



REVIEW

Artificial Intelligence in Bariatric Surgery: Current Status and Future Perspectives

Mustafa Bektaş¹ · Beata M. M. Reiber¹ · Jaime Costa Pereira² · George L. Burchell³ · Donald L. van der Peet¹

Received: 24 February 2022 / Revised: 3 June 2022 / Accepted: 3 June 2022 / Published online: 17 June 2022
© The Author(s) 2022

Abstract

Background Machine learning (ML) has been successful in several fields of healthcare, however the use of ML within bariatric surgery seems to be limited. In this systematic review, an overview of ML applications within bariatric surgery is provided.

Methods The databases PubMed, EMBASE, Cochrane, and Web of Science were searched for articles describing ML in bariatric surgery. The Cochrane risk of bias tool and the PROBAST tool were used to evaluate the methodological quality of included studies.

Results The majority of applied ML algorithms predicted postoperative complications and weight loss with accuracies up to 98%.

Conclusions In conclusion, ML algorithms have shown promising capabilities in the prediction of surgical outcomes after bariatric surgery. Nevertheless, the clinical introduction of ML is dependent upon the external validation of ML.

Keywords Artificial intelligence · Machine learning · Deep learning · Bariatric surgery

Introduction

Artificial intelligence (AI) is a new field in medicine gaining major interest within healthcare, but its development in clinical settings is already referred to as a digital revolution for healthcare [1].

Artificial intelligence is defined as computer science capable of imitating several aspects of human intelligence and behavior [2]. With the use of large datasets, AI models can be trained to conduct several complicated tasks [3]. Machine learning, one of the domains of AI, is a computer system in which models are trained to form new predictions or decisions by analyzing large quantities of data [4]. A specific subclass of machine learning known as deep learning uses multiple layers to analyze imported data. In each layer, weights are calculated for several factors from the data. After repeating this process, a final model is trained and ready to be applied on new data. Examples of both machine and deep learning techniques are presented in Table 1.

Several potentials of AI models have already been demonstrated in clinical practice [14, 15]. For example, machine learning algorithms have been applied to MRI, X-ray, and CT images to detect tumors in various organs. Additionally, input from large numbers of electronic health records enabled AI models to identify risk factors for multifactorial outcomes such as length of stay, mortality, and early hospital readmission after surgery [16]. Recently, in colorectal surgery, machine learning was used to predict outcomes such as lymph node metastasis, response to chemoradiotherapy, and postoperative complications. For these outcomes,

Key Points

Machine learning has been used to predict surgical outcomes in bariatric surgery.
Surgical outcomes have been predicted with accuracies up to 98%.
External validation is required for the clinical introduction of machine learning.

✉ Mustafa Bektaş
m.bektas@amsterdamumc.nl

- ¹ Department of Gastrointestinal Surgery, Amsterdam UMC Location Vrije Universiteit Amsterdam, De Boelelaan 1117, 1081 HV Amsterdam, the Netherlands
- ² Department of Computer Science, Vrije Universiteit Amsterdam, De Boelelaan 1105, 1081 HV Amsterdam, the Netherlands
- ³ Medical Library Department, Amsterdam UMC Location Vrije Universiteit Amsterdam, De Boelelaan 1117, 1081 HV Amsterdam, the Netherlands

Table 1 Definitions of subclasses within AI

Subclass	Definition
Machine learning (ML)	ML involves computer science that is able to perform desired tasks based on input data. When provided with sufficient data, algorithms can recognize patterns in data and train the model to perform better. After completion of the final model, the algorithm can be applied to new unknown data [5]
Decision tree (DT)	Within a DT model, multiple factors are classified into tree branches. Based on the algorithm, these branches are divided into nodes, forming several tree pathways. In the end, this model tends to find the smallest tree that optimally fits the data [6]
Gradient boosting (GBM)	In GBM, weights are added to several factors after classification. Afterwards an assessment of weights occurs, in which weights are modified based on the difficulty to classify the factors. this process is repeated until a final optimal model is generated [7]
Random forest (RF)	RF involves the formation of multiple decision trees with specific values for predictors. This technique combines all decision trees in order to build an accurate model for predictions [8]
Support vector machine (SVM)	SVM models use mapped input data to discover the optimal boundary to separate several classes and values [9]
Deep learning	As a specific branch of machine learning, deep learning can recognize patterns within datasets by using multiple processing layers. Within each layer, weights are present for several factors within the model. After the training process, an optimal model is built to perform on new data [10]
Artificial neural networks (ANNs)	Similar to our brain system, data is passed through multiple processing layers within ANNs. Each layer contains weights in order to make decisions for the resulting output. By repeat of this process, this model can improve results and produce the most accurate model in the end [11]
Convolutional neural networks (CNNs)	CNNs are a specific type of neural networks, however no weights are used in the layers. Instead, multiple layers are functioning as filters to register patterns or regions of images [12]
Radiomics	A radiomics model analyzes images in order to retrieve specific texture features that are registered as a 0 or 1. By detecting these features, various pathologies could be recognized [13]

Abbreviations: ML, machine learning; DT, decision tree; GBM, gradient boosting machine; RF, random forest; SVM, support vector machine; ANN, artificial neural networks; CNN, convolutional neural networks

predictions were performed with accuracies up to 96%. This could emphasize the potential of machine learning to support risk stratification and facilitate clinical decision-making for general surgeons [17–19].

Currently, bariatric surgery has evolved to being a key in treating the worldwide pandemic of morbid obesity. Optimal postoperative weight loss including resolution of obesity-related comorbidities leads to a decreased burden of disease and related mortality [20, 21]. Despite an increasing amount of large data set studies in bariatric surgery, several factors such as short- and long-term complication rates and weight loss remain unpredictable. An example in which AI could benefit bariatric surgery is insufficient weight loss after surgery. Ten to thirty percent of patients show insufficient weight loss after bariatric surgery [22]. Risk factors for this are extremely diverse varying from socio-economic factors such as insurance policy to a specific type of microbiome [23, 24]. A complete overview of all risk factors and ideally an algorithm to calculate the risk of insufficient weight loss for each patient separately is still missing. Assembling an algorithm to identify both patients at major risk of insufficient weight loss and high risk of postoperative complications would assist the bariatric surgeon as well as the patient to reach a well-informed decision.

Despite the potential benefits of AI, the scope of machine learning applications is rarely reported. Therefore, this

systematic review aims to provide an extensive overview of (potential) machine learning applications within bariatric surgery.

