

Review

Performance of Machine Learning in Diagnosing KRAS (Kirsten Rat Sarcoma) Mutations in Colorectal Cancer: Systematic Review and Meta-Analysis

Kaixin Chen^{1*}, BS; Yin Qu^{1*}, MD; Ye Han¹, BS; Yan Li², MM; Huiyan Gao³, MM; De Zheng¹, MD

¹Department of Anorectal Surgery, Shuguang Hospital, Shanghai University of Traditional Chinese Medicine; Anorectal Disease Institute of Shuguang Hospital, Shanghai, China

²Department of Traditional Chinese Medicine Anorectal Surgery, Gongli Hospital of Shanghai Pudong New Area, Shanghai, China

³Traditional Chinese Medicine Department, Shanghai Pudong New Area Beicai Community Health Service Center, Pudong New Area, Shanghai, China

*these authors contributed equally

Corresponding Author:

De Zheng, MD

Department of Anorectal Surgery

Shuguang Hospital, Shanghai University of Traditional Chinese Medicine; Anorectal Disease Institute of Shuguang Hospital

528 Zhangheng Road

Shanghai, 201203

China

Phone: 86 13641763662

Email: zd1232@sina.com

Abstract

Background: With the widespread application of machine learning (ML) in the diagnosis and treatment of colorectal cancer (CRC), some studies have investigated the use of ML techniques for the diagnosis of KRAS (Kirsten rat sarcoma) mutation. Nevertheless, there is scarce evidence from evidence-based medicine to substantiate its efficacy.

Objective: Our study was carried out to systematically review the performance of ML models developed using different modeling approaches, in diagnosing KRAS mutations in CRC. We aim to offer evidence-based foundations for the development and enhancement of future intelligent diagnostic tools.

Methods: PubMed, Cochrane Library, Embase, and Web of Science were systematically retrieved, with the search cutoff date set to December 22, 2024. The encompassed studies are publicly published research papers that use ML to diagnose KRAS gene mutations in CRC. The risk of bias in the encompassed models was evaluated via the PROBAST (Prediction Model Risk of Bias Assessment Tool). A meta-analysis of the model's concordance index (c-index) was performed, and a bivariate mixed-effects model was used to summarize sensitivity and specificity based on diagnostic contingency tables.

Results: A total of 43 studies involving 10,888 patients were included. The modeling variables were derived from clinical characteristics, computed tomography (CT), magnetic resonance imaging (MRI), positron emission tomography/computed tomography, and pathological histology. In the validation cohort, for the ML model developed based on CT radiomic features, the c-index, sensitivity, and specificity were 0.87 (95% CI 0.84-0.90), 0.85 (95% CI 0.80-0.89), and 0.83 (95% CI 0.73-0.89), respectively. For the model developed using MRI radiomic features, the c-index, sensitivity, and specificity were 0.77 (95% CI 0.71-0.83), 0.78 (95% CI 0.72-0.83), and 0.73 (95% CI 0.63-0.81), respectively. For the ML model developed based on positron emission tomography/computed tomography radiomic features, the c-index, sensitivity, and specificity were 0.84 (95% CI 0.77-0.90), 0.73, and 0.83, respectively. Notably, the deep learning (DL) model based on pathological images demonstrated a c-index, sensitivity, and specificity of 0.96 (95% CI 0.94-0.98), 0.83 (95% CI 0.72-0.91), and 0.87 (95% CI 0.77-0.92), respectively. The DL model MRI-based model showed a c-index of 0.93 (95% CI 0.90-0.96), sensitivity of 0.85 (95% CI 0.75-0.91), and specificity of 0.83 (95% CI 0.77-0.88).

Conclusions: ML is highly accurate in diagnosing KRAS mutations in CRC, and DL models based on MRI and pathological images exhibit particularly strong diagnosis accuracy. More broadly applicable DL-based diagnostic tools may be developed in the future. However, the clinical application of DL models remains relatively limited at present. Therefore, future research

should focus on increasing sample sizes, improving model architectures, and developing more advanced DL models to facilitate the creation of highly efficient intelligent diagnostic tools for KRAS mutation diagnosis in CRC.

J Med Internet Res 2025;27:e73528; doi: [10.2196/73528](https://doi.org/10.2196/73528)

Keywords: colorectal cancer; Kirsten rat sarcoma viral oncogene; deep learning; machine learning; radiomics; PRISMA

Introduction

Colorectal cancer (CRC) is the third most prevalent malignancy around the globe, with around 2 million new cases and 935,000 deaths in 2020, making up 10.7% of global cancer incidence and 9.5% of cancer-related mortality [1]. Its incidence significantly varies across different regions, at 36.4 per 100,000 population in America, 28.9 per 100,000 in Europe, and 28.8 per 100,000 in China [1,2]. Although the incidence has declined among the old in high-income countries, an upward trend is observed in emerging economies and among individuals under the age of 50 years worldwide [3]. This pattern reflects the critical public health challenge posed by CRC, prompting nations to intensify preventive efforts, promote early screening, and optimize therapeutic strategies to mitigate its societal and individual health impacts.

Currently, colorectal surgeries mainly involve the employment of laparoscopic minimally invasive techniques and robotic assistance [4], while modulation of the gut microbiota has proven effective in preventing anastomotic leakage [5]. Moreover, classical chemotherapy regimens combining oxaliplatin with 5-fluorouracil have demonstrated improved survival outcomes [6]. Immune checkpoint inhibitors exhibit specific efficacy in patients with mismatch repair deficiency [7], and targeted therapies aimed at suppressing A20 enhance immune responses and overcome drug resistance [8]. Despite continuous therapeutic advancements, overall patient prognosis remains suboptimal, with genetic mutations and molecular subtypes identified as key prognostic determinants [9]. Specifically, KRAS (Kirsten rat sarcoma) mutations are among the most common driver mutations, present in nearly 40% of CRC patients. Such mutations cause tumor cell activation, increase drug resistance, and are closely associated with poorer survival rates and diminished therapeutic efficacy [10]. However, current diagnostic methods have limitations. Tissue biopsy remains the gold standard for KRAS mutation diagnosis, yet its invasiveness and sampling constraints hinder comprehensive assessment of tumor heterogeneity, particularly in reflecting mutational discrepancies between primary and metastatic lesions [11]. Additionally, noninvasive circulating tumor DNA assays face sensitivity challenges in early-stage disease owing to low tumor DNA concentrations and dilution effects. Although combining circulating tumor DNA with trans-renal tumor DNA from urine samples may enhance diagnosis, such approaches are not yet widely adopted [12]. At the same time, emerging technologies such as surface-enhanced Raman scattering polymerase chain reaction and droplet digital polymerase chain reaction offer high sensitivity for KRAS mutation diagnosis but face limitations in clinical application owing to high costs,

operational complexity, and insufficient clinical validation [13]. Therefore, there is an urgent need to explore novel methodologies to help with the diagnosis of KRAS mutations in CRC.

As artificial intelligence (AI) technologies advance, machine learning (ML) has garnered considerable attention from clinicians owing to its capability to integrate high-dimensional data effectively [14]. In clinical practice, ML is mainly used for biomarker diagnosis and prognostic analysis support [15]. As a critical subset of ML, deep learning (DL) excels in image data processing, significantly enhancing the diagnosis accuracy of cancer-specific survival in CRC histopathological analyses compared to traditional methods [16]. Some studies have used ML techniques to diagnose KRAS mutation status in CRC, revealing that DL can not only assess microsatellite instability to predict chemotherapy responses [17] but also integrate imaging, genomic, and clinical data to accurately diagnose KRAS mutations, thereby facilitating personalized targeted therapies [18]. Nevertheless, ML models encounter various challenges related to modeling discrepancies during application. The complexity and variability of clinical data can result in model performance instability and an increased risk of overfitting [19]. Differences in image data quality, resolution, and equipment can affect feature extraction and diagnosis capabilities, leading to poor performance on external datasets [20]. Moreover, limited pathological image annotation and heterogeneity constrain model training efficacy, leading to inconsistent predictions across different pathological subtypes [21].

At present, systematic evidence on how modeling variations impact the diagnosis of KRAS mutations in CRC remains limited. Therefore, this study aims to review the effectiveness of ML under varying modeling parameters and present evidence-based insights to facilitate the future application of AI technologies in CRC diagnosis and treatment.

Methods

Study Registration

This study was undertaken as per the PRISMA (Preferred Reporting Items for Systematic Reviews and Meta-Analyses, 2020) guidelines and has been prospectively registered on the PROSPERO (International Prospective Register of Systematic Reviews) platform (CRD42025610919) [22].

