

Review

Not peer-reviewed version

Applying Artificial Intelligence to Childhood Obesity: DMT2 and MASLD Risk Predictive Models

[Marianna Amitrano](#)*, Mario Emiliano, [Umberto Paolo Santoro](#)

Posted Date: 4 June 2026

doi: 10.20944/preprints202606.0359.v1

Keywords: artificial intelligence; childhood obesity; diabetes type 2; MASLD; predictive models; risk assessment; prevention



Preprints.org is a free multidisciplinary platform providing preprint service that is dedicated to making early versions of research outputs permanently available and citable. Preprints posted at Preprints.org appear in Web of Science, Crossref, Google Scholar, Scilit, Europe PMC, OpenAlex.

Copyright: This open access article is published under a [Creative Commons CC BY 4.0 license](#), which permit the free download, distribution, and reuse, provided that the author and preprint are cited in any reuse.

Disclaimer/Publisher's Note: The statements, opinions, and data contained in all publications are solely those of the individual author(s) and contributor(s) and not of MDPI and/or the editor(s). MDPI and/or the editor(s) disclaim responsibility for any injury to people or property resulting from any ideas, methods, instructions, or products referred to in the content.

Review

Applying Artificial Intelligence to Childhood Obesity: DMT2 and MASLD Risk Predictive Models

Marianna Amitrano ^{1,*}, Mario Emiliano ¹ and Umberto Paolo Santoro ²

¹ Hospital "San Giovanni di Dio", Unit of Pediatrics and Neonatology, Frattamaggiore, Italy

² University of Naples "Federico II", Department of Public Health, Naples, Italy

* Correspondence: marianna.amitrano@aslnapoli2nord.it

Abstract

Background: Pediatric obesity represents a public health problem, given its pandemic spread in recent decades and the presence of complications already in childhood that affect the prognosis in adulthood. Obesity is a complex and multifactorial disease with a bio-psycho-social etiology, and several predisposing factors are still being discovered. Many obesity-related comorbidities are detected already in childhood: pre-diabetes and type 2 diabetes (T2DM), Metabolic Dysfunction-Associated Steatotic Liver Disease (MASLD), hypertension, sleep apnea, sarcopenia, and osteoarticular disorders. To date, lifestyle modification and the Mediterranean diet remain the cornerstones of primary and secondary obesity prevention. Artificial intelligence (AI) represents a new tool for healthcare professionals to improve the early diagnosis and treatment of childhood obesity complications. **Methods:** Articles, reviews, consensus statements, guidelines, meta-analyses, and editorials published on "PubMed," "NIH-National Library of Medicine," and "Google Scholar" were analyzed. The keywords used were: "obesity", "globesity", "children", "comorbidities", "healthcare", "socioeconomic status," "MASLD," "metabolic comorbidities," "diabetes," "artificial intelligence", "machine learning", "deep learning", and "multiomics". **Results:** By synthesizing multi-omic information—comprising the genome, epigenome, metabolome, transcriptome, and microbiota—alongside social and psychological metrics, artificial intelligence (AI) can forecast the probability of obesity development and facilitate the timely detection of complications within vulnerable populations. Utilizing machine learning (ML) and deep learning techniques, researchers have pinpointed specific metabolites, intestinal flora, neurotransmitters, and neurological areas linked to obesity's emergence. Furthermore, these methods have detected SNPs in genes related to carbohydrate and lipid metabolism, energy balance, and the hunger-satiety regulatory systems that contribute to obesity susceptibility. Various risk assessments and prognostic models rely on ML; these include algorithms designed to evaluate the likelihood of MASLD or diabetes progression in pediatric patients with obesity. Consequently, through ML-driven software, it is feasible to conduct remote surveillance of a patient's nutritional habits and exercise routines, or to suggest bespoke dietary regimens tailored to individual patient profiles. **Conclusions:** Artificial intelligence has the potential to aid healthcare providers in enhancing the management of childhood obesity through personalized medical strategies, although various concerns regarding access, data privacy, and digital literacy remain unresolved.

Keywords: artificial intelligence; childhood obesity; diabetes type 2; MASLD; predictive models; risk assessment; prevention

1. Introduction

Childhood obesity represents a therapeutic challenge for healthcare professionals, and its prevalence poses a public health problem for current and future generations. In pediatric age, we define obesity in different ways, depending on the age. Up to 24 months, obesity is defined as a weight-to-length ratio above the 99th percentile on the 2006 WHO growth charts. Between the ages

of 2 and 5 years, obese children are defined as having a body mass index (BMI) above the 99th percentile on the 2006 WHO charts. Over the age of 5 years, obese children are defined as having a BMI above the 97th percentile on the 2007 WHO charts. If this value falls between the 85th and 97th percentile, they are considered overweight; if it exceeds the 99th percentile, they are considered severely obese, as the incidence of complications is higher in this population group[1].

Obesity is divided into primary or idiopathic forms, related to a greater intake of calories than energy expenditure, and secondary forms caused by diseases for which obesity is the phenotype. Recognizing the signs and symptoms of diseases that manifest with secondary obesity is essential for proper therapeutic and care management. Among the signs that suggest the presence of underlying diseases, the most significant are: onset of the clinical picture before 5 years of age, rapid progression of obesity, association with signs attributable to a genetic syndrome, dysmorphic features, cognitive delay, short stature or slow growth associated with weight gain, and use of psychotropic medications. Among the genetic forms of obesity, we distinguish between syndromic and monogenic forms. The most common cause of genetic obesity is Prader-Willi syndrome, due in most cases to a paternal deletion of the 15q11-q13 region. Rarer forms include Bardet-Biedl syndrome, Alstrom syndrome, and Cohen syndrome: each is characterized by pathognomonic features that aid the clinician in making the diagnosis. Aneuploidies such as Turner syndrome, Down syndrome, and Klinefelter syndrome can also be associated with obesity. However, secondary forms of obesity represent only a small portion of all obese individuals. The vast majority of obese patients have a primary, multifactorial form, due to a completely inappropriate nutrient intake compared to the energy expended through basal metabolism and exercise. It is now known that not only do obese children and adolescents have a worse quality of life and prognosis due to complications in adulthood than the non-obese population, but that these comorbidities and reduced quality of life also start in childhood[1]. Given the prevalence of childhood obesity not only in the West but worldwide, worsened by the COVID-19 pandemic, primary and secondary prevention has become a public health goal, and more and more initiatives are being employed to this end, including the definition of obesity as a disease of great importance. Italy was the first country to recognize obesity as a chronic, progressive, and relapsing disease: in Italy the Pella Law (n. 183) was approved in October 2025, which ensures obese patients free access to obesity care and funds prevention programs.

Globesity: the Trend and the Pillars

According to the last survey in 2022, 5.6% of children under 5 years of age are overweight globally (37 million). 48% of overweight children are concentrated in upper-middle-income countries, demonstrating its prevalence in developing countries as well[2].

From a pathogenetic standpoint, childhood obesity can be divided into two groups: syndromic and nonsyndromic obesity. Nonsyndromic obesity, in turn, includes monogenic forms and polygenic obesity, which is the most common. Polygenic childhood obesity is a complex and multifactorial disease: a combination of predisposing genetic factors, combined with epigenetic changes, hormones and neurobiological mediators, social and psychological factors[1]. Among the genetic risk factors that predispose to obesity, SNPs in different genes have emerged thanks to Genome Wide Association Studies (GWAS): BDNF, NTRK2, SIM1, BBS2, BBS4, SH2B1, SDCCAG8, POMC, PCSK1, MC4R[3], MC3R, NTRK2, PRKD1, FTO, IL6, FHIT and TNF α are genes implicated in the regulation of food intake, fatty acid metabolism, and hormonal regulation. In particular, mutations in POMC, MC4R and PCSK1 genes cause severe monogenic obesity, while the presence of SNPs appears to be associated with a predisposition to an early increase in BMI[4]. However, it seems that the presence of SNPs alone is not sufficient to explain the genesis of obesity, but that these variants must be incorporated into a broader and more interactive model.

Talking about the epigenetic background predisposing to childhood obesity, it's necessary to start from the first 1000 days of life. What happens in this time period, including parents' lifestyle, influences a child's susceptibility to obesity by reprogramming the cell's epigenome, according to the theory of transgenerational epigenetic inheritance: a poor-quality maternal diet, maternal or paternal

obesity, maternal stress, smoking, exposure to environmental pollutants, delivery by c-section, absent or reduced breastfeeding, low or high birth weight for gestational age are associated with greater risk of childhood obesity. Methylation has been found in histone deacetylase genes as well as in the NPY, CR1, CART, CHST8 and FTO genes, in obese subjects compared to controls. Paternal obesity appears to alter the sperm epigenome, and in subjects who have undergone bariatric surgery, these changes appear to be partially reversed. During pregnancy and breastfeeding, obese women show methylation of genes associated with energy metabolism, glucose homeostasis, insulin signaling, and fat storage. For example, methylation of genes, such as the promoter of OR2L12 or CYP2E1, has been found in the cord blood of infants born to obese mothers or those with gestational diabetes[5].

An emerging issue in recent decades is the impact of endocrine disruptors on the metabolism of children and young adults of reproductive age: exogenous chemicals of various kinds that mimic the effects of endogenous hormones and interfere with hormonal signaling pathways. Endocrine disruptors, such as bisphenol A and phthalates, are present in plastics and can be absorbed through food, water, or skin. They can also influence the onset of obesity, among several other diseases. Prenatal exposure to phthalates, parabens, and other phenols appears to influence a child's BMI. Endocrine disruptors such as bisphenol A affect fetal growth and cause epigenetic changes that are transmitted transgenerationally, contributing to the onset of obesity and altered glucose metabolism[4].

From a neurobiological perspective, various mechanisms are involved in childhood obesity, which involves an alteration in the homeostasis of glucose and lipid metabolism, the secretion of digestive hormones, and the signaling pathways for hunger and satiety. Gastrointestinal hormones interact with the central nervous system through receptors on the vagal pathway (ghrelin is the main orexigenic hormone, PYY, GLP-1, OXM, glicentin, CCK, GIP, PP, amylin are anorexigenic) and act synergistically with each other. In obese individuals, energy homeostasis is disrupted due to altered circadian secretion of these hormones, promoting further weight gain. This occurs, for example, through reduced postprandial ghrelin suppression, loss of preprandial ghrelin peaks, and reduced diurnal ghrelin variability, along with reduced fasting and postprandial levels of potent anorexigenic peptides such as PYY and GLP-1, compared to normal-weight subjects. This translates into a reduced sense of satiety and a reliance on frequent snacking. An alteration in the microbiota could also partially explain the altered secretion of intestinal hormones[6].

Leptin, a hormone produced primarily by white adipose tissue and capable of crossing the blood-brain barrier in its free form, is secreted proportionally to an individual's fat mass and decreases during fasting. It can stimulate anorexigenic neurons and inhibit orexigenic neurons in the arcuate nucleus of the hypothalamus, the center of appetite control. In obese individuals, however, hyperleptinemia is accompanied by resistance to its action at this level, likely through diet-induced expression of the SOCS3 and STAT3 genes in POMC neurons. Furthermore, several SNPs have been identified in the leptin signaling pathway that predispose to the onset of obesity by altering the sense of satiety[7].