Materials and Methods

Search Strategy

A systematic search was performed in accordance with the *Cochrane Handbook for Systematic Reviews of Interventions* version 6.0 and PRISMA guidelines. To identify all relevant publications, systematic searches were conducted in the bibliographic databases PubMed, Embase.com, Clarivate Analytics/Web of Science Core Collection, and the Wiley/Cochrane Library from inception up to the 7th of July 2021. The search included keywords and free text terms for (synonyms of) ‘machine learning’ combined with (synonyms of) ‘digestive system surgical procedures’ and ‘bariatric surgery’. The full search strategy can be found in the Supplementary information (see Appendix).

Selection Process

Two reviewers (MB and JCP) conducted the title and abstract screening independently in accordance with the

inclusion and exclusion criteria. Studies were only selected for full-text assessment if both reviewers agreed on inclusion. Controversies between reviewers were resolved by discussions, resulting in consensus. Studies were included if they met the following criteria: (i) describing machine learning algorithms within bariatric surgery, (ii) clinical study, (iii) including adults. Studies were excluded if they (i) did not describe bariatric surgery specifically, (ii) were not written in English, (iii) were certain publication types: reviews, editorials, letters, legal cases, or interviews.

Risk of Bias Evaluation

The ROBINS-I assessment tool was applied by two reviewers (MB and JCP) to evaluate the methodological quality of included non-randomized studies [25]. Additionally, the PROBAST tool was used by two reviewers (MB and JCP) to assess the quality of machine learning models [26]. Conflicts between reviewers were solved by discussions.

Data Synthesis and Outcome Assessment

Following full-text screening, the following data were extracted from the included studies; first author, year of publication, country, number of patients included, mean age of the study population, percentage of female patients, study design, follow-up time, surgical procedure, type of machine learning, external validation, purpose of machine learning, outcome measurements, and prediction performance. The categorization of studies was based on machine learning purposes and results were demonstrated separately.

Results

Study Selection and Characteristics

The systematic literature search generated a total of 1821 references after removal of duplicates. After screening of titles and abstracts, 21 studies remained for full-text assessment. Eleven full texts were included. The flow chart of the search and selection process is presented in Fig. 1. Table 2 summarizes the general characteristics of the included studies.

Risk of Bias Evaluation

As all included studies were either retrospective ($n = 10$) or prospective ($n = 1$) cohort studies, the ROBINS-I assessment tool was used for quality assessment of all included studies (Fig. 2a). Since the primary outcome of this study was the type of machine learning techniques being used, domains such as bias due to confounding and bias in outcome measurements obtained low risk of bias scores. However, due to

the retrospective design of these studies, a moderate risk of bias was found in the intervention classification domain. Furthermore, results of the Probst score per domain are demonstrated in Fig. 2b.

Categorization of Machine Learning Techniques

Purposes of machine learning were prediction of postoperative complications ($n = 5$), prediction of the amount of postoperative weight loss ($n = 3$), aid in decision-making preoperatively ($n = 1$), predicting presence of hiatal hernias ($n = 1$), and prediction of quality of life ($n = 1$). The frequency at which each form of machine learning technique was used in the included studies is summarized in Fig 3.

Postoperative Complications

Five studies demonstrated the use of machine learning algorithms to predict postoperative complications.

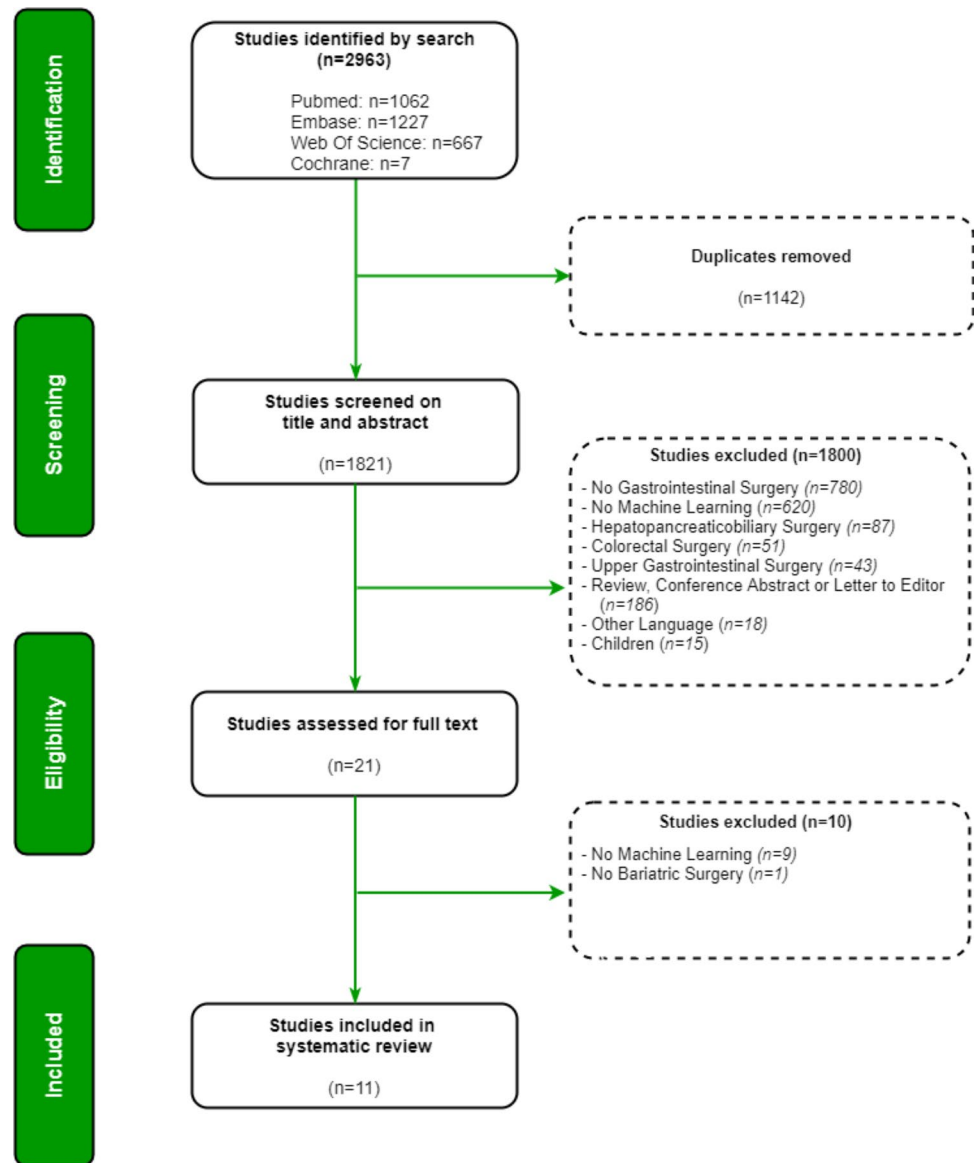
Sheikhtaheri et al. developed a model to predict postoperative complications within 90 days after one anastomosis gastric bypass surgery (OAGB), by using an ANN algorithm [27]. These complications included bleeding, anastomotic leakage, obstruction, intraabdominal abscess, and pulmonary embolism. Thirty-two factors ranging from age and BMI to smoking and laboratory test results were considered important in this prediction model. For the postoperative period of 10 days, the highest accuracy of the model was obtained; an AUC of 0.98 was observed.