Eligibility Criteria

The inclusion criteria were as follows:

1. Studies involving patients diagnosed with CRC;

2. Case-control, cohort, and cross-sectional studies about study designs;
3. An ML model was well constructed to differentiate KRAS mutation status. The model development process involved feature selection (via traditional ML techniques), model training, performance evaluation, and, where applicable, model validation; and
4. Publications reported in English.

The exclusion criteria were as follows:

1. Studies that fail to differentiate CRC from malignancies of other systems;
2. Meta-analyses, reviews, guidelines, expert opinions, or publicly available conference abstracts not subjected to peer review;
3. To evaluate the model's performance, the study incorporated at least one key metric, such as the concordance index (c-index), sensitivity, specificity, accuracy, precision, confusion matrix, F_1 -score, or calibration curve, during the ML model construction process. If these metrics were absent, the model was excluded from the analysis; and
4. Those with an inadequate sample size (fewer than 20 cases). This requirement was set because of the stringent case number demands during model development. Typically, each additional variable necessitates an increase of 10 times in the number of positive cases (KRAS mutations), that is, ensuring an event per variable (EPV) greater than 10. Model development is generally based on 2 or more variables. However, when the number of cases is fewer than 20, the EPV requirement is not met.

Data Sources and Search Strategy

The retrieval of PubMed, Cochrane Library, Embase, and Web of Science databases was thoroughly performed, covering studies published up to December 22, 2024. The search strategy integrated both Medical Subject Headings and free-text terms, with no restrictions on geographical region or publication year. The search strategies are detailed in [Multimedia Appendix 1](#).

Study Selection and Data Extraction

The retrieved studies were uploaded to EndNote (version 21.4, Thomson ResearchSoft). After duplicates were removed, titles and abstracts were checked to identify potentially eligible studies. Full-text papers were subsequently reviewed to determine final inclusion based on predefined criteria.

Before our data extraction, a standardized electronic data extraction form was developed, encompassing the variables as follows: first author, year of publication, country of the authors, study design, patient source, tumor staging, number of KRAS-mutant cases, total number of cases, approach of validation set generation, strategies to prevent overfitting, number of KRAS-mutant cases and total cases in the training and validation sets, sorts of models used, and variables used for model construction.

Data extraction was independently undertaken by 2 researchers (KC and DZ), which was followed by cross-verification. In cases of dissent, a third researcher (YQ) was asked for a decision.

Risk of Bias in Studies

The risk of bias in eligible studies was assessed via the PROBAST (Prediction Model Risk of Bias Assessment Tool) [23], which evaluates 4 key domains: participants, predictors, outcomes, as well as analysis, to determine the overall risk of bias and applicability of these studies. Every domain comprises multiple specific items, with response options including “yes or probably yes” (low risk), “no or probably no” (high risk), and “no information” (unclear risk). A domain is categorized as high risk if any item within it is classified as high risk, and as low risk if all items are rated as low risk.

Furthermore, as our study incorporated a substantial amount of radiomics research, the radiomics quality score (RQS) was used for quality assessment. The RQS consists of 16 criteria, with a maximum score of 36 points. These criteria include: image protocol quality, multi-segmentation, modality studies, image acquisition time, feature dimensionality reduction, model construction using both omics and nonimaging features (prognosis and molecular subtyping), diagnosis and discussion of biological relevance, threshold analysis, discriminative statistics, calibration statistics, prospective studies registered in trial databases, validation, comparison with the “gold standard,” potential clinical utility, cost-benefit analysis, and open science or data availability [24].

The risk of bias assessment and quality assessment were independently carried out by 2 researchers (KC and DZ), with cross-checking undertaken after completion. All disagreements were settled through consultation with a third researcher (YQ).

Synthesis Methods

A meta-analysis was undertaken on the c-index, a key metric for evaluating the overall accuracy of ML models. For studies lacking data on 95% CIs or SEs (SDs), SDs were estimated based on the methodology proposed by Debray et al [25]. Heterogeneity across studies was evaluated via the I^2 statistic. A random-effects model was applied if I^2 exceeded 50%, indicating substantial heterogeneity; otherwise, a fixed-effects model was used for I^2 values below 50%.

Furthermore, a meta-analysis of sensitivity and specificity was carried out via a bivariate mixed-effects model. This analysis was based on diagnostic 2×2 contingency tables. However, since these data were not reported directly in most primary studies, the necessary information was derived by integrating reported sensitivity, specificity, precision, and case numbers. The meta-analysis was enabled by Stata (version 15.1; StataCorp LLC).

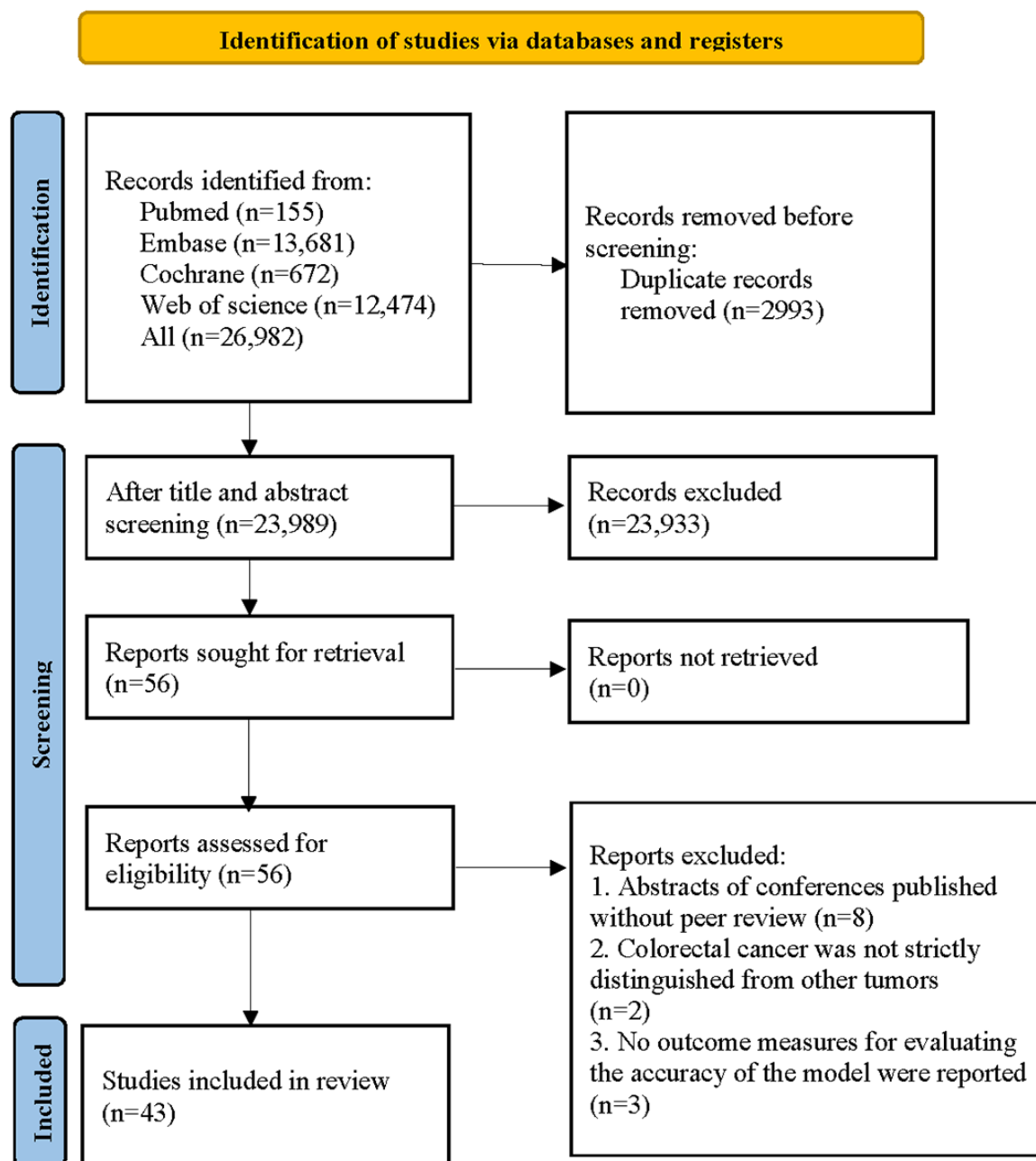
Results

Study Selection

A total of 26,982 papers were retrieved after our literature search. After 2993 duplicates were removed, the titles and abstracts of the rest were meticulously screened, leading to the removal of 23,933 studies deemed irrelevant to

the research topic. This initial screening process identified 56 potentially eligible articles. Following this, a thorough full-text evaluation was conducted. The full-text conference abstracts published without peer review did not rigorously distinguish between CRC and other cancers in the original research, nor did they report outcome measures assessing the accuracy of the evaluation models. Ultimately, 43 eligible studies were included in the final analysis [26-68] (Figure 1).