The hedonistic aspect of appetite instead is regulated by the orbitofrontal cortex, the amygdala, the nucleus accumbens, the dorsal striatum, and other structures of the limbic system. Food addiction, or the compulsion to eat, especially foods high in fat and carbohydrates, fried foods, and sweets, is increasingly being discussed in the literature. The neurobiological mechanisms underlying this behavior have been studied, which appear to focus on the opioid and dopamine systems, which are stimulated by artificially sweetened, high-fat, and salty foods, which trick the brain into perceiving a false sense of satiety[8].

Numerous social factors contribute to the onset of obesity: studies on the family context of obese and overweight children have highlighted poor family nutrition education (from parents and grandparents); family eating habits associated with obesity, such as skipping breakfast, eating little fruit and vegetables, frequent snacking, foods high in fat and carbohydrates, and sugary drinks, especially while sitting in front of screens; poorly defined family routines, such as not establishing mealtimes, sleeptimes, or time spent together; insufficient encouragement of children to participate

in outdoor activities; family dysfunction, social insecurity, low self-esteem, stress; the imposition of restrictive diets, which leads children to consume "forbidden" foods when unsupervised; exposure to junk food advertising; the convivial role of food, and in particular, associating positive emotions such as parties or happy events with overeating; and the consumption of fast food as a sense of emancipation, or in some poorer communities as a legacy of the economic crisis and reduced food availability. In some contexts, mothers are seen as over-nourishing and measure affection in the amount of food they give their children, thus resisting lifestyle changes proposed as a treatment for obesity. Psychosocial stress can also lead to the consumption of food as compensation, especially comfort food, that is, ultra-refined and ultra-processed foods rich in carbohydrates and saturated fats. Comorbidities psychopathological, such as anxiety, depression, autism spectrum disorder, but also reduced self-regulation, impatience, "difficult" temperament can lead to overeating disorders, such as binge eating disorder, associated with overweight and obesity[9].

Complications of Obesity: Focus on Metabolic Dysfunction

There are numerous comorbidities associated with childhood obesity and, contrary to what was once believed, they do not only affect adulthood, but begin and progress already in childhood and the frequency and severity are proportional to the precocity and extent of obesity[1]. The most frequent complications of childhood obesity are cardiometabolic comorbidities, the central drive of which appears to be obesity-related low-grade inflammation, which contributes to the onset of Metabolic Dysfunction-Associated Steatotic Liver Disease (MASLD), insulin resistance and hypertension. Other complications associated with obesity include gastrointestinal comorbidities, such as gastroesophageal reflux disease and cholelithiasis; respiratory comorbidities, such as obstructive sleep apnea syndrome (OSAS), asthma, and hypoventilation syndrome; orthopedic comorbidities, such as slipped capital femoral epiphysis, Blount disease, genu valgum, flat feet and increased susceptibility to fractures; sarcopenia; neurological comorbidities, such as idiopathic intracranial hypertension and chronic migraine; psychological comorbidities like anxiety-depressive symptoms, eating disorders, body dysmorphic disorder, low self-esteem, and suicidal risk. Binge Eating Disorder (BED) is the most common eating disorder among obese patients and can sometimes be associated with bulimia or attention deficit hyperactivity disorder[1,9].

This review will focus on metabolic comorbidities, the most important ones in childhood for prognosis and quality of life in adulthood. Excess weight, hypertension, dyslipidemia, systemic inflammation, insulin resistance, increased coagulability and oxidative stress are the main factors contributing to metabolic and cardiovascular risk in obese children[10]. This cellular metabolic dysregulation represents the complex pattern underlying metabolic syndrome (MS). Unlike adults, the diagnosis of MS in children and adolescents is not yet unequivocal: most assessments have been based on adaptations based on adult criteria. The current prevalence of MS in childhood and adolescence is 4.5-8.4%[11]. The recent article by Zong et al. proposed univocal diagnostic criteria for metabolic syndrome in pediatric age [12].

A key factor appears to be excessive central obesity, with visceral adipocytes dysfunction, that causes macrophage migration into adipose tissue, altered adipokine secretion, oxidative stress, endoplasmic reticulum stress and proinflammatory cytokine secretion that fuels meta-inflammation [13,14]. Obesity is associated with elevated levels of free fatty acids and adipokines, including leptin, resistin, interleukin-6 (IL-6), and tumor necrosis factor- α (TNF- α), which increase hepatic glucose production and reducing glucose uptake by skeletal muscle. This is accompanied by reduced expression of adiponectin, which has an insulin-sensitizing effect in muscle and liver[14]. In peripheral tissues, elevated levels of free fatty acids and triglycerides impair mitochondrial function and increase the degree of oxidative stress, with the overall effect of reducing the ability of insulin to stimulate glucose transporters to the cell surface. Insulin resistance leads to an increased need for insulin production, and glucose levels rise as resistance exceeds the capacity of pancreatic beta cells to release adequate amounts of insulin, contributing to the onset of T2DM. Further downstream effects include hypertension and reduced HDL cholesterol levels, both of which contribute to an

increased risk of cardiovascular disease[11]. In the pathogenesis of T2DM associated with childhood obesity, once again the proposed model is the interaction between a predisposing genetics and an exposome, i.e. the set of factors to which the individual is exposed, which constitutes the trigger environment [13,14]. Through GWAS, the Progress in Diabetes Genetics in Youth (ProDiGY) consortium identified seven genomically significant loci, including rs7903146 in TCF7L2, rs72982988 near MC4R, rs200893788 in CDC123, rs2237892 in KCNQ1, rs937589119 in IGF2BP2, rs113748381 in SLC16A11, and rs2604566 in CPEB2. In the population studied, the trigger environment was found to be a poor-quality hypercaloric diet combined with reduced physical activity[15]. The clinical course of pediatric-onset T2DM differs significantly from that of adult-onset disease. In a study comparing glucose tolerance in 34 obese adolescents to 17 adults matched for BMI, sex, and ethnicity, youth had nearly 50% lower insulin sensitivity and twice the insulin levels. This may explain the worse glycemic control in adolescents with T2DM compared to adults, due to the greater β -cell burden, which leads to premature β -cell failure and worsening glycemic status and thus an accelerated progression to micro- and macrovascular complications compared to adults[14].

Metabolic Dysfunction-Associated Steatotic Liver Disease (MASLD), formerly known as nonalcoholic fatty liver disease (NAFLD), has become the most common chronic liver disease in children and adults in industrialized countries, coinciding with the obesity epidemic. MASLD encompasses a spectrum from simple steatosis to steatohepatitis with varying degrees of fibrosis, called metabolic-associated steatohepatitis (MASH), to liver cirrhosis. The majority of individuals with MASLD do not progress to advanced stages of MASH with fibrosis, but given its high prevalence, MASLD is becoming a major reason for liver transplantation in adulthood. Furthermore, advanced MASLD increases the incidence of non-hepatic complications, such as type 2 diabetes and cardiovascular risk. Currently, there are no non-invasive gold standard methods for the diagnosis and monitoring of MASLD, and there are no approved targeted treatments for MASLD. The estimated prevalence of MASLD in obese children is 34.2%. The exact prevalence of MASH or more advanced stages is unknown, as this diagnosis requires liver biopsy, which is rarely performed in pediatric age for staging alone: the prevalence of liver fibrosis is estimated to be around 16% of cases and cirrhosis in 0-1% of cases. From a genetic point of view, multiple SNP variants have been identified that influence the risk of MASLD, primarily by altering intrahepatic lipid metabolism. The I148M allele in the PNPLA3 gene rs738409[G] is the genetic variant most strongly associated with hepatic fat and inflammation. Obesity-related metabolic disorders, including insulin resistance, type 2 diabetes, dyslipidemia, and hypertension, are risk factors for advanced MASLD. Insulin resistance and type 2 diabetes are the major risk factors for its onset, as insulin resistance plays a key role in the pathophysiology of MASLD: the prevalence of prediabetes and type 2 diabetes is double in children with obesity and MASLD compared to obesity alone, and the development of type 2 diabetes is strongly associated with MASLD progression. The natural course of pediatric MASLD remains incomplete due to limited long-term follow-up studies. Follow-up data from clinical trials have shown that one-third of children with baseline steatosis or borderline MASH experienced histological progression to definite MASH and/or advanced fibrosis within 2 years. A recent 10-year follow-up study of 51 adolescents showed progression of fibrosis in 16% of cases and advanced fibrosis in 6%. Although cirrhosis, liver failure, and hepatocellular carcinoma (HCC) are rare in pediatric age, cases have been described. In adults, approximately 5% of patients with MASLD progress to cirrhosis, with 1-2% dying from liver-related causes. Extrahepatic complications associated with MASLD in pediatric age include type 2 diabetes, dyslipidemia, hypertension, renal impairment, and polycystic ovary syndrome, all of which are considered risk factors and consequences of MASLD. It has been shown that pediatric-onset MASLD is independently associated with an increased risk of mortality from cancer, liver disease, and cardiometabolic diseases in adults: adults with MASLD have a mortality risk 1.3 times higher than the general population, mainly due to cardiovascular causes[16].

Artificial Intelligence

Artificial intelligence (AI) is the ability of a computer to perform operations associated with human intelligence, through a model or algorithm. AI devices are capable of learning, which is taught by the operator through training: the training makes the device capable of responding to the input with an adequate output, to achieve a pre-established purpose[17].

AI, due to its potential and particularly the amount of data it can store and process, represents a significant support in the healthcare field: it can assist healthcare professionals in a range of tasks, including compiling individual patient clinical data reviews, risk stratification of individual patients based on clinical characteristics, processing radiological or histological clinical images, and predicting a patient's clinical response to treatment. AI potentially enables precision medicine: personalized treatment for each patient will take into account genomic variants, age, gender, geographic location, ethnicity, cellular metabolic profile, microbiome, and ecological susceptibility to medical treatments. This involves the acquisition of patient data, such as genetic information, data from continuous telemetry monitoring devices, or data from electronic medical records. In this way, each patient will receive personalized, tailored treatment, reducing adverse effects and optimizing benefits [17,18]. However, before AI systems can be used in healthcare, they need to be trained with health data generated by clinical activities (screening, diagnosis, follow-up visits), so that the AI device can learn the characteristics of the subjects, do associations, classify them, identify patterns and disease indicators, and make predictions of various outcomes. The challenge lies in integrating widely varying data into a single language: epidemiological data, medical history, recordings from electronic devices (such as telemetry monitoring devices), physical examinations, radiological and/or histological images, laboratory data, genetics, epigenetics, and ultrasound data. All this information is expressed in quantitative (vital parameters, laboratory data) and qualitative (radiological reports) data, and therefore, in the eyes of the machine, it constitutes a heterogeneous code that is difficult to standardize. Various techniques, increasingly complex over time, have been developed to analyze and integrate this data, enabling the development of diagnostic and predictive algorithms applicable to healthcare[18].