Cao et al. (2019) applied multiple machine learning algorithms to detect severe complications within 30 days after bariatric surgery [28]. Machine learning techniques included decision tree, random forest, gradient boosting, SVM, and ANN models. Results have revealed the following performances for the models (accuracy, AUC): decision tree (92%, 0.5), random forest (95%, 0.51), gradient boosting (96%, 0.58), SVM (96%, 0.5), and ANN (96%, 0.54).

Consequently, Cao et al. (2020) applied ANN, and CNN models to predict serious complications within 30 days after bariatric surgery. Serious complications were defined as Clavien–Dindo classification grade 3b and higher (i.e., anastomotic leakage, organ failure, or death) [29]. For each model, the predictive performance was described by means of the accuracy, and AUC. The ANN model showed an accuracy of 84%, and an AUC of 0.54. For the CNN model, the accuracy was 95% and the AUC appeared to be 0.57 for predicting postoperative complications.

The authors of the 4th study used ANN and GBM models to predict gastrointestinal leak and venous thromboembolism in patients undergoing a laparoscopic gastric bypass or laparoscopic sleeve gastrectomy [30]. For gastrointestinal leakage, the ANN and GBM model showed the following predictive capabilities, respectively; an AUC of 0.75 and

Fig. 1 PRISMA flow diagram of the search



0.70. In predicting venous thromboembolisms, the ANN algorithm and gradient boosting model achieved the following values, respectively; an AUC of 0.65 and 0.67. Out of 37 variables, the most important factors in predicting both gastrointestinal leakage and venous thromboembolisms were age, height, and weight-related measures, hematocrit, albumin, and assistant training level. A history of deep vein thrombosis was an additional important variable for prediction of venous thromboembolisms.

Wise et al. (2019) aimed to predict the readmission rate of 3.1%, the reoperation and reintervention rate of 8.7%, and the mortality rate of 0.07% within 30 days after laparoscopic sleeve gastrectomy in a large cohort [31]. For this ANN model, an AUC of 0.59 was detected. Moreover, the

following seven factors appeared to be important for the prediction of 30-day morbidity and mortality: age, race, BMI, hypertension, diabetes mellitus, functional status, and previous surgery.

Weight Loss

All three studies aimed to predict postoperative weight loss by applying ANN models.

Piaggi et al. aimed to predict the percentage excess weight loss (%EWL) in women with severe obesity, 2 years after the laparoscopic adjustable gastric banding procedure [32]. %EWL at 2 years postoperatively was 48.2%. The ANN model developed was based on preoperative data including

Table 2 General characteristics of included studies

Authors	Year	Country	Patients	Age (mean)	Female (%)	Study design	Follow-up	Surgical procedures	Type of machine learning	External validation	ML Purpose	Study outcomes	Prediction performance (ACC/AUC)
Sheikhtaheri et al	2019	Iran	1509	39	NS	Retrospective Cohort	30 days	OAGB	Neural network	Yes	Predict post-operative complications	Accuracy; AUC	0.98/0.97
Cao et al	2019	Sweden	37,811	41	75.9	Retrospective Cohort	30 days	NS	Multiple machine learning	No	Predict post-operative complications	AUC	NA
Cao et al	2020	Sweden	44,061	42	NS	Retrospective Cohort	30 days	NS	Neural network	No	Predict post-operative complications	Accuracy; AUC	0.95/0.57
Nudel et al	2021	USA	436,807	45	79.3	Retrospective Cohort	30 days	Lap gastric bypass; LSG	Multiple machine learning	No	Predict post-operative complications	AUC	-/0.69
Wise et al	2020	USA	101,721	44	79.4	Retrospective Cohort	30 days	LSG	Neural network	No	Predict post-operative complications	AUC	-/0.59
Piaggi et al	2010	Italy	235	42	100	Retrospective Cohort	2 years	Gastric Banding	Neural network	No	Predict weight loss	AUC	-/0.80
Wise et al	2016	USA	647	47	79.6	Retrospective Cohort	1 year	Lap gastric bypass	Neural network	No	Predict weight loss	AUC	-/0.83
Lee et al	2007	Taiwan	249	33	71.1	Prospective Cohort	2 years	OAGB; Gastric Banding	Neural network	No	Predict weight loss	Accuracy	0.94/-
Aminian et al	2020	USA	13,722	54	65	Retrospective Cohort	4 years	Lap gastric bypass; LSG; Gastric Banding; Duodenal Switch	Random forest	No	Assist in decision-making	AUC	-/0.71
Assaf et al	2021	Israel	2482	43	62.7	Retrospective Cohort	-	LSG	Decision tree	No	Predict diagnosis of hiatal hernia	Accuracy	0.88/-
Cao et al	2019	Sweden	6687	43	77	Retrospective Cohort	5 years	Lap gastric bypass	Neural network	No	Predict post-operative Quality of Life	Mean squared error	NA

Abbreviations: LSG, laparoscopic sleeve gastrectomy; Lap gastric bypass, laparoscopic gastric bypass; OAGB, one-anastomosis gastric bypass; NS, not specified; ACC, accuracy; AUC, area under the curve; NA, not applicable