Figure 1. Literature screening process.



Study Characteristics

Of the 43 eligible studies, all were case-control studies. Among them, 30 studies originated from China, 3 studies from South Korea, and the rest were conducted in countries such as the United States, Argentina, and France. Data from 27 studies were derived from single-center sources, 9 were

multicenter studies, and 4 were based on registry databases, with some studies not explicitly reporting their data sources. Regarding the focus of the research, 12 studies concentrated on rectal cancer (RC), while the rest addressed CRC.

In terms of validation cohort generation, 33 studies provided relevant methodological details. The majority used

random sampling, while some used cross-validation, internal validation, or external validation methods, with a minority applying the leave-one-out approach. The modeling variables were predominantly derived from computed tomography (CT) and magnetic resonance imaging (MRI) imaging data. Twelve studies developed models based on pathological slides or clinical characteristics. Additionally, 8 studies focused on multigene analyses involving KRAS, NRAS (neuroblastoma ras viral oncogene homolog), and BRAF (v-raf murine sarcoma viral oncogene homolog B), whereas the remaining studies concentrated on single-gene KRAS analyses. All characteristics of the included studies are detailed in [Multimedia Appendix 2](#).

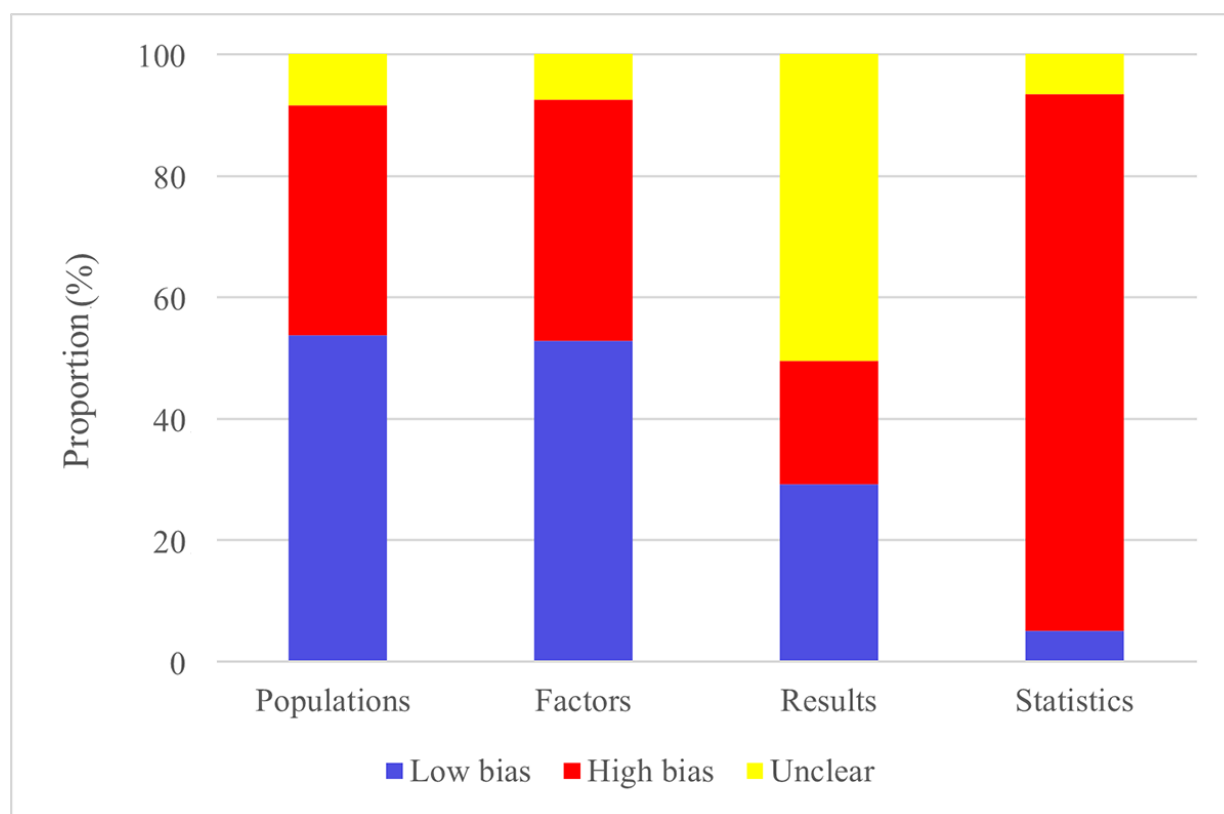
Risk of Bias in Studies

A risk of bias assessment was performed for the 74 ML models encompassed in our review. Of these, 63 models used data from case-control studies that were not based on registry databases, reflecting a high risk of bias in study participant

selection. Regarding predictor variables, since all models in this review were based on case-control studies, 4 models that relied on clinical characteristics may present a high risk of bias. However, most models were based on radiomics and pathological specimen data, which generally pose a lower risk of bias.

In our outcome assessment, as KRAS mutations were primarily diagnosed through pathological examinations serving as the gold standard, none of the encompassed models exhibited a risk of bias in this domain. Furthermore, for some ML models based on radiomics, the risk of bias related to case number selection remains unclear owing to the inability to derive the EPV ratio. Overall, data derived from radiomics and pathological specimens are relatively reliable, whereas data from nonregistry databases may affect the representativeness of study populations and introduce potential bias. Detailed risk of bias assessment results are presented in [Figure 2](#).

Figure 2. Risk assessment of the included models.



Additionally, a quality assessment was conducted for the 33 radiomics studies included in the analysis. The results indicated that the following 6 items, “study of modality,” “image acquisition time,” “dimensionality reduction of features,” “diagnosis and discussion of biological relevance,” “prospective studies registered in trial databases,” and “comparison with the ‘gold standard,’” were not addressed in any of the studies, resulting in a score of 0 for these categories. All studies included a complete imaging protocol, thereby earning a score. Five studies did not perform multiple segmentations and received no score in this regard.

Three studies explicitly stated the use of radiomics combined with clinical features for model construction and were accordingly awarded points. Eleven studies used threshold analysis, earning points for this criterion. All studies reported diagnostic statistics, resulting in full points for this item. Ten studies calibrated their statistical data, thereby receiving points. Most studies involved validation of their datasets, with multicenter studies scoring 3 points, single-center studies scoring 2 points, and 4 studies scoring 0 due to unclear validation methods. Eleven studies discussed the potential clinical utility of the models, earning points for this

aspect. Only 1 study discussed cost-effectiveness analysis, and 1 point is thus earned. Four studies did not provide the scientific and data foundations for the models they developed, while the remaining studies offered open-source region-of-interest segmentation, gaining 1 point each. The final average score was 7.4 points (18.90%, SD 4.0; [Multimedia Appendix 3](#)).

Meta-Analysis: KRAS Mutation

In the CRC training cohort, 36 models reported the c-index effect size for isolated KRAS mutations, with 34 models providing data directly extractable or indirectly calculable through 2*2 diagnostic tables. The c-index of models based on clinical characteristics was 0.69 (95% CI 0.68-0.70), with

sensitivity and specificity ranging from 0.66 to 0.79 and 0.57 to 0.58. For models based on CT radiomics, the c-index, sensitivity, and specificity were 0.86 (95% CI 0.82-0.89), 0.92 (95% CI 0.84-0.96), and 0.73 (95% CI 0.62-0.82). Models using MRI presented a c-index of 0.73 (95% CI 0.65-0.81), with a sensitivity and specificity of 0.75 (95% CI 0.68-0.81) and 0.65 (95% CI 0.49-0.79). Positron emission tomography/computed tomography (PET/CT)-enabled models showed a c-index of 0.66 (95% CI 0.52-0.81), with sensitivity and specificity ranges of 0.62-0.65 and 0.55-0.80. Additionally, pathology-based models exhibited a c-index of 0.87 (95% CI 0.79-0.94), sensitivity of 0.82 (95% CI 0.73-0.88), and specificity of 0.81 (95% CI 0.70-0.88; [Tables 1 and 2](#) and [Figures S1-S5 in Multimedia Appendix 4](#)).

Table 1. Meta-analysis of the concordance index (c-index) for ML^a diagnosis of MSI^b in CRC^c.