Through a branch of AI, called Machine Learning (ML) it is possible to train the machine to learn and analyze data and perform a specific task, to solve a desired problem[17]. The machine is trained by entering a large amount of data, and the higher the amount, the more efficient the machine's training. ML can be used to analyze structured data such as clinical, genetic, and biochemical data. In healthcare, machine learning models attempt to group patient characteristics into clusters or infer the likelihood of disease outcomes based on provided indicators (Figure 1). However, it is not possible to analyze unstructured data, such as radiological or histological images, and it is not possible with classical machine learning to integrate heterogeneous data into complex models from which to develop outcomes and predictive models. Therefore, classical machine learning (ML) has evolved into deep learning (DL), a more complex technique capable of integrating a significant amount of data and processing results from complex data. A DL technique that overcomes the limitations of machine learning on unstructured data is natural language processing (NLP)[18]. With NLP, you can automate question responses, perform text analysis and synthesis, as well as evaluate and extract information using unstructured sources[17]. NLP procedures aim to transform texts (such as medical records, radiological and histological reports) into structured machine-readable data, which can then be analyzed using ML techniques [18].

Machine Learning

Machine learning (ML) is the ability of a machine to provide an appropriate outcome in response to input and to learn automatically, using data to improve its performance based on experience[17]. ML builds data analysis algorithms to extract features from data: in healthcare, patient characteristics represent the input, that is the starting point from which to extract the result. Inputs typically include basic data, such as age, gender, and past medical history, as well as disease-specific data, such as

radiological tests, biochemical values, genetic variants, physical examination data, clinical symptoms, medication use, and so on[18].

ML algorithms can be divided into three main categories: unsupervised learning, supervised learning, and reinforcement learning. More recently, semi-supervised learning has been proposed as a hybrid between unsupervised and supervised learning[17,18].

In unsupervised learning, a computer is trained with a nonspecific data set and independently predicts the results. In this case, models are developed on a data set that has not been categorized or labeled, and the machine acts on it autonomously. The primary goal of this learning is to classify or group the unsorted data set based on commonalities, patterns, and variances. The machine must be able to identify hidden patterns in the input data set[17]. Clustering and principal component analysis (PCA) are two examples of unsupervised learning. Clustering groups subjects with similar traits into clusters, generating labels without using outcome information. PCA is primarily used to handle large numbers of repeated inputs, such as the number of genes in a genome-wide association study. PCA synthesizes a large amount of data into principal components without losing too much information about the individual inputs. It is possible to combine the use of PCA to reduce the data dimensionality and then use clustering to group the items[18].

In supervised ML, a specific, labeled dataset is used to train the machine, which, after training, is able to predict an outcome based on the learned model. For good efficiency, the input must be correct and the dataset must be correctly mapped by the machine[17]. The result is included in the training to improve the model's efficiency. Compared to unsupervised ML, it provides more reliable results, which is why supervised ML is preferred in healthcare. Unsupervised ML however can be used as part of the pre-processing phase, to reduce dimensionality or identify subgroups, which in turn makes the subsequent supervised learning phase more efficient[18]. Supervised ML allows, for example, to estimate based on patient characteristics the probability of contracting a particular clinical event, the value of a disease marker, or the expected survival time. Supervised ML techniques include linear regression, logistic regression, naïve Bayes, decision trees, K-nearest neighbors, random forest, linear discriminant analysis, support vector machine (SVM), and artificial neural network (ANN)[17]. SVM is one of the most widely used techniques and is primarily used to classify subjects into two groups. The training objective is to find the optimal discrimination so that the resulting classifications are as close as possible to the results, i.e., with minimal classification error (e.g., benign vs. non-benign lesion)[18]. An ANN is a model inspired by the human nervous system, consisting of artificial neurons or nodes (chips) that establish a complex interaction between each other, forming a network, connected by a system of connection parameters ("weights") that allows the passage of signals between one chip and another. An ANN allows for the capture of complex nonlinear relationships between input variables and an outcome[17]. An ANN can generally be divided into three layers of neurons: input (which collects inputs), hidden (which discovers patterns and handles most of the internal computations), and output (which creates and displays the final network outputs). Computation is fueled by connections between neurons, and each neuron has an output layer, a transfer function, and a weighted input. The weighted sum of the inputs a neuron receives activates it, and the activation signal then travels through a transfer function to generate an output. The neural network's architecture, learning rules, and transfer functions influence the behavior of the ANN. Several factors can influence the performance of a neural network: connection parameters, bias, learning speed, and the size of the processed subsets[17].

Therefore, in reinforcement learning, an artificial intelligence agent, called a software component, automatically executes its actions, learning from experience and improving performance. Reinforcement learning operates with a feedback-based approach, with a reward for positive feedback and a penalty for negative feedback. The goal of the software component is to obtain the maximum reward and the minimum penalty[17].

Deep Learning

While machine learning requires the acquisition of elementary data to provide results, according to the training provided by the data manager, deep learning (DL) represents a step beyond this "elementary" function. In DL, the machine is fed raw data and develops its own pattern recognition scheme. The model the machine uses to identify the various patterns is developed using a series of simple parallel operations as a basis, integrated into sets of increasingly complex operations. The most widely used DL models are based on supervised and reinforced ML models. The input data can be heterogeneous, a feature of enormous advantage in the healthcare sector: for example, results from laboratory tests, radiological examinations, disease indicators, histological images, and so on can be integrated. Furthermore, deep learning models can work on large data sets, integrating the results obtained from the processing models into a single final result. This feature represents an evolution compared to traditional DL models and allows for the identification of clinical and therapeutic outcomes and predictive models in the healthcare sector[19]. Literature has recently highlighted the potential of DL in complex diagnostic areas, such as dermatology, radiology, ophthalmology, and pathology. An example is the development of the field of artificial vision - computer vision (CV), which focuses on understanding images and videos and performs feature and pattern recognition and segmentation and classification, useful for determining the presence of a certain item (e.g., a tumor lesion) in a radiological or histological image[18,19].

DL has provided an evolution of ANN: convolutional neural networks (CNN). This DL algorithm is a specialized subcategory of ANN capable of extracting features from unstructured data such as images (in the healthcare sector, radiological images, cytological or histological). CNNs learn to classify objects in an image, enhancing clinicians' ability to identify features of interest (e.g., fibrosis, calcifications, tumor lesions). CNNs have demonstrated excellent performance in transfer learning, where a CNN initially trained on a huge dataset, unrelated to the desired outcome, is further refined on a much smaller dataset related to the task of interest (e.g., medical images). In the training phase, the algorithm leverages large amounts of data to learn statistically relevant, non-specific findings in the proposed images; in the second phase, higher levels of the algorithm are retrained to distinguish cases of medical interest. Additionally, object detection and segmentation algorithms identify specific parts of an image that correspond to a particular outcome (e.g. a particular cytoarchitecture). In this way CNN algorithms take the image data as input and process them until the original raw data matrix is transformed into a probability distribution[19].

Another new DL tool is natural language processing (NLP). Images, clinical reports, and genetic data must be machine-readable so that ML algorithms can be run directly after appropriate preprocessing or quality control. However, much of the clinical information is in the form of narrative text, which is unstructured and incomprehensible to computer programs. In this context, NLP aims to extract useful information from narrative text and classify it to support clinical decision-making. Through DL algorithms such as using recurrent neural networks (RNN), NLP can perform text analysis, understand spoken or transcribed language, analyze the temporal sequence of events, and generate image captions or text consistent with input (similar to the ChatGpt system). In the healthcare sector, the data processed by NLP increasingly comes from electronic health records (EHRs), veritable databases containing millions of data on millions of patients. The raw data for each patient must be standardized and standardized to create a comprehensive, analyzable system. In this way, the algorithm through the ML component it can manage structured data (quantitative, linear and continuous values) and through the NLP component it can extract unstructured texts and in turn these techniques can be integrated into systems such as CNNs to formulate diagnoses and predictive models [18,19].

One field where ML and DL are significantly contributing to scientific progress is genomics. GWAS analysis requires algorithms for very large patient cohorts and that manage confounding factors, characteristics where DL algorithms have proven effective. Potentially, algorithms could integrate complex phenotypic data, such as clinical, radiological, or histological images, or continuous remote monitoring data, such as that from sensors, into genotypic variants. In other

words, the study of genetic or epigenetic variants could benefit from DL techniques for phenotype prediction. For example, a challenge for clinical geneticists is determining whether a variant is clinically relevant. In this regard, DL-based predictive models could support this classification by predicting the pathogenicity of mutations. Genomic data can also directly serve as a biomarker for the onset and progression of a disease: for example, data extracted from circulating cell-free DNA may contain biomarkers useful in the early diagnosis of chronic diseases or tumors; or prenatal diagnosis can be performed using specific biomarkers from fetal DNA. DL systems can improve the quality of biomarker testing by providing predictive models based on the presence of SNPs, methylation profiles, and other epigenetic modifications[19].

2. Materials and Methods

This review was conducted using the following databases: PubMed, NIH-National Library of Medicine and Google Scholar. The keywords used were: "obesity," "globesity," "children," "comorbidities," "healthcare," "socioeconomic status," "MASLD," "metabolic comorbidities," "diabetes," "artificial intelligence," "machine learning", "deep learning," and "multiomics." The inclusion criteria used were: articles in English published since 2004; full-length articles, narrative reviews, systematic reviews, scoping reviews, consensus, guidelines, meta-analyses, editorials. 91 papers were selected, while 17 papers were excluded since they were beyond the scope of this review.

3. Results

Several scientific works in the literature have demonstrated the usefulness of AI in various fields of healthcare and in recent years AI models have been applied to different aspects of childhood obesity. In particular AI techniques have been used for the assessment of nutritional status, early diagnosis of complications, the therapeutic response to drugs and bariatric surgery, for the formulation of personalized lifestyle modification treatments[20]. This results in a significant contribution of AI to precision medicine, which will be able to provide personalized treatments to patients, based on their risk category and will be able to direct pharmacological treatment to those patients who are most likely to benefit from it. The goal of precision medicine in childhood obesity is to identify a targeted treatment based on a precise diagnosis based on the individual patient's specific characteristics: genetic and epigenetic factors, environmental factors such as diet, environmental pollutants, a sedentary lifestyle, smoking, insomnia (which constitute the exposome, i.e., the set of factors to which the individual is exposed), and metabolic and meta-inflammatory status. This process of integrating the patient's global characteristics is called "deep phenotyping" and allows for the creation of a risk profile for each patient and the individualization of treatment strategies[13].