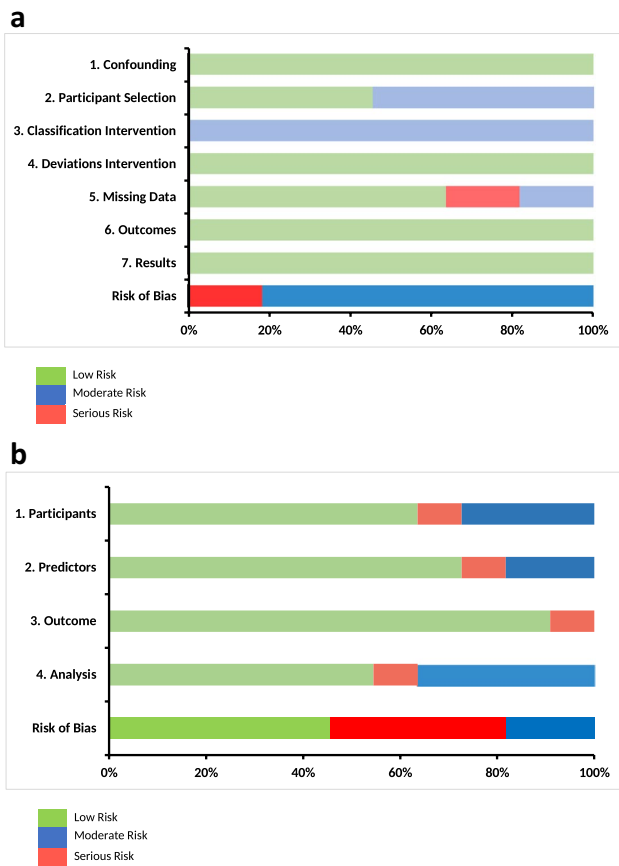


Fig. 2 **a** Methodological quality assessment of the non-randomized studies, according to ROBINS-I assessment tool. **b** Quality of machine learning models according to the Probst tool

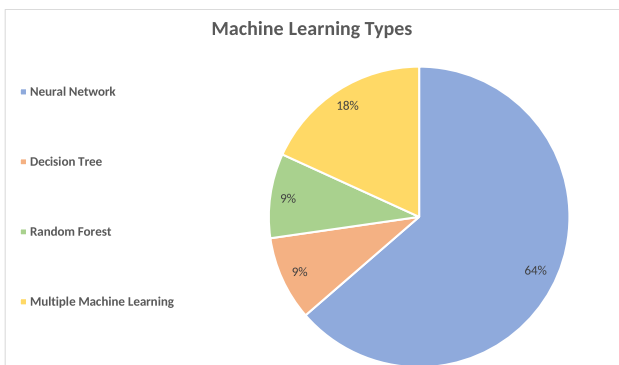


Fig. 3 Applied forms of machine learning

the comprehensive test of psychopathology Minnesota Multiphasic Personality Inventory-2. The model showed an AUC of 0.80 for this prediction. Age, paranoia, antisocial practices, and Type A behavior were independent predictors of %EWL.

Wise et al. (2016) used an ANN model to predict the percentage excess body mass index loss (%EBMIL) 180 and 360 days after laparoscopic Roux-en-Y gastric bypass surgery based on preoperative variables such as BMI, race, and gender [33]. The %EBMIL was 73.5% 1 year, postoperatively. The AUC for this model was observed to be 0.83. The variables gender, race, BMI, and diabetes mellitus appeared to be the key factors for postoperative weight loss.

Lastly, Lee et al. used 17 preoperative factors to predict successful %EWL 2 years after laparoscopic OAGB or gastric banding [34]. Success in %EWL was defined as %EWL > 50% which was accomplished by 84% of the patients. The ANN model showed an accuracy of 94% and the type of operation, HbA1c, and triglyceride levels appeared to be essential for predicting successful %EWL at 2 years postoperatively.

Decision-Making

Aminian et al. developed a prediction model using an RF algorithm to estimate the risk of long-term end-organ complications in patients with type 2 diabetes and obesity when considering bariatric surgery [35]. The discriminating ability at 10 years was measured in the area under the curve (AUC) and resulted in the following for the surgical and non-surgical groups, respectively; all-cause mortality 0.79 and 0.81, coronary artery events 0.66 and 0.67, heart failure 0.73 and 0.75, and nephropathy 0.73 and 0.76. The five most important variables in the prediction models of all-cause mortality were age, BMI at enrollment, history of heart failure, insulin use, and smoking status.

Diagnosis

Assaf et al. developed a decision tree model for the preoperative prediction of the presence of hiatal hernias (HH) in patients undergoing a laparoscopic sleeve gastrectomy procedure [36]. This is relevant as the presence of a hiatal hernia may impose per-operative technical challenges which is why foreknowledge is beneficial. The model showed an accuracy of 88.2% for the prediction of hiatal hernias. Thirteen variables were observed to be influencing the prediction of hiatal hernias, in which reflux symptoms, higher age, and BMI were discovered to be associated with a higher risk of hiatal hernias. Additionally, lower age and BMI have been discovered to be related to shorter operation lengths.

Postoperative Quality of Life

Cao et al. (2019) built a CNN model to predict the postoperative health-related quality of life 1, 2, and 5 years after

a primary gastric bypass procedure [37]. The postoperative quality of life was measured by the RAND-SF-36 questionnaire and the obesity-related problems scale (OP). Performance of the machine learning algorithm was presented as the mean squared error, indicating the discrepancy between the observed value and predicted value. The mean squared error for the CNN model was 0.035 in predicting the postoperative quality of life.

Discussion

From this systematic review, it can be concluded that artificial intelligence has potentials in several fields within bariatric surgery. Various models have been created to predict severe complications with AUCs up to 0.98. Secondly, weight loss was predicted by AUCs ranging from 0.80 to 0.83. Lastly, an AUC up to 0.81 was observed in predicting the postoperative quality of life, diagnosis, and end-organ complications of patients with morbid obesity.

Five studies have applied machine learning models to predict postoperative complications for patients undergoing bariatric surgery. Among several models, neural networks have shown the highest accuracy of 98% in predicting postoperative complications. Ideally, by using machine learning models, bariatric surgeons will be able to better predict (severe) postoperative complications for each unique patient. These predictions can, in theory, influence the decision towards a different type of bariatric operation or different timing of the operation, more specific prophylactic measures to prevent a certain type of complication, or a shared decision with complete informed consent.

In a recent study, the “low-risk bariatric patient” was defined by the absence of factors such as a medical history of thromboembolic events, diabetes mellitus, and kidney or pulmonary disease [38]. In this review, overlapping risk factors have been identified in the included studies predicting postoperative complications and weight loss (Table 3). It is of no surprise that age, BMI, previous intra-abdominal surgery, diabetes, and cardiovascular disease were identified as risk factors for postoperative severe complications.

However, other factors such as race, inflammatory bowel disease, laboratory results, and functional status are more controversial. Not all clinical variables were included in a similar or homogeneous manner across the included studies. This is despite the hypothesis that inclusion of previously excluded variables may improve the accuracy of machine learning models to predict postoperative complications and related risk factors. In the field of breast cancer surgery, the exclusion of variables in machine learning models was prevented by determining many variables based on pre-operative, intra-operative, and post-operative means [39]. These findings could suggest that guidelines are needed to secure a comprehensive list of clinical factors that can be used for an optimal training process of machine learning models.

