This review has shown that information technology can be used to successfully predict and prevent obesity to some extent. However interventions to tackle obesity such as improving diets and increasing physical activity can help to reduce the obesity to great extent.

REFERENCES

[4] Jean-Philippe Chaput, Claude Leblanc, Louis Perusse, Jean-Pierre Despres, Claude Bouchard, and Angelo Tremblay, Risk Factors for Adult Overweight and Obesity in the Quebec Family Study: Have We Been Barking Up the Wrong Tree?, Obesity;volume 17; number 10; October, 2009.
[13] Jean-Philippe Chaput, Jean-Pierre Despres, Claude Bouchard, and Angelo Tremblay, Short Sleep Duration is Associated with Reduced Leptin Levels and Increased Adiposity: Results from the Quebec Family Study, Obesity, Vol. 15; No. 1: January 2007.
AUTHORS PROFILE

Lalitha Saroja Thota received M.Tech in Software Engineering from Jawaharlal Nehru Technological University, Hyderabad, India in 2010 and M.Sc in Computer Science from Anna University, Chennai, Tamilnadu, India in 2008. She is pursuing Ph.D in Computer Science and Engineering from Acharya Nagarjuna University, Guntur, A.P, India. Her areas of interest are bioinformatics and data mining. She has more than 7 research papers in international journals and conferences to her credit.

Allam Appa Rao received Ph.D in Computer Engineering from Andhra University, Vishakapatnam, India in 1984. He has four decades of professional experience. His primary area of research is Bioinformatics. He has 60+ strong research team. 40 scholars were awarded Ph.D degrees under his guidance. He is currently Director, CRRao AIMSCS, University of Hyderabad campus, Hyderabad, A.P, India. He has more than 150 research papers in international journals and conferences to his credit.