Modeling variables and model types	Training set			Validation set		
	Models, n	c-index (95% CI)	I ²	Models, n	c-index (95% CI)	I ²
Clinical features						
LR ^d	3	0.69 (0.68-0.70)	0	3	0.64 (0.57-0.71)	32.8
DL ^e	— ^f	—	—	1	0.67 (0.57-0.77)	—
Overall	3	0.69 (0.68-0.70)	0	4	0.65 (0.60-0.70)	4.5
CT^g						
LR	9	0.81 (0.72-0.91)	96.9	7	0.81 (0.70-0.92)	94.5
LASSO ^h	4	0.94 (0.89-0.99)	91.7	3	0.87 (0.80-0.93)	0
RF ⁱ	1	0.98 (0.97-1.00)	—	2	0.89 (0.69-1.08)	84.1
SVM ^j	1	0.86 (0.81-0.91)	—	1	0.68 (0.53-0.83)	—
Boosting	2	0.89 (0.72-1.05)	95.3	4	0.91 (0.85-0.97)	86.3
Ensemble	—	—	—	2	0.93 (0.82-1.03)	78.6
KNN ^k	1	0.67 (0.59-0.75)	—	—	—	—
Bayes	1	0.73 (0.65-0.81)	—	—	—	—
Overall	19	0.86 (0.82-0.89)	96	19	0.87 (0.84-0.90)	89.2
MRI^l						
LR	1	0.68 (0.60-0.75)	—	3	0.74 (0.50-0.98)	94.6
LASSO	1	0.80 (0.69-0.91)	—	2	0.68 (0.62-0.75)	0
LDA ^m	—	—	—	1	0.67 (0.58-0.76)	—
ANN ⁿ	1	0.71 (0.63-0.79)	—	1	0.68 (0.42-0.94)	—
SVM	1	0.72 (0.65-0.79)	—	2	0.70 (0.62-0.78)	0
DT ^o	2	0.74 (0.47-1.02)	96.7	2	0.61 (0.54-0.68)	0
DL	—	—	—	5	0.93 (0.90-0.96)	70.5
Overall	6	0.73 (0.65-0.81)	85.3	16	0.77 (0.71-0.83)	94.1
PET^p/CT						
LR	2	0.66 (0.52-0.81)	72.7	—	—	—
DL	—	—	—	1	0.84 (0.77-0.90)	—
Overall	2	0.66 (0.52-0.81)	72.7	1	0.84 (0.77-0.90)	—
Pathology						
LR	1	0.77 (0.68-0.86)	—	1	0.64 (0.40-0.87)	—
ANN	1	0.71 (0.65-0.77)	—	—	—	—
RF	2	0.93 (0.87-1.00)	69.1	2	0.81 (0.72-0.89)	0
Boosting	2	0.92 (0.83-1.01)	82.5	2	0.80 (0.68-0.92)	0
DL	—	—	—	4	0.96 (0.94-0.98)	96.7

Modeling variables and model types	Training set			Validation set		
	Models, n	c-index (95% CI)	I^2	Models, n	c-index (95% CI)	I^2
Overall	6	0.87 (0.79-0.94)	92.7	9	0.94 (0.91-0.96)	93.4

^aML: machine learning.^bMSI: microsatellite instability.^cCRC: colorectal cancer.^dLR: logistic regression.^eDL: deep learning.^fNot applicable.^gCT: computed tomography.^hLASSO: least absolute shrinkage and selection operator.ⁱRF: random forest.^jSVM: support vector machine.^kKNN: k-nearest neighbors.^lMRI: magnetic resonance imaging.^mLDA: linear discriminant analysis.ⁿANN: artificial neural network.^oDT: decision tree.^pPET: positron emission tomography.**Table 2.** Meta-analysis of the sensitivity and specificity of ML^a in diagnosing MSI^b in CRC^c.

Modeling variables and model types	Training set			Validation set		
	Models, n	Sen ^d (95% CI)	Spe ^e (95% CI)	Models, n	Sen (95% CI)	Spe (95% CI)
Clinical features						
LR ^f	2	0.66-0.79	0.57-0.58	3	0.55-0.79	0.47-0.55
DL ^g	— ^h	—	—	1	0.64	0.65
Overall	2	0.66-0.79	0.57-0.58	4	0.67 (0.57-0.76)	0.54 (0.47-0.61)
CT ⁱ						
LR	8	0.84 (0.76-0.90)	0.77 (0.66-0.85)	7	0.80 (0.74-0.84)	0.72 (0.63-0.79)
LASSO ^j	4	0.97 (0.66-1.00)	0.74 (0.36-0.93)	3	0.71-0.86	0.23-1.00
RF ^k	1	1.00	0.93	2	0.82-0.85	0.75-0.95
SVM ^l	3	0.82-0.86	0.71-0.80	2	0.77-0.78	0.63-0.81
Boosting	2	1.00-1.01	0.18-0.61	6	0.92 (0.84-0.97)	0.90 (0.74-0.97)
KNN ^m	1	0.80	0.55	—	—	—
Ensemble	—	—	—	2	0.93-0.97	0.58-0.93
Overall	19	0.92 (0.84-0.96)	0.73 (0.62-0.82)	22	0.85 (0.80-0.89)	0.83 (0.73-0.89)
MRI ⁿ						
LR	1	0.78	0.59	3	0.63-0.88	0.45-0.91
LASSO	1	0.64	0.85	2	0.56-0.58	0.64-1.00
DT ^o	2	0.83-0.84	0.38-0.8	2	0.71-0.76	0.44-0.54
SVM	1	0.71	0.66	2	0.71-0.72	0.63-0.69
LDA ^p	—	—	—	1	0.77	0.51
DL	—	—	—	5	0.85 (0.75-0.91)	0.83 (0.77-0.88)
Overall	5	0.75 (0.68-0.81)	0.65 (0.49-0.79)	15	0.78 (0.72-0.83)	0.73 (0.63-0.81)
PET ^q /CT						
LR	2	0.62-0.65	0.55-0.80	—	—	—
DL	—	—	—	1	0.73	0.83
Overall	2	0.62-0.65	0.55-0.80	1	0.73	0.83
Pathology						
LR	1	0.76	0.7	1	0.69	0.8
ANN ^r	1	0.67	0.61	—	—	—
RF	2	0.85-0.89	0.80-0.92	2	0.63-0.83	0.80-0.85
Boosting	2	0.76-0.92	0.84-0.88	2	0.63-0.75	0.90-0.91

Modeling variables and model types	Training set			Validation set		
	Models, n	Sen ^d (95% CI)	Spe ^e (95% CI)	Models, n	Sen (95% CI)	Spe (95% CI)
DL	—	—	—	4	0.92 (0.83-0.97)	0.88 (0.71-0.96)
Overall	6	0.82 (0.73-0.88)	0.81 (0.70-0.88)	9	0.83 (0.72-0.91)	0.87 (0.77-0.92)

^aML: machine learning.

^bMSI: microsatellite instability.

^cCRC: colorectal cancer.

^dSen: sensitivity.

^eSpe: specificity.

^fLR: logistic regression.

^gDL: deep learning.

^hNot applicable.

ⁱCT: computed tomography.

^jLASSO: least absolute shrinkage and selection operator.

^kRF: random forest.

^lSVM: support vector machine.

^mKNN: k-nearest neighbors.

ⁿMRI: magnetic resonance imaging.

^oDT: decision tree.

^pLDA: linear discriminant analysis.

^qPET: positron emission tomography.

^rANN: artificial neural network.

In the validation cohort, the c-index was extracted from 49 models, while sensitivity and specificity were calculable from 51 models. The c-index, sensitivity, and specificity for clinical characteristic-based models were 0.65 (95% CI 0.60-0.70), 0.67 (95% CI 0.57-0.76), and 0.54 (95% CI 0.47-0.61). For CT radiomics-based models, the corresponding values were 0.87 (95% CI 0.84-0.90), 0.85 (95% CI 0.80-0.89), and 0.83 (95% CI 0.73-0.89). MRI-based models demonstrated a c-index of 0.77 (95% CI 0.71-0.83), sensitivity of 0.78 (95% CI 0.72-0.83), and specificity of 0.73 (95% CI 0.63-0.81). PET/CT-based models had a c-index of 0.84 (95% CI 0.77-0.90), with sensitivity and specificity of 0.73 and 0.83. Pathology-based models showed a c-index of 0.94 (95% CI 0.91-0.96), sensitivity of 0.83 (95% CI 0.72-0.91), and specificity of 0.87 (95% CI 0.77-0.92). These findings indicate significant performance differences across data types, with pathology and CT radiomics models demonstrating superior diagnosis accuracy. It was found that models based solely on clinical features exhibited significantly lower diagnosis accuracy compared to those constructed using MRI, CT, and PET/CT radiomic features, as well as those built from pathological images. Furthermore, the model based on pathological images demonstrated the highest diagnosis accuracy (Tables 1 and 2 and Figures S6-S10 in Multimedia Appendix 4).

For the RC training set, 7 models analyzed isolated KRAS mutations, yielding overall c-index, sensitivity, and specificity effect sizes of 0.77 (95% CI 0.63-0.91), 0.77 (95% CI 0.67-0.85), and 0.59 (95% CI 0.42-0.74). In the validation cohort, 18 models provided a c-index of 0.76 (95% CI 0.71-0.82), while sensitivity and specificity from 14 models were 0.78 (95% CI 0.70-0.85) and 0.70 (95% CI 0.60-0.79; Tables 3 and 4 and Figures S11 and S12 in Multimedia Appendix 4).