Three studies have attempted to predict postoperative weight loss. Neural networks demonstrated the highest AUC of 0.94 in predicting postoperative weight loss. For decades now, researchers in the bariatric field have attempted to identify all risk factors for insufficient weight loss after bariatric surgery. Multiple studies have shown that postoperative weight loss is dependent on multiple factors, both objective measures such as BMI and subjective measures such as patient-related measures. It could therefore be specifically beneficial and interesting for bariatric surgeons to implement AI as a means of identifying risk factors for, for example, insufficient WL. However, as Nudel et al. noted [30], external validation of the machine learning model was missing due to insufficient data. Therefore, more large datasets are needed before accurate and valid models can be developed.

For predicting the risk of long-term end-organ complications, such as coronary artery events, heart failure, and nephropathy in patients suffering from type 2 diabetes and morbid obesity, a random forest model showed an AUC of 0.66, 0.73, and 0.73, respectively. According to Aminian et al. [35], this random forest model may support and accelerate the process of decision-making toward bariatric surgery. This is desirable as the duration of obesity itself and the presence of its related comorbidities have repeatedly been reported to lead to less postoperative weight loss and higher comorbidity-related mortality [40–42]. As weight loss after bariatric surgery is not always associated

Table 3 Summary of overlapping factors for postoperative complications and weight loss

Postoperative complications		Postoperative weight loss	
Protective factors	Risk factors	Helping factors	Inhibiting factors
Low BMI	Non-White race Diabetes mellitus* Older age Previous bariatric surgery	Female gender	Older age Diabetes mellitus* High BMI

BMI, body mass index

* = type not specified

with health-related quality of life, predicting the increase in quality of life after bariatric surgery is a welcome algorithm in the process of expectation management and shared decision-making, preoperatively [43, 44]. Neural networks have shown a mean squared error of 0.035 in predicting the postoperative health-related quality of life 1, 2, and 5 years after bariatric surgery, indicating an accurate estimation, since the mean squared error was close to 0. This neural network model might provide the opportunity to improve postoperative care and rehabilitation for patients undergoing bariatric surgery. However, due to missing patient information, the generalizability of this model might be uncertain. Missing data could be solved by imputation, as this was done in the study of Tseng et al. [45], in which machine learning models were used to predict acute kidney injury after cardiac surgery. One study predicted the presence of hiatal hernias. The importance of hiatal hernia (HH) present at the time of bariatric surgery remains controversial but is increasingly recommended to be corrected simultaneously with the laparoscopic sleeve gastrectomy [46]. Nevertheless, gastroesophageal reflux symptoms may worsen or persist, and a secondary operation with conversion from sleeve gastrectomy to LRYGB may be necessary [47]. The foreknowledge of the presence of HH may both influence the patient and surgeon in decision-making towards LRYGB and predict a longer operation time. However, as the authors of this study mention, the accuracy of the models developed is not impressive and the study should be regarded as proof of concept, exploring the possibilities with AI.

Due to the missing external validation in most studies, the first step for future studies in bariatric surgery should be the inclusion of external validation cohorts to gain more generalizability of machine learning models. Afterwards, clinical trials should be conducted to facilitate the implementation of ML models within bariatric surgery. For both steps, large amounts of data are required for the training process of these models. This data could be retrieved from available patient databases or robotic surgery, eventually facilitating the training process of machine learning [48, 49].

This review has revealed that machine learning models have potentials to predict postoperative complications, weight loss, end-organ complications, quality of life, and preoperative diagnosis. After the necessary steps to improve generalizability and clinical validation, machine learning

models may have a significant impact on decision-making within bariatric surgery. As machine learning models are improved and validated, surgeons could be one step closer to achieving personalized decision-making for patients undergoing bariatric procedures.

To use machine learning models for the prediction of surgical outcomes in bariatric surgery, data from laparoscopic bariatric surgery should be accessible [50]. Laparoscopic videos of bariatric procedures could be collected to serve as a training database for machine learning models. By providing accurate image navigation during surgery, anatomical landmarks and unexpected intraoperative findings such as adhesions and abdominal wall hernias could be identified efficiently by machine learning models [51]. In addition, perioperative data collected from anesthesiologists could be collected such as continuous blood pressure measures or oxygen saturation as factors possibly predicting postoperative complications. Furthermore, as robotic surgery is often performed in bariatric surgery, machine learning models could also improve the performance of robotic surgery by providing 3D mapping during surgery and evaluating surgical skills afterward [52].

This review has several limitations. External validation cohorts seem to be missing for most studies, indicating the uncertainty of machine learning models. Therefore, big data from clinical settings are required to achieve appropriate generalizability and accuracy for machine learning models [53]. Additionally, due to the presence of inconsistencies in reported accuracies and AUCs, a meta-analysis could not be conducted.

Conclusion

In this review, promising predictive capabilities of machine learning have been discovered within bariatric surgery. Machine learning has predominantly been used for prediction of postoperative complications and weight loss. However, ML algorithms have mainly been applied to datasets without external validation. To overcome this problem, additional data from large patient databases, laparoscopic surgery, or robotic surgery should be used. By validating ML models, the clinical implementation of ML will be facilitated.

Appendix

Table 4

Table 4 Search strategy in PubMed

Search	Query	Results
#3	#1 AND #2	1062
#2	“Digestive System Surgical Procedures”[Mesh] OR “Bariatric Surgery”[Mesh] OR “Laparotomy”[Mesh] OR “Roux-en-Y”[Tiab] OR “Cholecystostom*”[Tiab] OR “Choledochostom*”[Tiab] OR “Gastroenterostom*”[Tiab] OR “Jejunioleal Bypass*”[Tiab] OR “Pancreaticojejunostom*”[Tiab] OR “Peritoneovenous Shunt”[Tiab] OR “Portoenterostom*”[Tiab] OR “Gastric Bypass*”[Tiab] OR “Appendectom*”[Tiab] OR “Cholecystectom*”[Tiab] OR “Sphincterotom*”[Tiab] OR “Colectom*”[Tiab] OR “Cecostom*”[Tiab] OR “Colostom*”[Tiab] OR “Duodenostom*”[Tiab] OR “Ileostom*”[Tiab] OR “Jejunostom*”[Tiab] OR “Esophagectom*”[Tiab] OR “Hemorrhoidectom*”[Tiab] OR “Hepatectom*”[Tiab] OR “Liver Transplant*”[Tiab] OR “Pancreas Transplant*”[Tiab] OR “Pancreatectom*”[Tiab] OR “Pancreaticoduodenectom*”[Tiab] OR “Proctectom*”[Tiab] OR “gastrectom*”[tiab] OR “Gastrostom*”[tiab] OR “Esophagoplast*”[tiab] OR “Esophagostom*”[tiab] OR “Hepatectom*”[tiab]	489,138
#1	“Machine Learning”[Mesh] OR “Machine Learning”[Tiab] OR “machine intelligen*”[tiab] OR “machine vision*”[tiab] OR “machine learning”[tiab] OR “transfer learning”[tiab] OR “deep learning”[tiab] OR “neural network*”[tiab] OR “support vector machine*”[tiab] OR “automatic segmentation*”[tiab] OR “Long short term memory”[tiab] OR “LSTM”[tiab] OR “supervised learning”[tiab] OR “unsupervised learning”[tiab] OR “reinforcement learning*”[tiab] OR “hierarchical learning*” [tiab] OR “Image Interpretation*”[tiab] OR “Prediction model*”[tiab] OR “image recognition”[tiab] OR “perceptron”[tiab]	154,754