AI-based models can integrate data from electronic health records (EHRs), growth curves, family lifestyle data, including dietary habits, physical activity and social and environmental determinants[21]. For example, computer vision applied to satellite imagery allows for the quantification of green spaces, physical activity areas, bike paths, and the density of commercial establishments selling junk food, which influence the behavior of certain groups of individuals[22]. By integrating these heterogeneous data sources, AI algorithms can build high-resolution maps of obesogenic environments, identifying social, geographic, and economic risk areas, to support targeted intervention initiatives for obesity prevention. Advanced machine learning algorithms, such as RNNs, can capture complex nonlinear relationships between the various factors that contribute to the onset of complex pathologies, allowing a better understanding of the mechanisms underlying childhood obesity. In this regard, a study has identified some structural neurological characteristics and biological mediators typical of obesity: using ML techniques, obese individuals were identified from overweight individuals, using independent fecal metabolites and neuroimaging data, incorporated into a neural network model reconstructed from structural and functional magnetic resonance imaging (ANN) studies. By optimizing the data using SVM, the model also achieved an accuracy of 90.25% in discriminating obese from severely obese patients. It was also possible to

highlight structural differences in obese individuals in brain regions involved in reward circuits and default network [23], already examined in another preclinical study that demonstrated an increase in activity in overweight mice compared to lean mice[24]. Microbiome data were also analyzed in the obese cohort compared to the non-obese cohort: several amino acid-derived metabolites emerged as important in differentiating obese individuals [23]. A preclinical study has already demonstrated that mice transplanted with the microbiomes of human twin pairs, with different BMIs, showed differences in body composition and a microbiota in obese mice with a predominance of branched-chain amino acids [25]. Among amino acid derivatives, agmatine has demonstrated a neuromodulatory role, thanks to its activity on serotonergic and glutaminergic neurotransmitters. It is released as a compensatory and protective mechanism in response to stress and intestinal inflammation and can interact with vagal afferents, constituting the crosstalk of the brain-gut axis in obese and overweight patients[23].

AI has also proven effective in diagnosing psychopathological comorbidities of obesity: food addiction poses significant challenges to obesity treatment, being often associated with higher BMI, a higher rate of treatment dropout, and psychological distress related to guilt. A predictive nomogram using Naive Bayes (ML) was tested to screen for food addiction, based on the Yale Food Addiction Scale, a widely validated questionnaire for food addiction screening[23]. This tool, used on an outpatient basis, can identify patients requiring more targeted psychological intervention, reducing dropouts and maximizing therapeutic outcomes.

AI algorithms can predict the risk of obesity in normal-weight or overweight patients, thus enabling primary prevention[26] and some advanced models can classify patients as low, medium or high risk for certain outcomes[27]. Various studies have been conducted using ML and computer programming techniques to predict the onset of obesity in cohorts of overweight patients, to predict the onset of complications in obese patients, or to establish the role of biomarkers in the onset of obesity or the role of environmental factors in individuals with a certain genetic background[28,29]. Various ML and DL techniques have been compared and the models with the highest diagnostic accuracy are based on ANN, KNN, SVM, decision trees and random forests: these models have been used on large datasets that include environmental, clinical and behavioral characteristics and they are able to predict the risk of obesity and the onset of complications in high-risk groups[19,28]. The study by Wang and colleagues used genetic variants identified by next-generation sequencing (NGS) to predict obesity risk using machine learning models such as SVM, achieving good efficacy (approximately 71% accuracy in risk classification)[30]. Advanced AI methods (CNN, RNN) can integrate genomic variants, epigenomics, and environmental characteristics to generate a risk classification, enabling the identification of high-risk subgroups. A ML model was proposed that captures the complex multifactorial interactions underlying obesity, using genome, epigenome, diet, and lifestyle data. The model achieved an accuracy of approximately 70%, identifying 21 SNPs, 230 DNA methylation sites (e.g., CPT1A, ABCG1, SLC7A11, RNF145, SREBF1), and 26 dietary components, such as processed meat, carbonated beverages, high-fat dairy products, and flavonoid intake, as risk factors. This study demonstrates how multi-omics integration with artificial intelligence can reveal patterns of interaction between biology and nutrition and enable obesity prevention through nutrition, from a nutrigenomics perspective[26]. Several complex predictive models, based on ML and DL techniques, have been created with the help of AI. In 2015, Dugan and colleagues used ML models to build a clinical dataset for a pediatric clinical decision support system called CHICA (Child Health Improvement via Computer Automation). This system has proven useful for identifying risk factors for obesity, such as early weight gain between one and two years of age, and protective factors, such as Caucasian ethnicity[31]. In another study, Allen and colleagues investigated the interaction effect of an obesogenic environment on the development of obesity in adolescents, using 120 individual characteristics as predictors of each child's waist-to-height ratio. The analysis revealed that children from less educated, single-parent families, or from more disadvantaged and poorer areas had higher waist-to-height ratio z-scores. Furthermore, among these

children, those who played less than 23 minutes of sports per week showed higher waist-to-height ratio z-scores[32].

Another study used ML to predict obesity risk, building a model capable of planning meals to encourage changes in eating habits[33]. AI methods such as random forests and logistic regression have also proven useful for identifying predictive factors for complications associated with metabolic syndrome, such as type 2 diabetes, cardiovascular disease, and cancer. Thanks to ML, new predictive indicators have been identified, in addition to traditional diagnostic criteria, such as resting heart rate, plasma C-reactive protein concentration, sex, age, lipidomics data, and SNPs[34]. Another application of AI in the treatment of childhood obesity is provided by predictive models of therapeutic response: it is possible to predict protein-protein interactions and therefore the efficacy of drugs through ML techniques, allowing to overcome phase II clinical trials with a rapid and economical method[35].

The use of technologies based on AI, such as smart movement games or smart devices with applications that suggest healthy meals and physical exercise, generating advice based on data collected from the patient, could significantly contribute to the therapeutic management of obesity in children and adolescents[36]. According to the 2023 American Academy of Pediatrics (AAP) guidelines, a targeted lifestyle intervention for at least 26 hours over a 3-12 month period in obese children can lead to a reduction in BMI. This intervention should be multimodal, including dietary changes, increased physical activity, and lifestyle improvements, to achieve maximum results in the prevention and treatment of childhood obesity[37]. AI has enabled the creation of digital supports for the treatment of childhood obesity, such as applications connected to smart devices (smartwatches) or gaming platforms or digital coaches that provide personalized recommendations for diet, physical activity and sleep hygiene[38]. Data can be shared with family or case managers to monitor the progress of lifestyle changes. Reinforcement learning algorithms dynamically adapt recommendations based on feedback from the user's behavior (e.g., a food diary), promoting motivation and adherence. Integration with wearable sensors and monitoring devices supports continuous assessment of physical activity and energy expenditure. This form of monitoring an obese child's habits, both in terms of eating habits and physical activity, can improve both short- and long-term weight management outcomes. Several platforms have been developed with the help of AI to support obesity treatment, particularly lifestyle modification, which remains the primary treatment strategy today, but is burdened by a high rate of failure or relapse after initial therapeutic success. Examples include large language models (LLMs), such as ChatDiet, ChatGPT meal planner, CHARLIE, GPT4 exercise planner, and Paola; SlimME: these models generate interactive responses, providing empathetic support similar to a human life coach; however, these models are limited by the risk of providing false information and failing to interpret input other than text. Another example is provided by so-called Digital Treatments (DTx), such as Omada health, Liva healthcare, and multimodal intervention apps: software used for therapeutic purposes, created for long-term behavioral interventions, limited by the subject's digital capabilities and, above all, by the individual's motivation to use the software on a daily basis. Several articles regarding DTx, the results of which have been summarized in the review by Lee et al., have been published to demonstrate the effectiveness of these tools in making obese patients lose weight[19]. Another useful resource is "active video games" or "exergames," movement games that integrate movement into children's daily lives, promoting healthier lifestyle choices in an engaging way. Several reviews of game-based interventions have demonstrated their effectiveness in improving nutritional knowledge, eating habits, and body composition, and in reducing body weight in overweight children[36]. The recent development of the so-called "Internet of Things" (IoT), a network of smart devices that collect and exchange information from sensors, has made it possible to increase the monitoring of the eating habits and lifestyle of overweight and obese patients[39]. These devices are able to analyze food and physical activity choices and generate personalized recommendations[40]. An example of a model that uses the IoT in the monitoring and treatment of childhood obesity is represented by the ETIOBE system, an intelligent platform that integrates data from monitoring sensors (electronic vest to

monitor physical activity, pulse oximeter for sleep monitoring, a blood pressure monitoring belt), data entered by the patient regarding diet, directives entered by the clinician[41]. This represents one of the most valid approaches for the treatment of obesity, also allowing the clinician to remotely monitor therapeutic success.

3.1 Algorithms for the Diagnosis and Management of T2DM and MASLD

Various AI techniques have been applied for the diagnosis and risk stratification of complications in obese patients. In particular, this section focuses on the use of AI for the early diagnosis and the formulation of predictive models for the risk of T2DM and MASLD, given the high prevalence in obese children and adolescents, which justifies the extensive study by several research groups in the last decade.

Validated applications of ML and DL in the diagnosis and management of T2DM are diverse: through decision trees, random forests, or neural networks, it is possible to predict the risk of diabetes in obese patients by analyzing lifestyle, clinical and psychological factors, and physical and social habits; to build predictive prognosis models and personalize treatment in diabetic patients; to predict long-term complications; for example, through DL (CNN) it is possible to perform automated retinal screening to identify early signs of diabetic retinopathy. Furthermore, it is possible to phenotype the patient through genomic and epigenetic studies, achieve autonomous patient management through monitoring devices that use AI software, sensors and telemedicine[42].

Several studies have identified the risk factors most associated with the onset of T2DM in obese patients: they are summarized in the Table 1 below.

Table 1. ML and DL techniques used to create risk prediction models of DMT2 and its complications, investigate microscopic alterations and personalize treatment in obese patients.

Articles	AI techniques ¹	Results
Ellahham et al, 2020	SVM, ANN, NB, DT, RF, classification and regression trees, KNN	stratify the risk of diabetes and identify patients with diabetes and controls
Sun et al, 2025	LR	Predict DMT2 risk from BMI, dietary habits, blood pressure
Nomura et al., 2021	LR, RF, GB	prediction of new-onset diabetes
Cardozo et al., 2022	KNN, SVM, NB, RF, ANN	DMT2 risk score based on laboratory tests
Yang et al., 2025	SVM	Early diagnosis of DMT2 in obese children, based on BMI, creatinine, prealbumin, glucose (180 min), glycosylated hemoglobin A1c, thyrotropin, total thyroxine (T4), and free T4 concentrations
Li et al., 2025	RF showed best accuracy among other techniques	Early prediction of diabetic nephropathy in DMT2 patients
Anjana et al., 2020	K-means clustering	Clusterization and clinical phenotyping of patients with DMT2

Xie et al., 2025	XGBoost	Investigating pancreatic cell dysregulation in DMT2 patients
Zeevi et al., 2025	GB	Predicting glycemic response; personalized nutrition to control glycemic response

Using a machine learning algorithm, the study by Sun et al. analyzes dietary patterns and identifies, among various diets, the one associated with obesity and the highest risk of T2DM, based on sugary and processed foods[43]. The review by Nomura et al. reports articles demonstrating that using models based on gradient boosting, random forests and logistic regression it is possible to predict the onset of T2DM up to 5 years before the onset[44]. Techniques such as ANN applied to a database of biochemical data have been found to be effective in identifying patients with prediabetes and T2DM[45]. Models using decision trees and random forests have proven effective in terms of specificity and accuracy in associating risk factors such as high total cholesterol, LDL cholesterol, and triglycerides with the onset of T2DM, as well as less considered risk factors such as less than 6 hours of sleep a night, high sodium intake, psychological stress, exposure to environmental pollutants, and living in urbanized areas[46]. Furthermore, through ML it was possible to study and identify biomarkers of beta cell dysregulation in patients with T2DM, such as the depletion of mature insulin-producing cells, the expansion of immature cells and cells produced in response to endoplasmic reticulum stress, modification of acinar cells towards an inflammatory pattern and of ductal cells towards a secretory pattern[47]. A new study by Yang et al. produced a machine learning model to predict the onset of T2DM in obese children: 292 children with obesity and T2DM were enrolled and their characteristics were studied. Eight machine learning models were compared for their ability to identify clinical and biochemical characteristics for the creation of predictive risk models. The SVM was the best model for predicting the onset of T2DM in obese children and identified eight characteristics on which the predictive model was based: BMI, creatinine, prealbumin, thyrotropin, total thyroxine, free thyroxine, glycosylated hemoglobin, and blood glucose 180 minutes after an oral glucose load[48].