Given the potential impact of model types on results, subgroup analyses were conducted. Some MRI and pathology-based models used DL algorithms. In the CRC MRI validation cohort, DL-based models demonstrated a c-index of 0.93 (95% CI 0.90-0.96), sensitivity of 0.85 (95% CI 0.75-0.91), and specificity of 0.83 (95% CI 0.77-0.88). In pathology-based models, the c-index was 0.96 (95% CI 0.94-0.98), with a sensitivity of 0.92 (95% CI 0.83-0.97) and specificity of 0.88 (95% CI 0.71-0.96).

For RC, the c-index effect size in the MRI validation cohort was calculated from 13 models, yielding 0.74 (95% CI 0.65-0.82). Among these, 10 models provided sensitivity and specificity estimates of 0.75 (95% CI 0.67-0.82) and 0.66 (95% CI 0.54-0.75).

Table 3. Meta-analysis of the c-index^a for ML^b in diagnosing MSI^c in RC^d.

Modeling variables	Training set			Validation set		
	Model, n	c-index (95% CI)	I ²	Model, n	c-index (95% CI)	I ²
Clinical features	— ^e	—	—	1	0.67 (0.57-0.77)	—
MRI ^f	5	0.73 (0.64-0.83)	88.2	13	0.74 (0.65-0.82)	94.6
PET ^g /CT ^h	—	—	—	1	0.84 (0.77-0.90)	—
CT	2	0.86 (0.61-1.11)	95.6	1	0.81 (0.65-0.98)	—
Pathology	—	—	—	2	0.88 (0.73-1.04)	91.3
Overall	7	0.77 (0.63-0.91)	97.3	18	0.76 (0.71-0.82)	95.2

Modeling variables	Training set			Validation set		
	Model, n	c-index (95% CI)	I^2	Model, n	c-index (95% CI)	I^2
^a c-index: concordance index.						
^b ML: machine learning.						
^c MSI: microsatellite instability.						
^d RC: rectal cancer.						
^e Not applicable.						
^f MRI: magnetic resonance imaging.						
^g PET: positron emission tomography.						
^h CT: computed tomography.						

Table 4. Meta-analysis of the sensitivity and specificity of ML^a in diagnosing MSI^b in RC^c.

Modeling variables	Training set			Validation set		
	Model, n	Sen ^d (95% CI)	Spe ^e (95% CI)	Model, n	Sen (95% CI)	Spe (95% CI)
Clinical features	— ^f	—	—	1	0.64	0.65
MRI ^g	5	0.75 (0.68-0.81)	0.65 (0.49-0.79)	10	0.75 (0.67-0.82)	0.66 (0.54-0.75)
PET ^h /CT ⁱ	—	—	—	1	0.73	0.83
CT	2	0.64-0.1	0.22-0.71	—	—	—
Pathology	—	—	—	2	0.91-0.98	0.61-0.94
Overall	7	0.77 (0.67-0.85)	0.59 (0.42-0.74)	14	0.78 (0.70-0.85)	0.70 (0.60-0.79)

^aML: machine learning.^bMSI: microsatellite instability.^cRC: rectal cancer.^dSen: sensitivity.^eSpe: specificity.^fNot applicable.^gMRI: magnetic resonance imaging.^hPET: positron emission tomography.ⁱCT: computed tomography.

KRAS in Conjunction With Other Gene Mutations

In the study of mixed genotypes, the training set included 5 models, with a c-index, sensitivity, and specificity of 0.87 (95% CI 0.77-0.96), 0.89 (95% CI 0.77-0.96), and 0.67

(95% CI 0.50-0.80). In the validation set, the pooled c-index derived from 6 models was 0.90 (95% CI 0.84-0.96), while the sensitivity and specificity, calculated from 4 models, were 0.83 (95% CI 0.73-0.90) and 0.70 (95% CI 0.51-0.84; [Tables 5 and 6](#) and [Figures S13 and S14 in Multimedia Appendix 4](#)).

Table 5. Meta-analysis of the c-index^a for ML^b-based diagnosis of KRAS^c and other gene mutation status in CRC^d and RC^e.

Modeling variables	Training set			Validation set		
	Model, n	c-index (95% CI)	I^2	Model, n	c-index (95% CI)	I^2
CT ^f	2	0.90 (0.79-1.02)	72.6	1	0.79 (0.66-0.92)	— ^g
Pathology	2	0.87 (0.68-1.06)	98.3	1	0.82 (0.66-0.99)	—
PET ^h /CT	1	0.76 (0.60-0.92)	—	1	0.70 (0.48-0.93)	—
MRI ⁱ	—	—	—	3	0.95 (0.91-0.98)	39.5
Overall	5	0.87 (0.77-0.96)	94.3	6	0.90 (0.84-0.96)	69.4

^ac-index: concordance index.^bML: machine learning.^cKRAS: Kirsten rat sarcoma.^dCRC: colorectal cancer.^eRC: rectal cancer.^fCT: computed tomography.^gNot applicable.^hPET: positron emission tomography.ⁱMRI: magnetic resonance imaging.

Table 6. Meta-analysis of sensitivity and specificity for ML^a-based diagnosis of KRAS^b and other gene mutation status in CRC^c and RC^d.

Modeling variables	Training set			Validation set		
	Model, n	Sen ^e (95% CI)	Spe ^f (95% CI)	Model, n	Sen (95% CI)	Spe (95% CI)
CT ^g	2	0.90-0.98	0.40-0.64	1	0.89	0.45
Pathology	2	0.75-0.91	0.66-0.88	1	0.75	0.90
PET ^h /CT	1	0.81	0.69	1	0.79	0.55
MRI ⁱ	— ^j	—	—	1	0.85	0.80
Overall	5	0.89 (0.77-0.96)	0.67 (0.50-0.80)	4	0.83 (0.73-0.90)	0.70 (0.51-0.84)

^aML: machine learning.
^bKRAS: Kirsten rat sarcoma.
^cCRC: colorectal cancer.
^dRC: rectal cancer.
^eSen: sensitivity.
^fSpe: specificity.
^gCT: computed tomography.
^hPET: positron emission tomography.
ⁱMRI: magnetic resonance imaging.
^jNot applicable.

Discussion

Principal Findings

Our study proves that the KRAS mutation prediction model for CRC, built on ML, predominantly relies on radiomics and pathology, exhibiting high overall accuracy. Among these, the CT-based radiomics model shows a c-index of 0.87 (95% CI 0.84-0.90), the MRI-based model has a c-index of 0.77 (95% CI 0.71-0.83), while the pathology-based model achieves the highest c-index of 0.94 (95% CI 0.91-0.96). Notably, the DL model significantly outperforms in terms of accuracy. In DL-based models, the MRI-based model has a c-index of 0.93 (95% CI 0.90-0.96), with a sensitivity of 0.85 (95% CI 0.75-0.91) and specificity of 0.83 (95% CI 0.77-0.88). The pathology-based DL model performs exceptionally well, with a c-index of 0.96 (95% CI 0.94-0.98), sensitivity of 0.92 (95% CI 0.83-0.97), and specificity of 0.88 (95% CI 0.71-0.96).

Previous reviews have highlighted certain advantages of DL-based KRAS mutation prediction for CRC. In 1 study, the highest c-index for the KRAS mutation in the validation set using ML was 0.58, while traditional pathological feature extraction methods face technical limitations. Additionally, the ambiguity surrounding gene mutation definitions restricts the training performance of AI models, leading to significant interpretative limitations of the results [69]. Conversely, Jia et al [70], based on radiomics, found the combined sensitivity, specificity, and c-index for the validation cohort (13 studies) to be 0.78 (95% CI 0.71-0.84), 0.84 (95% CI 0.74-0.90), and 0.86 (95% CI 0.83-0.89). Subgroup analysis based on imaging modality and segmentation methods revealed a c-index of 0.84 (95% CI 0.80-0.87), though the small sample size restricts the interpretation of the foregoing results. Furthermore, researchers have found that the co-occurrence of multiple gene mutations (eg, KRAS, NRAS, and BRAF) may further impact patient prognosis, yet relevant studies remain scarce and warrant further data support [71]. Overall, the studies included in the analysis

featured small sample sizes, considerable variations in the sources of radiomics images and ML methods, and insufficient systematic subgroup analyses, which may influence the interpretation and generalizability of our findings. Given these limitations, our study further explores different modeling variables and ML model types to provide a more comprehensive description of the strengths of various modeling approaches, offering evidence to enhance the clinical application value of KRAS mutation prediction models in CRC.