Table 5

Table 5 Search strategy in Embase.com

Search	Query	Results
#5	#3 NOT #4	1227
#4	#3 AND (‘chapter’/it OR ‘conference abstract’/it OR ‘conference paper’/it OR ‘conference review’/it OR ‘editorial’/it OR ‘erratum’/it OR ‘letter’/it OR ‘note’/it OR ‘short survey’/it OR ‘tombstone’/it)	857
#3	#1 AND #2	2084
#2	‘gastrointestinal surgery’/exp OR ‘laparotomy’/exp OR ‘biliary tract surgery’/exp OR (‘Roux-en-Y’ OR ‘Cholecystostom*’ OR ‘Choledochostom*’ OR ‘Gastroenterostom*’ OR ‘Jejunioleal Bypass*’ OR ‘Pancreaticojejunostom*’ OR ‘Peritoneovenous Shunt’ OR ‘Portoenterostom*’ OR ‘Gastric Bypass*’ OR ‘Appendectom*’ OR ‘Cholecystectom*’ OR ‘Sphincterotom*’ OR ‘Colectom*’ OR ‘Cecostom*’ OR ‘Colostom*’ OR ‘Duodenostom*’ OR ‘Ileostom*’ OR ‘Jejunostom*’ OR ‘Esophagectom*’ OR ‘Hemorrhoidectom*’ OR ‘Hepatectom*’ OR ‘Liver Transplant*’ OR ‘Pancreas Transplant*’ OR ‘Pancreatectom*’ OR ‘Pancreaticoduodenectom*’ OR ‘Proctectom*’ OR ‘gastrectom*’ OR ‘Gastrostom*’ OR ‘Esophagoplast*’ OR ‘Esophagostom*’ OR ‘Hepatectom*’):ti,ab,kw	720,511
#1	‘machine learning’/exp OR (‘Machine Learning’ OR ‘machine intelligen*’ OR ‘machine vision*’ OR ‘machine learning’ OR ‘transfer learning’ OR ‘deep learning’ OR ‘neural network*’ OR ‘support vector machine*’ OR ‘automatic segmentation*’ OR ‘Long short term memory’ OR ‘LSTM’ OR ‘supervised learning’ OR ‘unsupervised learning’ OR ‘reinforcement learning*’ OR ‘hierarchical learning*’ OR ‘Image Interpretation*’ OR ‘Prediction model*’ OR ‘image recognition’ OR ‘perceptron’):ti,ab,kw	335,846

Table 6 Search strategy in Clarivate Analytics/Web of Science Core Collection

Search	Query	Results
#4	#1 AND #2 Refined by: [excluding] DOCUMENT TYPES: (LETTER OR MEETING ABSTRACT OR EDITORIAL MATERIAL OR CORRECTION)	667
#3	#1 AND #2	747
#2	TS=(“Roux-en-Y” OR “Cholecystostom*” OR “Choledochostom*” OR “Gastroenterostom*” OR “Jejunioleal Bypass*” OR “Pancreaticojejunosom*” OR “Peritoneovenous Shunt” OR “Portoenterostom*” OR “Gastric Bypass*” OR “Appendectom*” OR “Cholecystectom*” OR “Sphincterotom*” OR “Colectom*” OR “Cecostom*” OR “Colostom*” OR “Duodenostom*” OR “Ileostom*” OR “Jejunostom*” OR “Esophagectom*” OR “Hemorrhoidectom*” OR “Hepatectom*” OR “Liver Transplant*” OR “Pancreas Transplant*” OR “Pancreatectom*” OR “Pancreaticoduodenectom*” OR “Proctectom*” OR “gastrectom*” OR “Gastrostom*” OR “Esophagoplast*” OR “Esophagostom*” OR “Hepatectom*”)	294,577
#1	TS=(“Machine Learning” OR “machine intelligen*” OR “machine vision*” OR “machine learning” OR “transfer learning” OR “deep learning” OR “neural network*” OR “support vector machine*” OR “automatic segmentation*” OR “Long short term memory” OR “LSTM” OR “supervised learning” OR “unsupervised learning” OR “reinforcement learning*” OR “hierarchical learning*” OR “Image Interpretation*” OR “Prediction model*” OR “image recognition” OR “perceptron”)	477,557

Table 7 Search strategy in Wiley/Cochrane Library

Search	Query	Results
#3	#1 AND #2	7
#2	(“Roux en Y” OR “Cholecystostom*” OR “Choledochostom*” OR “Gastroenterostom*” OR “Jejunioleal Bypass*” OR “Pancreaticojejunosom*” OR “Peritoneovenous Shunt” OR “Portoenterostom*” OR “Gastric Bypass*” OR “Appendectom*” OR “Cholecystectom*” OR “Sphincterotom*” OR “Colectom*” OR “Cecostom*” OR “Colostom*” OR “Duodenostom*” OR “Ileostom*” OR “Jejunostom*” OR “Esophagectom*” OR “Hemorrhoidectom*” OR “Hepatectom*” OR “Liver Transplant*” OR “Pancreas Transplant*” OR “Pancreatectom*” OR “Pancreaticoduodenectom*” OR “Proctectom*” OR “gastrectom*” OR “Gastrostom*” OR “Esophagoplast*” OR “Esophagostom*” OR “Hepatectom*”):ti,ab,kw	4113
#1	(“Machine Learning” OR “machine intelligen*” OR “machine vision*” OR “machine learning” OR “transfer learning” OR “deep learning” OR “neural network*” OR “support vector machine*” OR “automatic segmentation*” OR “Long short term memory” OR “LSTM” OR “supervised learning” OR “unsupervised learning” OR “reinforcement learning*” OR “hierarchical learning*” OR “Image Interpretation*” OR “Prediction model*” OR “image recognition” OR “perceptron”):ti,ab,kw	4733

Author contribution Mustafa Bektaş: participated in the design of the study, data collection and interpretation. Wrote and submitted the manuscript.