ML is effective in producing predictive models of risk of developing complications in obese patients with T2DM[49]. Already in 2018, Ahlqvist and colleagues had applied k-means and hierarchical clustering to reclassify adults with new-onset diabetes into subgroups based on clinical and biochemical characteristics, comparing the groups using logistic regression and Cox regression methods to compare the risk of complications across groups. Among the subgroups, individuals with insulinopenia showed a higher risk of diabetic retinopathy, while individuals with severe insulin resistance had a higher risk of diabetic nephropathy[50]. Another k-means-based model was used in a cohort of 19,084 individuals with T2DM, classified into 4 groups based on clinical variables (age at diagnosis, BMI, glycosylated hemoglobin, HDL cholesterol, C-peptide, waist circumference), identifying a group with early-onset T2DM, with poor glycemic control and high risk of nephropathy and retinopathy[51].

ML has been widely applied in the field of precision nutrition to tailor a personalized diet to the individual, aimed at preventing or managing diet-related diseases[52,53]. An example is provided by the study by Zeevi et al., who developed a personalized diet to predict glycemic response in 800 healthy and prediabetic adults, considering biochemical and anthropometric data, diet, physical activity, and gut microbiota data in an integrated approach. The researchers adopted a gradient-boosting regression (ML) algorithm that accurately predicted postprandial glycemic responses to meals: a randomized, blinded, controlled intervention based on an algorithm-predicted diet led to significantly lower postprandial glycemic values and changes in gut microbiota composition[54]. Similar results were obtained in two other similar studies on a population of American adults[55,56]. In the therapeutic management of T2DM, a platform (AdvisorPro) has been created that integrates data from continuous monitoring sensors with self-monitoring of blood glucose levels to suggest

insulin dose adjustments, with results no lower than those proposed by the clinician. Another software program (Guardian) allows for the prediction of hypoglycemia in diabetic patients up to 30 minutes in advance, allowing for better glycemic control and fewer adverse events[44].

AI has significantly contributed to many aspects of MASLD management and has contributed to new understanding of the complex pathophysiology of this disease. Table 2 summarizes the contribution of AI in this field.

Table 2. ML and DL techniques used to create risk prediction models of MASLD and its complications.

Articles	AI techniques ¹	Results
Lou et al., 2025	CNN, AI- based texture analysis	MASLD diagnosis and staging from ultrasound, CT and histology
Yang et al., 2026	DT, RF, SVM, XGBoost, neural networks	Identifying MASLD-related genes and lipidomic biomarkers, non-invasive screening technologies and predicting the risk of disease progression
Ji et al., 2022	RF, multinomial logistic regression analyses, recursive partitioning and regression tree algorithm	Integrating metabolomic and transcriptomic data to find new biomarkers of MASLD; early diagnosis of MASH
Noureddin et al., 2022	LR	Predicting MASLD risk in population based on male sex, hemoglobin A1c, age, and body mass index
Razmpour et al., 2023	RF showed best accuracy among other techniques	MASLD screening and early diagnosis from anthropometric data
Huang et al., 2023	LR	Prediction on MASLD 5 years before outset
Qin et al., 2023	SVM	MASLD early diagnosis from physical examination and laboratory tests
Sorino et al., 2020	SVM	MASLD early diagnosis
Ryou et al., 2021	CNN	MASLD diagnosis from ultrasound
Das et al., 2021	ML model comprising SVM, Neural Net and XGBoost algorithms	MASLD diagnosis from ultrasound
Graffy et al., 2019	CNN	Automated CT-based liver fat quantification tool
Vanderbeck et al., 2021	SVM	Identifying histological features of steatosis, ballooning, inflammation, fibrosis, etc.
Heinemann et al., 2019	CNN	Early diagnosis of MASH
Zhang et al., 2025	XGBoost showed best accuracy among other techniques	Early diagnosis of MASLD

Li et al., 2026	LR	Predicting MASLD using body roundness index (BRI) and the triglyceride-glucose (TyG) index
Zhou et al., 2025	RF, GB	MASLD early diagnosis based on clinical scores
Wang et al., 2025	LR, RF, DT, SVM, XGBoost, Light GBM	Development of a proteomic risk score (ProScore) to improve MASLD diagnostic accuracy
Li et al., 2025	RF, LASSO regression	Predictive score for MASLD based on clinical and biochemical features
Tavaglione et al., 2024	ANN	MASLD risk based on lipidome
Verma et al., 2024	RF	Early detection of fibrosis in MASLD patients
Gil-Rojas et al., 2024	XGBoost	Early diagnosis of hepatocellular carcinoma in MASLD
Shibata et al., 2024	CART	Cardiovascular risk in MASLD patients
Zhan et al., 2025	LR	non-invasive biomarkers for the early diagnosis of obesity-related MASLD
Lu et al., 2026	RF, LR, SVM, XGBoost, Adaptive Boosting	Predictive score of fibrosis based on lipidome
Zöggeler et al., 2025	XGBoost, RF	Identifying microbiota species in MASLD obese patients
Nychas et al., 2025	RF	Identifying highly specific microbiota signature in MASLD patients

Thanks to ML, advances have been made in screening, long-term prognosis, treatment, and monitoring of MASLD[57]. ML techniques can analyze the transcriptome and genes implicated in the pathophysiology of MASLD, identifying new genes and biomarkers associated with it, such as AXUD1, FOSB, GADD45B, and SOCS2. Metabolomics and lipidome can be studied, integrating data from a multi-omics perspective (in this field, metabolites such as glutamic acid, isocitric acid, and C4BPA have been identified as predictors of MASLD). It is also possible to study the role of exposure to environmental pollutants, such as pesticides, in the onset and progression of MASLD, or to conduct large-scale screening using EMR to identify the true incidence of this disease. ML also allows for risk stratification, prognosis prediction, and analysis of the role of comorbidities in the overall risk of the individual patient, with a view to precision medicine. It allows for treatment support by providing a personalized diet and a prediction of drug response[58]. Several studies have been published on the use of ML and DL in this sense: ML-based models, especially random forests and SVM, have been validated and proven effective in predicting the onset of MASLD based on clinical and/or demographic characteristics of obese patients[59–64]. DL-based models, such as CNN, have been shown to be useful in detecting hepatic steatosis on ultrasound images[65,66], in quantifying hepatic steatosis from computed tomography images[67], in identifying microscopic features typical of MASLD from histological images[68–71]. In the study by Heinemann et al., a CNN-based model created a predictive score for steatosis, inflammation, and fibrosis from biopsy images[72]. A model

using CNN has been shown to be able to recognize stages of fibrosis from elastography data. Furthermore, it is possible to integrate the data from radiomics (computed tomography, ultrasound, contrast-enhanced ultrasound, magnetic resonance imaging) with data derived from circulating DNA or biomarkers derived from metabolomics studies[73].

Several studies have reported ML-based models able to predict the onset of MASLD: in the study by Zhang et al. a predictive model based on an XGBoost algorithm was found to be effective in establishing the risk of MASLD using clinical and biochemical data as predictive factors, with particular predictive value of waist circumference[74]. In the work of Li et al., two combined indices (roundness index and triglyceride-glucose index) were identified as predictive factors of MASLD in obese patients through ML[75], while another working group realized a visceral adiposity index as a predictor of MASLD through random forests and XGBoost[76]. Another model integrated factors such as age and sex with a dietary inflammation score and the presence of diagnostic features of metabolic syndrome to estimate the risk of MASLD in obese individuals: this model suggests that consumption of high glycemic index foods, low intake of flavonoids and vitamin D, low perception of one's health status, poor psychological well-being and an increasing trend of liver enzymes are all predictors of the onset of MASLD[77]. Wang et al. study identified 15 circulating inflammatory proteins associated with MASLD in obese children, developing a proteomics-based risk score (ProScore) to increase the diagnostic accuracy of MASLD in overweight and obese children. In particular, a panel of 6 proteins (FGF21, CDCP1, CD244, OPG, FLT3L, MCP1) was found to be predictive, especially in children with low genetic risk, allowing further stratification of this group[78]. The group of Li et al. instead created a predictive nomogram in which BMI, waist circumference, fat mass and ALT were identified as risk factors positively associated with MASLD in obese children[79]. Another ML-based model highlighted BMI and triglycerides as independent predictors of steatosis and progression to MASH[80]. Regarding the prognosis of MASLD, ML demonstrated approximately 10% better performance than the FIB4 calculation method in discriminating liver fibrosis, using clinical, anthropometric and biochemical data[81]. Through ML it is possible to identify negative prognostic factors in patients with MASLD, including the risk of developing hepatocellular carcinoma[82]. A model called CART was also created to identify subjects with MASLD at high cardiovascular risk[83].

AI has also improved our understanding of the pathogenesis of MASLD in obese patients through the study of lipidomics, transcriptome, metabolomics, and the intestinal microbiota of obese patients with MASLD. The cellular alteration of liver cells with MASLD was studied using four ML algorithms[84]. Biomarkers were detected using ML algorithms, such as tryptophan derivatives, metabolites elevated in obese patients with MASLD, which induce steatosis and oxidative stress[85]. Another ML algorithm analyzes the predictive value of the lipidome: it identified diradylglycerol 34:1 and triradylglycerol 52:3 as biomarkers of fibrosis[86]. Some studies have analyzed the role of the microbiota in the pathogenesis of MASLD: alterations in the microbiota influence the development and severity of MASLD and MASH. AI has analyzed the microbiome from children with MASLD and MASH and many differences in species diversity was found. Using XGBoost-based models and random forests, it was possible to predict the onset of MASLD in obese children and MASH in children with MASLD, based on the characteristics of the microbiota[87]. A strain, *Prevotella Copri*, has been identified that worsens the progression of MASLD by down-regulating lipid metabolism genes, increasing the accumulation of triglycerides and cholesterol in the liver and down-regulating the expression of occludins, increasing intestinal permeability[88]. A specific microbiota typical of MASLD has been isolated through metagenomic studies. ML identified 12 bacterial species in obese patients with MASLD, while some strains are typically less present. *Eubacterium hallii* was less abundant in patients with MASLD, both obese and thin, suggesting a possible role as a probiotic. This strain is one of the producers of short-chain fatty acids, which have an anti-inflammatory effect, improve intestinal permeability and the growth of eutrophic bacterial flora, and regulate metabolism. Specific metabolites derived from the microbiota have also been

identified in patients with MASLD, which could participate in the brain-gut cross-talk and in the pathogenesis of the disease[89].