It was noted that the included literature mainly focused on the application of radiomics and pathological slides. In radiomics, common diagnosis methods include CT colonography, MRI, and PET/CT. Clinically, CT colonography is a safe and noninvasive examination with a sensitivity of up to 93.8%, capable of accurately diagnosing adenomatous polyps (≥10 mm) and excelling in gastrointestinal anatomical assessment, preoperative staging, and the diagnosis of liver, lung, and abdominal cavity lesions, particularly in identifying calcified metastases [72,73]. In contrast, although MRI is the gold standard for randomized controlled trial staging and performs excellently in assessing liver metastasis, its high cost, longer examination time, and dependence on the operator’s technical proficiency limit its extensive application [73,74]. PET/CT, primarily used for preoperative assessment of complex cases or advanced patients, is not suitable as a routine screening tool owing to its high cost [74]. In general, CT colonography is important in CRC diagnosis and staging due to its efficiency, safety, and multifaceted capabilities. Our study found that CT radiomics is the predominant technology in the included studies, performing well in diagnosing KRAS mutations. However, it was also noted that current research in CT radiomics has not explored the application of DL. Future studies should focus on expanding sample sizes and investigating DL applications in CT imaging to enhance the precision and clinical value of KRAS mutation diagnosis.

Traditional ML relies on manually designed features, which often results in poor adaptability and the potential

loss of important information. In contrast, DL uses convolutional neural networks for automatic extraction of image features, thereby retaining more information and enhancing both accuracy and robustness [75,76]. Furthermore, DL is well-suited for high-dimensional, nonlinear problems due to its reliance on large datasets and highly parameterized structures. When integrated with prior knowledge and data optimization, it can further improve accuracy and reliability [77]. In our study, DL is predominantly applied to the diagnosis of MRI and pathology-related specimens. Among these, MRI models based on DL demonstrated high diagnostic performance. Pathology models exhibited even higher c-index values. Additionally, in traditional ML, ensemble learning and boosting algorithms exhibited favorable performance on the validation set as well. However, the foregoing results are based on a limited sample size, which presents certain limitations. Therefore, future research should incorporate more image data to further optimize DL models and develop more advanced diagnosis tools.

In CRC diagnosis and treatment, mutations in key genetic loci such as NRAS and BRAF, in addition to KRAS mutations, are also of significant concern. Although NRAS mutations occur at a relatively low frequency in CRC, approximately 3%-7%, they are highly valuable in comprehensive RAS gene testing before epidermal growth factor receptor-targeted therapy [78]. BRAF mutations not only significantly influence the effects of anti-epidermal growth factor receptor monoclonal antibody treatment but are also closely associated with tumor invasiveness and metastasis, making them crucial indicators for evaluating disease progression and treatment prognosis [79]. The study of mutations in these genes holds considerable potential for future exploration. In our research, some studies have combined KRAS mutations with other gene mutations such as NRAS and BRAF for overall gene mutation prediction analysis. Upon evaluation, these studies showed a certain degree of diagnostic accuracy. However, these studies did not strictly distinguish between specific mutated genotypes, which may reduce the credibility and clinical applicability of our results. Therefore, future research should further strengthen the application of ML in diagnosing mutations at various genetic loci, with a clear distinction of genotypes, to increase the scientific validity and clinical relevance of the models. This will provide more reliable evidence for early diagnosis, precise prevention, and customized CRC treatment strategies.

During the application of models in clinical practice, several challenges must be acknowledged. First, there is a challenge related to the source of the population. A significant number of studies to date are single-center studies, which introduce challenges in the generalizability and interpretation of the models constructed [80]. Second, the selection of modeling variables presents another challenge. In existing studies, the variables are typically categorized into clinical features and radiomics features. Clinical features may involve sensitive personal information during extraction, leading to concerns regarding the authenticity of clinical variables [81]. Concurrently, studies based on radiomics features face

challenges in image segmentation, extraction, and filtering [82]. The segmentation process is often influenced by the researchers' prior knowledge and experience, which may result in substantial variations in the segmented areas [83,84]. Additionally, radiomics-based research is reliant on image parameters, with differences in parameters between images that could affect image quality and, consequently, the generalizability of the constructed models [85,86]. Furthermore, during texture feature selection, there is a significant risk of information loss from the images [87], which poses challenges to the accuracy of the resulting models. It was also noted that the generation of validation sets in model construction often relies on internal validation from a single center, which significantly limits the generalizability of the models [88]. Finally, the sample size should be considered in model construction, which also reflects a significant challenge. Many studies had limited cases, which led to validation sets being generated primarily through cross-validation, or the absence of an independent validation set, thus constraining model interpretation. Therefore, the inclusion of more cases is suggested to address these challenges.

In the studies included in our review, the overall risk of bias, as assessed by PROBAST, appears to be relatively high. Moreover, the use of the RQS tool for evaluating radiomics studies still presents challenges. This is because both PROBAST and RQS are stringent assessment tools. While PROBAST is suitable for multivariable models used in diagnosis and prediction, it is often applied to retrospective case-control studies, with few prospective studies included [88]. As such, PROBAST's evaluation results often indicate a high risk of bias, which is a significant challenge for diagnostic models. In our study, most of the included research pertains to diagnostic models, where bias may arise during case selection. Furthermore, the statistical analysis section mandates an EPV greater than 20, an independent validation set, and a validation set sample size greater than 100 for low bias. Many studies struggle to meet these stringent criteria, leading to a higher risk of bias in the statistical analysis [89]. Additionally, the statistical analysis should report whether the weights presented in the model align with those reported in practice [89]. This is particularly challenging for models with lower interpretability, such as neural networks, support vector machines, and XGBoost (Extreme Gradient Boosting), which often do not provide weights in the original validation set, making it difficult to assess consistency. These issues reflect that, while PROBAST is widely adopted to evaluate the risk of bias in ML, its criteria may be overly stringent for diagnostic models. Future research should consider updating the PROBAST tool accordingly [90].

Regarding the RQS assessment tool, it requires repeated experiments on images from different scanners, accounting for variations not only between different models of scanners but also between devices of the same model but from different manufacturers. Furthermore, repeated measurements at different time points are required. This is a stringent criterion, and it is difficult to accurately and effectively present such variations during the analysis, making it impossible to score these entries. Additionally, the tool

requires prospective registration, with multiple iterations of downscaling, where successful registration adds 7 points, and downscaling adds 3 points, both contributing substantial weight. However, in practice, effective prospective registration is challenging to implement. For validation, it is difficult to conduct robust validation, and some studies may overlook this aspect. For these studies, failure to perform any form of validation results in a deduction of 7 points. From these perspectives, it is evident that RQS is a stringent scoring tool. Consequently, models evaluated using this tool face challenges in terms of quality. Moreover, many studies evaluating radiomics models with the RQS tool report relatively low scores [91,92].

Advantages and Limitations of Our Study

The accuracy of KRAS mutation diagnosis via ML was comprehensively summarized across various modeling approaches. Our study also reveals that, currently, DL techniques have only been applied in the literature related to MRI and histopathology, while research in CT radiomics remains insufficient. This finding suggests that future research should further investigate the application of DL in CT radiomics to enhance the accuracy and clinical applicability of genetic testing, which holds significant research value and practical implications.

However, several limitations exist in this study. First, although subgroup analyses were conducted based on different modeling approaches, the limited number of eligible studies and a paucity of clear reporting of modeling details

regarding training and validation sets in some studies led to an imbalance in the statistical representation of these sets. Second, the validation sets in the included studies were primarily generated through random sampling, which may impose certain limitations in interpreting model generalizability and performance. Furthermore, our design incorporated a comprehensive set of diagnosis variables for model construction. Therefore, during the registration process, we used the widely accepted PROBAST tool to assess the bias risk of the diagnosis model. However, throughout the research process, a significant number of radiomics models developed based on imaging data were identified. To address this, the RQS was used to assess the quality of the relevant literature, which introduces some discrepancies from the initial registration content. Furthermore, the encompassed studies in our analysis had small sample sizes, which possibly affected the representativeness of the results. Therefore, future research should incorporate larger datasets to ameliorate the stability and accuracy of the models.

Conclusions

ML demonstrates ideal accuracy in diagnosing KRAS mutations in CRC, particularly DL models based on MRI and histopathological images. However, these conclusions are drawn from limited primary data. Future research should involve larger sample sizes and further optimize the development of more intelligent DL models to provide more precise and efficient tools for the intelligent diagnosis of CRC.

Acknowledgments

This research was supported by the Scientific Research Project of Health Planning and Health Care Commission of Pudong New Area, Shanghai (PW2022D-14), Shanghai Key Discipline Construction Project of Traditional Chinese Medicine (Clinical Category; shzyyzdxk-2024111), and the Project of Pudong New Area Health Commission: Construction of Famous TCM Workshop in Pudong New Area (PDZY-2025-0725).