Beata M.M. Reiber: participated in the design of the study, wrote parts of the manuscript, and revised the manuscript critically.

Jaime Costa Pereira: participated in the design and data collection of the study.

George L. Burchell: performed the literature search and wrote parts of the manuscript.

Donald L. van der Peet: participated in the design of the study, wrote parts of the manuscript, and revised the manuscript critically.

All authors approved the final version of the manuscript.

Declarations

Ethics approval and consent to participate This article does not contain any studies with human participants or animals performed by any of the authors. Informed consent does not apply.

Conflict of interest The authors declare no competing interests.

Open Access This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit <http://creativecommons.org/licenses/by/4.0/>.

References

- Briganti G, Le Moine O. Artificial intelligence in medicine: today and tomorrow. *Front Med*. 2020;7:27.
- Chen M, Decary M. Artificial intelligence in healthcare: an essential guide for health leaders. *Health Manage Forum*. 2020;33(1):10–8.
- Kulkarni S, Seneviratne N, Baig MS, Khan AHA. Artificial intelligence in medicine: where are we now? *Acad Radiol*. 2020;27(1):62–70.
- Davenport T, Kalakota R. The potential for artificial intelligence in healthcare. *Future Healthc J*. 2019;6(2):94–8.
- El Naqa I, Murphy MJ. What Is Machine Learning? In: El Naqa I, Li R, Murphy M, editors. *Machine learning in radiation oncology*. Cham: Springer; 2015. p. 3–11.
- Stiglic G, Kocbek S, Pernek I, Kokol P. Comprehensive decision tree models in bioinformatics. *PLoS ONE*. 2012;7(3): e33812.
- Friedman JH. (2002). Stochastic gradient boosting. *Computational Statistics & Data Analysis*. 2002;38(4):367–378.
- Breiman L. Random Forests. *Mach Learn*. 2001;45:5–32.
- Zhang L, Zhou W, Jiao L. Wavelet support vector machine. *IEEE Trans Syst Man Cybern B Cybern*. 2004;34(1):34–9.
- Rusk N. Deep learning. *Nat Methods*. 2016;13:35.
- Abraham A. Artificial neural networks. *Handbook of measuring system design*. New Jersey: John Wiley & Sons; 2005:901–908.
- Albawi S, Mohammed TA, Al-Zawi S. Understanding of a convolutional neural network. *2017 International Conference on Engineering and Technology (ICET)*; 2017 Aug 1–6.
- Kumar V, Gu Y, Basu S, Berglund A, Eschrich SA, Schabath MB, et al. Radiomics: the process and the challenges. *Magn Reson Imaging*. 2012;30(9):1234–48.
- Mintz Y, Brodie R. Introduction to artificial intelligence in medicine. *Minim Invasive Ther Allied Technol*. 2019;28(2):73–81.
- Topol EJ. High-performance medicine: the convergence of human and artificial intelligence. *Nat Med*. 2019;25(1):44–56.
- Hashimoto DA, Rosman G, Rus D, Meireles OR. Artificial intelligence in surgery: promises and perils. *Ann Surg*. 2018;268(1):70–6.
- Eresen A, Li Y, Yang J, Shangguan J, Velichko Y, Yaghmai V, et al. Preoperative assessment of lymph node metastasis in colon cancer patients using machine learning: a pilot study. *Cancer imaging: the official publication of the International Cancer Imaging Society*. 2020;20(1):30.
- Shaish H, Aukerman A, Vanguri R, Spinelli A, Armenta P, Jambawalikar S, et al. Radiomics of MRI for pretreatment prediction of pathologic complete response, tumor regression grade, and neoadjuvant rectal score in patients with locally advanced rectal cancer undergoing neoadjuvant chemoradiation: an international multicenter study. *Eur Radiol*. 2020;30(11):6263–73.
- Bunn C, Kulshrestha S, Boyda J, Balasubramanian N, Birch S, Karabayir I, et al. Application of machine learning to the prediction of postoperative sepsis after appendectomy. *Surgery*. 2021;169(3):671–7.
- Schauer PR, Bhatt DL, Kirwan JP, Wolski K, Aminian A, Brethauer SA, et al. Bariatric surgery versus intensive medical therapy for diabetes - 5-year outcomes. *N Engl J Med*. 2017;376(7):641–51.
- Syn NL, Cummings DE, Wang LZ, Lin DJ, Zhao JJ, Loh M, et al. Association of metabolic-bariatric surgery with long-term survival in adults with and without diabetes: a one-stage meta-analysis of matched cohort and prospective controlled studies with 174 772 participants. *Lancet*. 2021;397(10287):1830–41.
- Linke K, Schneider R, Gebhart M, Ngo T, Slawik M, Peters T, et al. Outcome of revisional bariatric surgery for insufficient weight loss after laparoscopic Roux-en-Y gastric bypass: an observational study. *Surg Obes Relat Dis*. 2020;16(8):1052–9.
- Karmali S, Brar B, Shi X, Sharma AM, de Gara C, Birch DW. Weight recidivism post-bariatric surgery: a systematic review. *Obes Surg*. 2013;23(11):1922–33.
- Faria SL, Santos A, Magro DO, Cazzo E, Assalin HB, Guadagnini D, et al. Gut microbiota modifications and weight regain in morbidly obese women after Roux-en-Y gastric bypass. *Obes Surg*. 2020;30(12):4958–66.
- Sterne JA, Hernán MA, Reeves BC, Savović J, Berkman ND, Viswanathan M, et al. ROBINS-I: a tool for assessing risk of bias in non-randomised studies of interventions. *BMJ*. 2016;355: i4919.
- Moons KGM, Wolff RF, Riley RD, Whiting PF, Westwood M, Collins GS, et al. PROBAST: a tool to assess risk of bias and applicability of prediction model studies: explanation and elaboration. *Ann Intern Med*. 2019;170(1):W1–33.
- Sheikhtaheri A, Orooji A, Pazouki A, Beitollahi M. A clinical decision support system for predicting the early complications of one-anastomosis gastric bypass surgery. *Obes Surg*. 2019;29(7):2276–86.
- Cao Y, Fang X, Ottosson J, Näslund E, Stenberg E. A comparative study of machine learning algorithms in predicting severe complications after bariatric surgery. *J Clin Med*. 2019;8(5):668.
- Cao Y, Montgomery S, Ottosson J, Näslund E, Stenberg E. Deep learning neural networks to predict serious complications after bariatric surgery: analysis of Scandinavian obesity surgery registry data. *JMIR Med Inform*. 2020;8(5): e15992.
- Nudel J, Bishara AM, de Geus SWL, Patil P, Srinivasan J, Hess DT, et al. Development and validation of machine learning models to predict gastrointestinal leak and venous thromboembolism after weight loss surgery: an analysis of the MBSAQIP database. *Surg Endosc*. 2021;35(1):182–91.