3.2. Figures

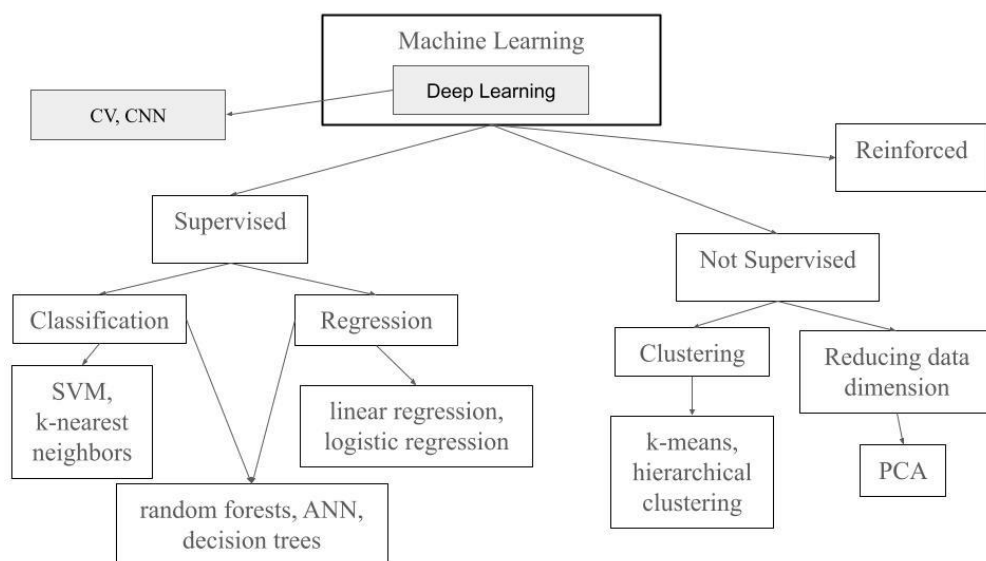


Figure 1. ML and DL algorithms applicable to the healthcare.

4. Discussion

Childhood obesity represents a therapeutic challenge for the clinician and the frequent drop out also makes early management of complications difficult: sometimes the child or adolescent is sent to the childhood obesity expert only when the complications are already present, when it is difficult to reverse metabolic effects. Childhood obesity is a complex and multifactorial disease and thanks to AI, through GWAS[15,30], studies about epigenome[5], metabolomics[84], transcriptome[78], microbiome[23,25,54], it has been possible to broaden the spectrum of factors predisposing to the onset of obesity in childhood. Thanks to ML, metabolites, intestinal bacteria, neurotransmitters and brain regions[23] involved in the onset of obesity have been identified, as well as SNPs in genes involved in lipid and carbohydrate metabolism, energy homeostasis and the hunger-satiety mechanism that predispose to an increase in BMI[3,4]. It is also possible to integrate ecological and demographic data[22,32], using ML and DL to estimate the risk of obesity in a given population, depending on the child's environment. Thus, it is possible to estimate the risk of developing complications for an individual based on multi-omics integration[20]. Among the most frequent complications, already present in pediatric age, are MS, prediabetes, DMT2, MASLD, all united by meta-inflammation and by insulin resistance[11]. It is known that the probability of developing these complications depends on the duration of obesity, however several studies have demonstrated a genetic and epigenetic susceptibility underlying this risk [26,28,29]. AI helps clinicians in the early diagnosis of such complications, as in the case of CV (DL technique) in retinal screening to identify diabetic retinopathy[42] or CNN (also a DL technique), which identifies steatosis and fibrosis from ultrasound images and elastography[72,73], in the latter case with a better performance than the FIB-4 score. Several scores and predictive risk models have been formulated using ML algorithms: among these, the most accurate (>90%) have proven to be models based on SVM and random forests[58]. Examples are the ProScore, based on proteomics, which, based on the presence of circulating inflammatory proteins, establishes a risk of developing MASLD in obese children[79]; or the model by Yang et al. based on SVM which, based on clinical and biochemical characteristics, identifies a risk of developing T2DM in obese children[48]. Another useful application of AI in the therapeutic management of obesity is represented by the ETIOBE platform, which integrates data entered by the patient on meals consumed and physical activity with data provided by motion sensors and vital

signs sensors, such as saturation and blood pressure, allowing for the active monitoring of lifestyle changes and the clinical conditions of the obese patient[41]. AI also allows for the formulation of personalized dietary plans based on the patient's characteristics, with a view to precision nutrition, opening up the possibility in the future of personalization based on deep phenotyping, i.e. the set of genetic, clinical and biochemical characteristics of the obese patient [21,26].

However, some controversial aspects of the use of AI remain to be highlighted. AI is possible thanks to the use of large databases. To enable data scientists and engineers to work, data collection, storage, integration, and maintenance are required, which entails significant costs. Furthermore, highly heterogeneous data must be validated by expert clinicians and standardized to be understandable by data scientists[26]. Furthermore, the use of AI raises fundamental ethical questions: AI systems require access to sensitive patient data, with risks of breaches and issues with informed consent; AI could widen social inequalities, both due to the gap in access to technology and the digital illiteracy affecting a large portion of the population; it is difficult to understand the reasons behind the results produced by AI, given the complexity of the technology; furthermore, the question of who is legally responsible for the errors made by AI algorithms remains unresolved. Another aspect to consider is that the effectiveness of software for monitoring and recommending targeted meals and exercise depends heavily on individual motivation. Without direct supervision, adherence is variable and significantly impacts the results. Finally, research is still in its early stages: large-scale, long-term clinical studies on diverse populations are lacking, as are economic evaluations that justify universal adoption [20,26,29].

5. Conclusions

ML and DL have made significant contributions to the diagnosis, risk stratification for complications, prognosis, and treatment of childhood obesity. By integrating multimodal data—clinical, biochemical, social, behavioral, ecological, genetic, epigenetic, radiological, and histological—it is possible to develop a unique profile for each patient. This phenotyping allows for the establishment of a personalized risk of progression to complications and therefore for targeted interventions tailored to the patient's clinical phenotype. This strategy represents an example of precision medicine. AI-based software can also optimize lifestyle interventions and offer intelligent monitoring of complications, also promoting greater patient awareness. However, several ethical and practical challenges remain: the accessibility and interpretation of these technologies, the management of large and complex databases, and the supervision of the results provided by expert personnel. AI represents the inevitable future of healthcare, but it is essential to outline a safety profile for large-scale application.

Funding: This research received no external funding

Data Availability Statement: All the sources are cited in the bibliography

Acknowledgments: nothing to declare

Conflicts of Interest: The authors declare no conflicts of interest.

Abbreviations

The following abbreviations are used in this manuscript:

T2DM	Type 2 Diabetes Mellitus
MASLD	Metabolic Dysfunction-Associated Steatotic Liver Disease
AI	Artificial Intelligence
ML	Machine Learning
SNPs	Single Nucleotide Polimorphisms
WHO	World Health Organization
BMI	Body Mass Index

COVID-19	Coronavirus Disease 2019
GWAS	Genome Wide Association Studies
PYY	Peptide YY
GLP-1	Glucagon-Like Peptide-1
OXM	Oxintomodulin
CCK	Cholecystokinin
GIP	Gastric Inhibitory Polypeptide
PP	Pancreatic Polypeptide
OSAS	Obstructive Sleep Apnea Syndrome
BED	Binge Eating Disorder
MS	Metabolic Syndrome
MASH	Metabolic-associated Steatohepatitis
DL	Deep Learning
NLP	Natural Language Processing
PCA	Principal Component Analysis
SVM	Support Vector Machine
ANN	Artificial Neural Network
CV	Computer Vision
CNN	Convolutional Neural Networks
RNN	Recurrent Neural Networks
EHRs	Electronic Health Records
NGS	Next Generation Sequencing
CHICA	Child Health Improvement via Computer Automation
AAP	American Academy of Pediatrics
LLMs	Large Language Models
DTx	Digital Treatments
IOT	Internet of Things
NB	Naïve Bayes
DT	Decision Trees
RF	Random Forests
KNN	K-Nearest Neighbor
LR	Logistic Regression
GB	Gradient Boosting
HDL	High Density Lipoprotein
CART	Classification and Regression Tree
LASSO	Least Absolute Shrinkage and Selection Operator
XGBoost	eXtreme Gradient Boosting

References

1. Valerio, G., Maffei, C., Saggese, G. et al. Diagnosis, treatment and prevention of pediatric obesity: consensus position statement of the Italian Society for Pediatric Endocrinology and Diabetology and the Italian Society of Pediatrics. *Ital J Pediatr*, 2018, 44, 88 .
2. Levels and trends in child malnutrition: UNICEF/WHO/World Bank Group joint child malnutrition estimates: key findings of the 2023 edition, 2023
3. Littleton SH, Berkowitz RI, Grant SFA. Genetic Determinants of Childhood Obesity. *Mol Diagn Ther*. 2020 Dec;24(6):653-663
4. Panera N, Mandato C, Crudele A, Bertrando S, Vajro P and Alisi A Genetics, epigenetics and transgenerational transmission of obesity in children. *Front. Endocrinol.*, 2022, 13:1006008
5. Alfano R, Robinson O, Handakas E, Nawrot TS, Vineis P, Plusquin M. Perspectives and challenges of epigenetic determinants of childhood obesity: A systematic review. *Obes Rev*. 2022 Jan;23 Suppl 1:e13389.
6. Koliaki, C., Liatis, S., Dalamaga, M. et al. The Implication of Gut Hormones in the Regulation of Energy Homeostasis and Their Role in the Pathophysiology of Obesity. *Curr Obes Rep* 9 2020, 255–271
7. Obradovic M, Sudar-Milovanovic E, Soskic S, Essack M, Arya S, Stewart AJ, Gojobori T and Isenovic ER Leptin and Obesity: Role and Clinical Implication. *Front. Endocrinol*. 2021 12:585887