Data Availability

All data generated or analyzed during this study are included in this published paper and its supplementary information files.

Authors' Contributions

All authors contributed to this study's conception and design. KC and YQ did the writing of the original draft preparation. KC, YQ, and DZ also did the review and editing of the writing, conceptualization, and methodology. KC, YQ, DZ, and YH handled the formal analysis and investigation. DZ and YH worked on the funding acquisition. KC, YQ, DZ, YH, YL, and HG assisted with the resources. DZ and YH performed on the supervision. All authors commented on previous versions of this paper. All authors read and approved the final paper.

Conflicts of Interest

None declared.

Multimedia Appendix 1

Literature search strategy.

[\[DOCX File \(Microsoft Word File\), 25 KB-Multimedia Appendix 1\]](#)

Multimedia Appendix 2

Basic characteristics of included studies.

[\[DOCX File \(Microsoft Word File\), 32 KB-Multimedia Appendix 2\]](#)

Multimedia Appendix 3

Radiomics quality score of the included models.

[\[DOCX File \(Microsoft Word File\), 28 KB-Multimedia Appendix 3\]](#)

Multimedia Appendix 4

Additional figures file.

[\[DOCX File \(Microsoft Word File\), 12793 KB-Multimedia Appendix 4\]](#)

Checklist 1

PRISMA checklist.

[\[DOCX File \(Microsoft Word File\), 31 KB-Checklist 1\]](#)

References

1. Morgan E, Arnold M, Gini A, et al. Global burden of colorectal cancer in 2020 and 2040: incidence and mortality estimates from GLOBOCAN. *Gut*. Feb 2023;72(2):338-344. [doi: [10.1136/gutjnl-2022-327736](#)] [Medline: [36604116](#)]
2. Siegel RL, Wagle NS, Cercek A, Smith RA, Jemal A. Colorectal cancer statistics, 2023. *CA Cancer J Clin*. 2023;73(3):233-254. [doi: [10.3322/caac.21772](#)] [Medline: [36856579](#)]
3. Eng C, Yoshino T, Ruíz-García E, et al. Colorectal cancer. *Lancet*. Jul 20, 2024;404(10449):294-310. [doi: [10.1016/S0140-6736\(24\)00360-X](#)] [Medline: [38909621](#)]
4. Derks ME, Te Groen M, Peters CP, et al. Endoscopic and surgical treatment outcomes of colitis-associated advanced colorectal neoplasia: a multicenter cohort study. *Int J Surg*. Jul 1, 2023;109(7):1961-1969. [doi: [10.1097/JS9.000000000000335](#)] [Medline: [37300890](#)]
5. Hajjar R, Olierio M, Fragoso G, et al. Modulating gut microbiota prevents anastomotic leak to reduce local implantation and dissemination of colorectal cancer cells after surgery. *Clin Cancer Res*. Feb 1, 2024;30(3):616-628. [doi: [10.1158/1078-0432.CCR-23-1601](#)] [Medline: [38010363](#)]
6. Brockmueller A, Buhrmann C, Moravejolahkami AR, Shakibaei M. Resveratrol and p53: how are they involved in CRC plasticity and apoptosis? *J Adv Res*. Dec 2024;66(181-95):181-195. [doi: [10.1016/j.jare.2024.01.005](#)] [Medline: [38190940](#)]
7. Pelka K, Hofree M, Chen JH, et al. Spatially organized multicellular immune hubs in human colorectal cancer. *Cell*. Sep 2, 2021;184(18):4734-4752. [doi: [10.1016/j.cell.2021.08.003](#)] [Medline: [34450029](#)]
8. Luo M, Wang X, Wu S, et al. A20 promotes colorectal cancer immune evasion by upregulating STC1 expression to block "eat-me" signal. *Signal Transduct Target Ther*. Aug 23, 2023;8(1):312. [doi: [10.1038/s41392-023-01545-x](#)] [Medline: [37607946](#)]
9. Mármol I, Sánchez-de-Diego C, Dieste AP, Cerrada E, Rodríguez Yoldi MJ. Colorectal carcinoma: a general overview and future perspectives in colorectal cancer. *Int J Mol Sci*. Jan 19, 2017;18(1):197. [doi: [10.3390/ijms18010197](#)] [Medline: [28106826](#)]
10. Zhu G, Pei L, Xia H, Tang Q, Bi F. Role of oncogenic KRAS in the prognosis, diagnosis and treatment of colorectal cancer. *Mol Cancer*. Nov 6, 2021;20(1):143. [doi: [10.1186/s12943-021-01441-4](#)] [Medline: [34742312](#)]
11. Biller LH, Schrag D. Diagnosis and treatment of metastatic colorectal cancer: a review. *JAMA*. Feb 16, 2021;325(7):669-685. [doi: [10.1001/jama.2021.0106](#)] [Medline: [33591350](#)]
12. Ohta R, Yamada T, Sonoda H, et al. Detection of KRAS mutations in circulating tumour DNA from plasma and urine of patients with colorectal cancer. *Eur J Surg Oncol*. Dec 2021;47(12):3151-3156. [doi: [10.1016/j.ejso.2021.07.017](#)] [Medline: [34315643](#)]
13. Alcaide M, Cheung M, Bushell K, et al. A novel multiplex droplet digital PCR assay to identify and quantify KRAS mutations in clinical specimens. *J Mol Diagn*. Mar 2019;21(2):214-227. [doi: [10.1016/j.jmoldx.2018.09.007](#)] [Medline: [30472330](#)]
14. Ngiam KY, Khor IW. Big data and machine learning algorithms for health-care delivery. *Lancet Oncol*. May 2019;20(5):e262-e273. [doi: [10.1016/S1470-2045\(19\)30149-4](#)] [Medline: [31044724](#)]
15. Reel PS, Reel S, Pearson E, Trucco E, Jefferson E. Using machine learning approaches for multi-omics data analysis: a review. *Biotechnol Adv*. 2021;49:107739. [doi: [10.1016/j.biotechadv.2021.107739](#)] [Medline: [33794304](#)]
16. Skrede OJ, De Raedt S, Kleppe A, et al. Deep learning for prediction of colorectal cancer outcome: a discovery and validation study. *Lancet*. Feb 1, 2020;395(10221):350-360. [doi: [10.1016/S0140-6736\(19\)32998-8](#)] [Medline: [32007170](#)]
17. Saba T. Recent advancement in cancer detection using machine learning: systematic survey of decades, comparisons and challenges. *J Infect Public Health*. Sep 2020;13(9):1274-1289. [doi: [10.1016/j.jiph.2020.06.033](#)] [Medline: [32758393](#)]
18. Myszczyńska MA, Ojamies PN, Lacoste AMB, et al. Applications of machine learning to diagnosis and treatment of neurodegenerative diseases. *Nat Rev Neurol*. Aug 2020;16(8):440-456. [doi: [10.1038/s41582-020-0377-8](#)] [Medline: [32669685](#)]