31. Wise ES, Amateau SK, Ikramuddin S, Leslie DB. Prediction of thirty-day morbidity and mortality after laparoscopic sleeve gastrectomy: data from an artificial neural network. *Surg Endosc.* 2020;34(8):3590–6.
32. Piaggi P, Lippi C, Fierabracci P, Maffei M, Calderone A, Mauri M, et al. Artificial neural networks in the outcome prediction of adjustable gastric banding in obese women. *PLoS ONE.* 2010;5(10): e13624.
33. Wise ES, Hocking KM, Kavic SM. Prediction of excess weight loss after laparoscopic Roux-en-Y gastric bypass: data from an artificial neural network. *Surg Endosc.* 2016;30(2):480–8.
34. Lee YC, Lee WJ, Lee TS, Lin YC, Wang W, Liew PL, et al. Prediction of successful weight reduction after bariatric surgery by data mining technologies. *Obes Surg.* 2007;17(9):1235–41.
35. Aminian A, Zajichek A, Arterburn DE, Wolski KE, Brethauer SA, Schauer PR, et al. Predicting 10-year risk of end-organ complications of type 2 diabetes with and without metabolic surgery: a machine learning approach. *Diabetes Care.* 2020;43(4):852–9.
36. Assaf D, Rayman S, Segev L, Neuman Y, Zippel D, Goitein D. Improving pre-bariatric surgery diagnosis of hiatal hernia using machine learning models. *Minimally invasive therapy & allied technologies: MITAT: official journal of the Society for Minimally Invasive Therapy.* 2021;1–7.
37. Cao Y, Raoof M, Montgomery S, Ottosson J, Näslund I. Predicting long-term health-related quality of life after bariatric surgery using a conventional neural network: a study based on the Scandinavian obesity surgery registry. *J Clin Med.* 2019;8(12):2149.
38. Gero D, Raptis DA, Vleeschouwers W, van Veldhuisen SL, Martin AS, Xiao Y, et al. Defining global benchmarks in bariatric surgery: a retrospective multicenter analysis of minimally invasive Roux-en-Y gastric bypass and sleeve gastrectomy. *Ann Surg.* 2019;270(5):859–67.
39. Juwara L, Arora N, Gornitsky M, Saha-Chaudhuri P, Velly AM. Identifying predictive factors for neuropathic pain after breast cancer surgery using machine learning. *Int J Med Inform.* 2020;141: 104170.
40. Guerreiro V, Neves JS, Salazar D, Ferreira MJ, Oliveira SC, Souteiro P, et al. Long-term weight loss and metabolic syndrome remission after bariatric surgery: the effect of sex, age, metabolic parameters and surgical technique - a 4-year follow-up study. *Obes Facts.* 2019;12(6):639–52.
41. Heller S, Lingvaj I, Marso SP, Tsimikas AP, Pieber TR, Poulter NR, et al. Development of a hypoglycaemia risk score to identify high-risk individuals with advanced type 2 diabetes in DEVOTE. *Diabetes Obes Metab.* 2020;22(12):2248–56.
42. Anand SS, Islam S, Rosengren A, Franzosi MG, Steyn K, Yusufali AH, et al. Risk factors for myocardial infarction in women and men: insights from the INTERHEART study. *Eur Heart J.* 2008;29(7):932–40.
43. Waljee JF, Ghaferi A, Cassidy R, Varban O, Finks J, Chung KC, et al. Are patient-reported outcomes correlated with clinical outcomes after surgery?: A population-based study. *Ann Surg.* 2016;264(4):682–9.
44. Doll HA, Petersen SE, Stewart-Brown SL. Obesity and physical and emotional well-being: associations between body mass index, chronic illness, and the physical and mental components of the SF-36 questionnaire. *Obes Res.* 2000;8(2):160–70.
45. Tseng PY, Chen YT, Wang CH, Chiu KM, Peng YS, Hsu SP, et al. Prediction of the development of acute kidney injury following cardiac surgery by machine learning. *Crit Care.* 2020;24(1):478.
46. Chen W, Feng J, Wang C, Wang Y, Yang W, Dong Z, et al. Effect of concomitant laparoscopic sleeve gastrectomy and hiatal hernia repair on gastroesophageal reflux disease in patients with obesity: a systematic review and meta-analysis. *Obes Surg.* 2021;31(9):3905–18.
47. Peng BQ, Zhang GX, Chen G, Cheng Z, Hu JK, Du X. Gastroesophageal reflux disease complicating laparoscopic sleeve gastrectomy: current knowledge and surgical therapies. *Surgery for obesity and related diseases: official journal of the American Society for Bariatric Surgery.* 2020;16(8):1145–55.
48. Chand M, Ramachandran N, Stoyanov D, Lovat L. Robotics, artificial intelligence and distributed ledgers in surgery: data is key! *Tech Coloproctol.* 2018;22(9):645–8.
49. Balla A, Batista Rodríguez G, Corradetti S, Balagué C, Fernández-Ananín S, Targarona EM. Outcomes after bariatric surgery according to large databases: a systematic review. *Langenbecks Arch Surg.* 2017;402(6):885–99.
50. Johnston SS, Morton JM, Kalsekar I, Ammann EM, Hsiao CW, Repts J. Using machine learning applied to real-world healthcare data for predictive analytics: an applied example in bariatric surgery. *Value in health: the journal of the International Society for Pharmacoeconomics and Outcomes Research.* 2019;22(5):580–6.
51. Kitaguchi D, Takeshita N, Hasegawa H, Ito M. Artificial intelligence-based computer vision in surgery: recent advances and future perspectives. *Annals of gastroenterological surgery.* 2021;6(1):29–36.
52. Bhandari M, Zeffiro T, Reddiboina M. Artificial intelligence and robotic surgery: current perspective and future directions. *Curr Opin Urol.* 2020;30(1):48–54.
53. Ngiam KY, Khor IW. Big data and machine learning algorithms for health-care delivery. *Lancet Oncol.* 2019;20(6):293.

Publisher's Note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.