8. Maqsood S, Ahmed F, Arshad MT, Ikram A, Abdullahi MA. Comparative Analysis of Food Addiction and Obesity: A Critical Review. *Food Sci Nutr*. 2025 Aug 15;13(8):e70799
9. Carlos Alberto Nogueira-de-Almeida, Virginia Resende Silva Weffort, Fábio da V. Ued, Ivan S. Ferraz, Andrea A. Contini, Edson Zangiacomini Martinez, Luiz A. Del Ciampo, What causes obesity in children and adolescents?, *Journal of Pediatrics*, Volume 100, Supplement 1, 2024, Pages S48-S56, ISSN 0021-7557
10. Kumar, Seema et al., Review of Childhood Obesity, *Mayo Clinic Proceedings*, 2017, Volume 92, Issue 2, 251 - 265
11. DeBoer MD. Assessing and Managing the Metabolic Syndrome in Children and Adolescents. *Nutrients*. 2019 Aug 2;11(8):1788.
12. Zong X, Bovet P and Xi B A Proposal to Unify the Definition of the Metabolic Syndrome in Children and Adolescents. *Front. Endocrinol*. 2022, 13:925976.
13. Subramanian, M., Wojtuszczyz, A., Favre, L. et al. Precision medicine in the era of artificial intelligence: implications in chronic disease management. *J Transl Med*, 2020, 18, 472
14. Salama M, Balagopal B, Fennoy I, Kumar S. Childhood Obesity, Diabetes. and Cardiovascular Disease Risk. *J Clin Endocrinol Metab*. 2023 Nov 17;108(12):3051-3066. doi: 10.1210/clinem/dgad361. Erratum in: *J Clin Endocrinol Metab*. 2024 Apr 19;109(5):e1422.
15. Srinivasan S, Chen L, Todd J, et al; ProDiGY Consortium. The first genome-wide association study for type 2 diabetes in youth: the Progress in Diabetes Genetics in Youth (ProDiGY) Consortium. *Diabetes*. 2021;70(4):996-1005
16. Stroes AR, Vos M, Benninga MA, Koot BGP. Pediatric MASLD: current understanding and practical approach. *Eur J Pediatr*. 2024 Nov 19;184(1):29.
17. Jayanti Mukherjee, Ramesh Sharma, Prasenjit Dutta & Biswanath Bhunia, Artificial intelligence in healthcare: a mastery, *Biotechnology and Genetic Engineering Reviews*, 2024, 40:3, 1659-1708
18. Jiang F, Jiang Y, Zhi H, et al. Artificial intelligence in healthcare: past, present and future. *Stroke and Vascular Neurology* 2017;2: e000101
19. Esteva A, Robicquet A, Ramsundar B, Kuleshov V, DePristo M, Chou K, Cui C, Corrado G, Thrun S, Dean J. A guide to deep learning in healthcare. *Nat Med*. 2019 Jan;25(1):24-29
20. Lee H, Hwang J, Yon DK, Rhee SY. Multimodal and Multidimensional Artificial Intelligence Technology in Obesity. *JOMES* 2025;34:394-404
21. Raphaeli, Orit; Singer, Pierre, Towards personalized nutritional treatment for malnutrition using machine learning-based screening tools, *Clinical Nutrition*, 2021, Volume 40, Issue 10, 5249 - 5251
22. K. Peng, Z. Peng, and R. Zhang, "Enhancing Neighborhood Food Availability for Safer and Healthier Urban Environments: A Cross-Sectional Investigation in Changsha, China," in *Designing Healthy Buildings and Communities: Shaping a Climate-Resilient Future*, eds A. Cheshmehzangi, J. Zuo, A. Sharifi, R. Zhang, A. Z. Bafarasat, and J. Zhao, Springer, 2025, 99-123.
23. Osadchiy, V., Bal, R., Mayer, E.A. et al. Machine learning model to predict obesity using gut metabolite and brain microstructure data. *Sci Rep*, 2023, 13, 5488
24. Tregellas, J. R. et al. Altered default network activity in obesity. *Obes. (Silver Spring)*, 2011, 19, 2316-2321
25. Ridaura, V. K. et al. Gut microbiota from twins discordant for obesity modulate metabolism in mice. *Science*, 2013, 341, 1241214.
26. Singer P, Robinson E, Raphaeli O. The future of artificial intelligence in clinical nutrition. *Curr Opin Clin Nutr Metab Care*. 2024 Mar 1;27(2):200-206
27. M. H. Wang, " Artificial Intelligence Across the Obesity Continuum: From Mechanistic Insights to Global Precision Prevention and Therapy," *Obesity* 34, no. 2, 2026, 294-316
28. Shabani Jafarabadi, G.; Busetto, L. Artificial Intelligence in Obesity Prevention. *Healthcare* 2025,13, 3262.
29. Azmi, S.; Kunnathodi, F.; Alotaibi, H.F.; Alhazzani, W.; Mustafa, M.; Ahmad, I.; Anvarbatcha, R.; Lytras, M.D.; Arafat, A.A. Harnessing Artificial Intelligence in Obesity Research and Management: A Comprehensive Review. *Diagnostics* 2025, 15, 396
30. Wang H.Y., Chang S.C., Lin W.Y., Chen C.H., Chiang S.H., Huang K.Y., Chu B.Y., Lu J.J., Lee T.Y. Machine Learning-Based Method for Obesity Risk Evaluation Using Single-Nucleotide Polymorphisms Derived from Next-Generation Sequencing. *J. Comput. Biol. A J. Comput. Mol. Cell Biol*. 2018;25:1347-1360.

31. Anand V., Biondich P.G., Liu G.C., Rosenman M.B., Downs S.M. Child health improvement through computer automation: The CHICA system. *Proc. Medinfo.* 2004;107:187–191.
32. Allen B., Lane M., Steeves E.A., Raynor H. Using Explainable Artificial Intelligence to Discover Interactions in an Ecological Model for Obesity. *Int. J. Environ. Res. Public Health.* 2022;19:9447
33. Kaur R., Kumar R., Gupta M. Predicting risk of obesity and meal planning to reduce the obese in adulthood using artificial intelligence. *Endocrine.* 2022;78:458–469
34. Liu, J., Liu, Z., Liu, C., Sun, H., Li, X. and Yang, Y., Integrating Artificial Intelligence in the Diagnosis and Management of Metabolic Syndrome: A Comprehensive Review. *Diabetes Metab Res Rev.* 2025 41: e70039
35. Lee YA, Huang Y, Dai H, Yuce TK, Shah V, Bian J, Guo J. Characterize Disease Progression Subphenotypes in Real World Populations with Overweight and Obesity using a Graph-based Neural Network Framework. *medRxiv* 2025 Nov 13, 2025.11.10.25339913.
36. Huang, L.; Huhulea, E.N.; Abraham, E.; Bienenstock, R.; Aifuwa, E.; Hirani, R.; Schulhof, A.; Tiwari, R.K.; Etienne, M. The Role of Artificial Intelligence in Obesity Risk Prediction and Management: Approaches, Insights, and Recommendations. *Medicina* 2025, 61, 358.
37. Sarah E. Hampl, Sandra G. Hassink, Asheley C. Skinner, Sarah C. Armstrong, Sarah E. Barlow, Christopher F. Bolling, Kimberly C. Avila Edwards, Ihuoma Eneli, Robin Hamre, Madeline M. Joseph, Doug Lunsford, Eneida Mendonca, Marc P. Michalsky, Nazrat Mirza, Eduardo R. Ochoa, Mona Sharifi, Amanda E. Staiano, Ashley E. Weedn, Susan K. Flinn, Jeanne Lindros, Kymika Okechukwu; Clinical Practice Guideline for the Evaluation and Treatment of Children and Adolescents With Obesity. *Pediatrics* February 2023; 151 (2): e2022060640.
38. Z. Huang, M. P. Berry, C. Chwyl, G. Hsieh, J. Wei, and E. M. Forman, “Comparing Large Language Model AI and Human-Generated Coaching Messages for Behavioral Weight Loss,” *Journal of Technology in Behavioral Science*, 2025,1–12.
39. Machorro-Cano, I.; Alor-Hernández, G.; Paredes-Valverde, M.A.; Ramos-Deonati, U.; Sánchez-Cervantes, J.L.; Rodríguez-Mazahua, L. PISIoT: A Machine Learning and IoT-Based Smart Health Platform for Overweight and Obesity Management. *Appl. Sci.* 2019, 9, 3037
40. Vazquez-Briseno, M.; Navarro-Cota, C.; Nieto-Hipólito, J.; Jiménez-García, E.; Sanchez-Lopez, J. A proposal for using the internet of things concept to increase children’s health awareness. In *Proceedings of the CONIELECOMP 2012, 22nd International Conference on Electrical Communications and Computers*, Puebla, Mexico, 27–29 February 2012; pp. 168–172.
41. Zaragoza, I.; Guixeres, J.; Alcañiz, M.; Cebolla, A.; Saiz, J.; Álvarez, J. Ubiquitous monitoring and assessment of childhood obesity. *Pers. Ubiquit. Comput.* 2013, 17, 1147–1157.
42. Ellahham S Artificial Intelligence: The Future for Diabetes Care *The American Journal of Medicine*, 2020; 133, 895-900
43. Sun H, Zhu L, Wang P, Yuan K, Nawrin SS, Cui Y and Li L Dietary patterns and obesity are associated with type 2 diabetes risk in elderly Chinese men: a machine learning approach. *Front. Nutr.* 2025 12:1705683
44. Nomura, A., Noguchi, M., Kometani, M. et al. Artificial Intelligence in Current Diabetes Management and Prediction. *Curr Diab Rep* 2021 21, 61
45. Cardozo, Glauco, Pintarelli, Guilherme Brasil, Andreis, Guilherme Rettore, Lopes, Annelise Correa Wengerkievicz, Marques, Jefferson Luiz Brum, Use of Machine Learning and Routine Laboratory Tests for Diabetes Mellitus Screening, *BioMed Research International*, 2022, 8114049
46. Wang DD, Hu FB. Precision nutrition for prevention and management of type 2 diabetes. *Lancet Diabetes Endocrinol.* 2018;6(5):416–26.
47. de Toro-Martin J, et al. Precision nutrition: a review of personalized nutritional approaches for the prevention and management of metabolic syndrome. *Nutrients.* 2017;9(8):913
48. Yang J-X, Liu Y, Huang R, Wu H-y, Wang Y-y, Cao S-y, Wang G-y, Zhang J-M, Ai Z-S and Zhou H-m Development and internal validation of a machine learning algorithm for the risk of type 2 diabetes mellitus in children with obesity. *Front. Endocrinol.* 2025 16:1649988.
49. Li Y, Jin N, Zhan Q, Huang Y, Sun A, Yin F, Li Z, Hu J and Liu Z Machine learning-based risk predictive models for diabetic kidney disease in type 2 diabetes mellitus patients: a systematic review and meta-analysis. *Front. Endocrinol.* 2025 16:1495306.

50. Ahlqvist, Emma et al., Novel subgroups of adult-onset diabetes and their association with outcomes: a data-driven cluster analysis of six variables, *The Lancet Diabetes & Endocrinology*, 2018, Volume 6, Issue 5, 361 - 369
51. Anjana RM, Baskar V, Nair ATN, et al. Novel subgroups of type 2 diabetes and their association with microvascular outcomes in an Asian Indian population: a data-driven cluster analysis: the INSPIRED study. *BMJ Open Diab Res Care* 2020;8:e001506.
52. Xie, X., Wu, C., Yang, Y. et al. Interpretable machine learning-guided single-cell mapping deciphers multi-lineage pancreatic dysregulation in type 2 diabetes. *Cardiovasc Diabetol* 2025, 24, 300
53. Zeevi D, et al. Personalized nutrition by prediction of glycemic responses. *Cell*. 2015;163(5):1079–94
54. Mistry S, Riches NO, Gouripeddi R, Facelli JC. Environmental exposures in machine learning and data mining approaches to diabetes etiology: A scoping review. *Artif Intell Med*. 2023 Jan;135:102461.
55. Mendes-Soares H, et al. Model of personalized postprandial glycemic response to food developed for an Israeli cohort predicts responses in Midwestern American individuals. *Am J Clin Nutr*. 2019;110(1):63–75.
56. Mendes-Soares H, et al. Assessment of a personalized approach to predicting postprandial glycemic responses to food among individuals without diabetes. *JAMA Netw Open*. 2019;2(2):e188102.
57. Lou JJ, Zeng J. Artificial intelligence applications for managing metabolic dysfunction-associated steatotic liver disease: Current status and future prospects. *World J Gastroenterol* 2025; 31(47): 111900
58. Yang, Fang, XueyueSun, Kuijiang, MingxinZhang, and ChaoSun. "Recent Advances in the Application of Machine Learning Models in Metabolic Dysfunction–Associated Steatotic Liver Disease," *Diabetes/Metabolism Research and Reviews*, 2026, e70129.
59. M. Ji, Y. Jo, S. J. Choi, et al., "Plasma Metabolomics and Machine Learning-Driven Novel Diagnostic Signature for Non-Alcoholic Steato-hepatitis," *Biomedicines* 10, no. 7, 2022, 1669
60. M. Noureddin, F. Ntanos, D. Malhotra, et al., "Predicting NAFLD Prevalence in the United States Using National Health and Nutrition Examination Survey 2017–2018 Transient Elastography Data and Application of Machine Learning," *Hepatology Communications* 6, no. 7, 2022, 1537–1548
61. F. Razmpour, R. Daryabeygi-Khotbehsara, D. Soleimani, et al., "Application of Machine Learning in Predicting Non-Alcoholic Fatty Liver Disease Using Anthropometric and Body Composition Indices," *Scientific Reports* 13, no. 1, 2023, 4942
62. G. Huang, Q. Jin, and Y. Mao, "Predicting the 5-Year Risk of Nonalcoholic Fatty Liver Disease Using Machine Learning Models: Prospective Cohort Study," *Journal of Medical Internet Research* 25, 2023, e46891,
63. S. Qin, X. Hou, Y. Wen, et al., "Machine Learning Classifiers for Screening Nonalcoholic Fatty Liver Disease in General Adults," *Scientific Reports* 13, no. 1, 2023, 3638
64. P. Sorino, M. G. Caruso, G. Misciagna, et al., "Selecting the Best Machine Learning Algorithm to Support the Diagnosis of Non-Alcoholic Fatty Liver Disease: A Meta Learner Study," *PLoS One* 15, no. 10, 2020, e0240867
65. S. Y. Rhyou and J. C. Yoo, "Cascaded Deep Learning Neural Network for Automated Liver Steatosis Diagnosis Using Ultrasound Images," *Sensors* 21, no. 16, 2021, 5304
66. A. Das, M. Connell, and S. Khetarpal, "Digital Image Analysis of Ultrasound Images Using Machine Learning to Diagnose Pediatric Nonalcoholic Fatty Liver Disease," *Clinical Imaging* 77, 2021, 62–68
67. P. M. Graffy, V. Sandfort, R. M. Summers, and P. J. Pickhardt, "Automated Liver Fat Quantification at Nonenhanced Abdominal CT for Population-Based Steatosis Assessment," *Radiology* 293, no. 2, 2019, 334–342
68. D. Sethunath, S. Morusu, M. Tuceryan, et al., "Automated Assessment of Steatosis in Murine Fatty Liver," *PLoS One* 13, no. 5, 2018, e0197242
69. Y. Ramot, G. Zandani, Z. Madar, S. Deshmukh, and A. Nyska, "Utilization of a Deep Learning Algorithm for Microscope-Based Fatty Vacuole Quantification in a Fatty Liver Model in Mice," *Toxicologic Pathology* 48, no. 5, 2020, 702–707
70. S. Vanderbeck, J. Bockhorst, R. Komorowski, D. E. Kleiner, and S. Gawrieh, "Automatic Classification of White Regions in Liver Biopsies by Supervised Machine Learning," *Human Pathology* 45, no. 4, 2014, 785–792

71. R. Forlano, B. H. Mullish, N. Giannakeas, et al., "High-Throughput, Machine Learning-Based Quantification of Steatosis, Inflammation, Ballooning, and Fibrosis in Biopsies From Patients With Nonalcoholic Fatty Liver Disease," *Clinical Gastroenterology and Hepatology* 18, no. 9, 2020, 2081
72. F. Heinemann, G. Birk, and B. Stierstorfer, "Deep Learning Enables Pathologist-Like Scoring of NASH Models," *Scientific Reports* 9, no. 1, 2019, 18454
73. Chaulagain RP, Dinislam K, Shrestha Y, Yadav DK, Ali A. Advancing Diagnosis of Liver Cirrhosis: Why Non-invasive Methods Are the Future? *Cureus*. 2025 Dec 12;17(12):e99071.
74. Zhang Y, Liu X, Zhang X, Fei Y, Li X. Machine learning-based prediction of metabolic dysfunction-associated steatotic liver disease using National Health and Nutrition Examination Survey (NHANES) data. *PLoS One*. 2025 Nov 12;20(11):e0335656.
75. Li X, Liu S, Zhao Q, An M, Hou C, Hu S, Niu Y. Combining body roundness index and triglyceride-glucose index to enhance MASLD prediction: insights from NHANES and machine learning. *Hormones (Athens)*. 2025 Sep 24.
76. Zhou T, Ding X, Chen L, Huang Q, He L. Visceral adiposity index as a predictor of metabolic dysfunction-associated steatotic liver disease: a cross-sectional study. *BMC Gastroenterol*. 2025 May 1;25(1):326.
77. Wu X, Zhang T, Park S. Dietary quality, perceived health, and psychological status as key risk factors for newly developed metabolic dysfunction-associated steatotic liver disease in a longitudinal study. *Nutrition*. 2025 Feb;130:112604.
78. Wang Y, Huang DQ, Zhang P, Wang M, Wu Y, Nur E, Li L, Wang H. Plasma inflammatory proteome profiles identify MASLD among children with overweight or obesity. *Cardiovasc Diabetol*. 2025 Nov 27;24(1):450.
79. Li Y, Liu R, An Y, He F. Markers of body fat, the mediating role of alanine aminotransferase, and their association with the risk of metabolic dysfunction-associated steatotic liver disease. *Eur J Pediatr*. 2025 Aug 2;184(8):524.
80. Tavaglione F, Marafioti G, Romeo S, Jamialahmadi O. Machine Learning Reveals the Contribution of Lipoproteins to Liver Triglyceride Content and Inflammation. *J Clin Endocrinol Metab*. 2024 Dec 18;110(1):218-227.
81. Verma N, Duseja A, Mehta M, De A, Lin H, Wong VW, Wong GL, Rajaram RB, Chan WK, Mahadeva S, Zheng MH, Liu WY, Treeprasertsuk S, Prasoppokakorn T, Kakizaki S, Seki Y, Kasama K, Charatcharoenwitthaya P, Sathirawich P, Kulkarni A, Purnomo HD, Kamani, Lee L, Lee YY, Wong MS, Tan EXX, Young DY. Machine learning improves the prediction of significant fibrosis in Asian patients with metabolic dysfunction-associated steatotic liver disease - The Gut and Obesity in Asia (GO-ASIA) Study. *Food Pharmacol Ther*. 2024 Mar;59(6):774-788.
82. Gil-Rojas S, Suárez M, Martínez-Blanco P, Torres AM, Martínez-García N, Blasco P, Torralba M, Mateo J. Prognostic Impact of Metabolic Syndrome and Steatotic Liver Disease in Hepatocellular Carcinoma Using Machine Learning Techniques. *Metabolites*. 2024 May 27;14(6):305.
83. Shibata N, Morita Y, Ito T, Kanzaki Y, Watanabe N, Yoshioka N, Arai Y, Yasuda S, Koshiyama Y, Toyoda H, Morishima I. A machine learning algorithm for stratification of risk of cardiovascular disease in metabolic dysfunction-associated steatotic liver disease. *Eur J Intern Med*. 2024 Nov;129:62-70.
84. Wang C, Chen Y, Xiao H, Cai J, Wang R, Zeng X, Lin M, Liu W, Chi X, Chen Q. Metabolomics-guided machine learning reveals diagnostic and mechanistic biomarkers in CHB with MASLD. *PLoS One*. 2026 Feb 11;21(2):e0331529.
85. Zhan S, Wang X, Wang C, Zhu B, Gao J, Peng Z, Wang R, Yang Y, Zhang L, Wang T, Wu J, Wu W, Huang K, Dong G, Ren Q, Wang S, Wang S, Zhou X, Xu L, Fu J, Guo X. Tryptophan derivatives as non-invasive diagnostic indicators for obesity-related MASLD in children and adolescents. *Diabetes Obes Metab*. 2025 Dec;27(12):7544-7560.
86. Lu CH, Hsieh YR, Huang SY, Wang W, Chang CW, Panunggal B, Chang IW, Chen CL, Chang CC, Kao WY. Serum Lipidome as a Predictor of Significant Liver Fibrosis in Patients with Severe Obesity Undergoing Bariatric Surgery. *Obes Surg*. 2026 Feb;36(2):652-665.

87. Zöggeler T, Kavallar AM, Pollio AR, Aldrian D, Decristoforo C, Scholl-Bürgi S, Müller T, Vogel GF. Meta-analysis of shotgun sequencing of gut microbiota in obese children with MASLD or MASH. *Gut Microbes*. 2025 Dec;17(1):2508951.
88. Zhang D, Leitman M, Pawar S, Shera S, Hernandez L, Jacobs JP, Dong TS. The Association Between *Prevotella* covers and Advanced Fibrosis in the Progression of Metabolic Dysfunction-Associated Steatotic Liver Disease. *Nutrients*. 2025 Jun 27;17(13):2145.
89. Nychas E, Marfil-Sánchez A, Chen X, Mirhakkak M, Li H, Jia W, Xu A, Nielsen HB, Nieuwdorp M, Loomba R, Ni Y, Panagiotou G. Discovery of robust and highly specific microbiome signatures of non-alcoholic fatty liver disease. *Microbiome*. 2025 Jan 14;13(1):10.

Disclaimer/Publisher's Note: The statements, opinions and data contained in all publications are solely those of the individual author(s) and contributor(s) and not of MDPI and/or the editor(s). MDPI and/or the editor(s) disclaim responsibility for any injury to people or property resulting from any ideas, methods, instructions or products referred to in the content.s