19. Tran KA, Kondrashova O, Bradley A, Williams ED, Pearson JV, Waddell N. Deep learning in cancer diagnosis, prognosis and treatment selection. *Genome Med.* Sep 27, 2021;13(1):152. [doi: [10.1186/s13073-021-00968-x](https://doi.org/10.1186/s13073-021-00968-x)] [Medline: [34579788](https://pubmed.ncbi.nlm.nih.gov/34579788/)]
20. Yu G, Sun K, Xu C, et al. Accurate recognition of colorectal cancer with semi-supervised deep learning on pathological images. *Nat Commun.* Nov 2, 2021;12(1):6311. [doi: [10.1038/s41467-021-26643-8](https://doi.org/10.1038/s41467-021-26643-8)] [Medline: [34728629](https://pubmed.ncbi.nlm.nih.gov/34728629/)]
21. Kiehl L, Kuntz S, Höhn J, et al. Deep learning can predict lymph node status directly from histology in colorectal cancer. *Eur J Cancer.* Nov 2021;157:464-473. [doi: [10.1016/j.ejca.2021.08.039](https://doi.org/10.1016/j.ejca.2021.08.039)] [Medline: [34649117](https://pubmed.ncbi.nlm.nih.gov/34649117/)]
22. PROSPERO home page. National Institute for Health and Care Research. URL: <https://www.crd.york.ac.uk/PROSPERO/> [Accessed 2025-06-03]
23. Wolff RF, Moons KGM, Riley RD, et al. PROBAST: a tool to assess the risk of bias and applicability of prediction model studies. *Ann Intern Med.* Jan 1, 2019;170(1):51-58. [doi: [10.7326/M18-1376](https://doi.org/10.7326/M18-1376)] [Medline: [30596875](https://pubmed.ncbi.nlm.nih.gov/30596875/)]
24. Lambin P, Leijenaar RTH, Deist TM, et al. Radiomics: the bridge between medical imaging and personalized medicine. *Nat Rev Clin Oncol.* Dec 2017;14(12):749-762. [doi: [10.1038/nrclinonc.2017.141](https://doi.org/10.1038/nrclinonc.2017.141)] [Medline: [28975929](https://pubmed.ncbi.nlm.nih.gov/28975929/)]
25. Debray TP, Damen JA, Riley RD, et al. A framework for meta-analysis of prediction model studies with binary and time-to-event outcomes. *Stat Methods Med Res.* Sep 2019;28(9):2768-2786. [doi: [10.1177/0962280218785504](https://doi.org/10.1177/0962280218785504)] [Medline: [30032705](https://pubmed.ncbi.nlm.nih.gov/30032705/)]
26. Ma Y, Guo Y, Cui W, et al. SG-Transunet: a segmentation-guided Transformer U-Net model for KRAS gene mutation status identification in colorectal cancer. *Comput Biol Med.* May 2024;173:108293. [doi: [10.1016/j.compbiomed.2024.108293](https://doi.org/10.1016/j.compbiomed.2024.108293)] [Medline: [38574528](https://pubmed.ncbi.nlm.nih.gov/38574528/)]
27. Li M, Yuan Y, Zhou H, Feng F, Xu G. A multicenter study: predicting KRAS mutation and prognosis in colorectal cancer through a CT-based radiomics nomogram. *Abdom Radiol (NY).* Jun 2024;49(6):1816-1828. [doi: [10.1007/s00261-024-04218-7](https://doi.org/10.1007/s00261-024-04218-7)] [Medline: [38393357](https://pubmed.ncbi.nlm.nih.gov/38393357/)]
28. Huang Z, Huang X, Huang Y, et al. Identification of KRAS mutation-associated gut microbiota in colorectal cancer and construction of predictive machine learning model. *Microbiol Spectr.* May 2, 2024;12(5):e0272023. [doi: [10.1128/spectrum.02720-23](https://doi.org/10.1128/spectrum.02720-23)] [Medline: [38572984](https://pubmed.ncbi.nlm.nih.gov/38572984/)]
29. Cai M, Zhao L, Qiang Y, Wang L, Zhao J. CHNet: a multi-task global-local collaborative hybrid network for KRAS mutation status prediction in colorectal cancer. *Artif Intell Med.* Sep 2024;155:102931. [doi: [10.1016/j.artmed.2024.102931](https://doi.org/10.1016/j.artmed.2024.102931)] [Medline: [39094228](https://pubmed.ncbi.nlm.nih.gov/39094228/)]
30. Zhao H, Su Y, Wang Y, et al. Using tumor habitat-derived radiomic analysis during pretreatment ¹⁸F-FDG PET for predicting KRAS/NRAS/BRAF mutations in colorectal cancer. *Cancer Imaging.* Feb 12, 2024;24(1):26. [doi: [10.1186/s40644-024-00670-2](https://doi.org/10.1186/s40644-024-00670-2)] [Medline: [38342905](https://pubmed.ncbi.nlm.nih.gov/38342905/)]
31. Li XJ, Chi XD, Huang PJ, Liang Q, Liu JP. Deep neural network for the prediction of KRAS, NRAS, and BRAF genotypes in left-sided colorectal cancer based on histopathologic images. *Comput Med Imaging Graph.* Jul 2024;115:102384. [doi: [10.1016/j.compmedimag.2024.102384](https://doi.org/10.1016/j.compmedimag.2024.102384)]
32. Wesdorp N, Zeeuw M, van der Meulen D, et al. Identifying genetic mutation status in patients with colorectal cancer liver metastases using radiomics-based machine-learning models. *Cancers (Basel).* Nov 29, 2023;15(23):5648. [doi: [10.3390/cancers15235648](https://doi.org/10.3390/cancers15235648)] [Medline: [38067353](https://pubmed.ncbi.nlm.nih.gov/38067353/)]
33. Cao Y, Zhang J, Huang L, et al. Construction of prediction model for KRAS mutation status of colorectal cancer based on CT radiomics. *Jpn J Radiol.* Nov 2023;41(11):1236-1246. [doi: [10.1007/s11604-023-01458-3](https://doi.org/10.1007/s11604-023-01458-3)] [Medline: [37311935](https://pubmed.ncbi.nlm.nih.gov/37311935/)]
34. Alshuhri MS, Alduhayim A, Al-Mubarak H, et al. Investigating the feasibility of predicting KRAS status, tumor staging, and extramural venous invasion in colorectal cancer using inter-platform magnetic resonance imaging radiomic features. *Diagnostics (Basel).* Nov 27, 2023;13(23):3541. [doi: [10.3390/diagnostics13233541](https://doi.org/10.3390/diagnostics13233541)] [Medline: [38066782](https://pubmed.ncbi.nlm.nih.gov/38066782/)]
35. Zhao L, Song K, Ma Y, et al. A segmentation-based sequence residual attention model for KRAS gene mutation status prediction in colorectal cancer. *Appl Intell.* May 2023;53(9):10232-10254. [doi: [10.1007/s10489-022-04011-3](https://doi.org/10.1007/s10489-022-04011-3)]
36. Xiang Y, Li S, Song M, et al. KRAS status predicted by pretreatment MRI radiomics was associated with lung metastasis in locally advanced rectal cancer patients. *BMC Med Imaging.* Dec 12, 2023;23(1):210. [doi: [10.1186/s12880-023-01173-5](https://doi.org/10.1186/s12880-023-01173-5)] [Medline: [38087207](https://pubmed.ncbi.nlm.nih.gov/38087207/)]
37. Wu KC, Chen SW, Hsieh TC, et al. Imaging prediction of KRAS mutation in patients with rectal cancer through deep metric learning using pretreatment [¹⁸F]Fluorodeoxyglucose positron emission tomography/computed tomography. *Br J Radiol.* Nov 2023;96(1151):20230243. [doi: [10.1259/bjr.20230243](https://doi.org/10.1259/bjr.20230243)] [Medline: [37750945](https://pubmed.ncbi.nlm.nih.gov/37750945/)]
38. Lara MAR, Esposito MI, Aineseder M, et al. Radiomics and machine learning for prediction of two-year disease-specific mortality and KRAS mutation status in metastatic colorectal cancer. *Surg Oncol.* Dec 2023;51:101986. [doi: [10.1016/j.suronc.2023.101986](https://doi.org/10.1016/j.suronc.2023.101986)] [Medline: [37729816](https://pubmed.ncbi.nlm.nih.gov/37729816/)]
39. Porto-Álvarez J, Cernadas E, Aldaz Martínez RA, et al. CT-based radiomics to predict KRAS mutation in CRC patients using a machine learning algorithm: a retrospective study. *Biomedicines.* Jul 29, 2023;11(8):2144. [doi: [10.3390/biomedicines11082144](https://doi.org/10.3390/biomedicines11082144)] [Medline: [37626641](https://pubmed.ncbi.nlm.nih.gov/37626641/)]

40. Heckenauer R, Weber J, Wemmert C, et al. Comparison of deep learning architectures for colon cancer mutation detection. Presented at: 2023 IEEE 36th International Symposium on Computer-Based Medical Systems (CBMS); Jun 22-24, 2023; L'Aquila, Italy. [doi: [10.1109/CBMS58004.2023.00244](https://doi.org/10.1109/CBMS58004.2023.00244)]
41. Hu J, Xia X, Wang P, et al. Predicting Kirsten rat sarcoma virus gene mutation status in patients with colorectal cancer by radiomics models based on multiphasic CT. *Front Oncol*. 2022;12:848798. [doi: [10.3389/fonc.2022.848798](https://doi.org/10.3389/fonc.2022.848798)] [Medline: [35814386](https://pubmed.ncbi.nlm.nih.gov/35814386/)]
42. Guo Y, Lyu T, Liu S, et al. Learn to estimate genetic mutation and microsatellite instability with histopathology H&E slides in colon carcinoma. *Cancers (Basel)*. Aug 27, 2022;14(17):4144. [doi: [10.3390/cancers14174144](https://doi.org/10.3390/cancers14174144)] [Medline: [36077681](https://pubmed.ncbi.nlm.nih.gov/36077681/)]
43. Xue T, Peng H, Chen QL, Li MM, Duan SF, Feng F. Preoperative prediction of KRAS mutation status in colorectal cancer using a CT-based radiomics nomogram. *Br J Radiol*. Jun 1, 2022;95(1134):20211014. [doi: [10.1259/bjr.20211014](https://doi.org/10.1259/bjr.20211014)] [Medline: [35312376](https://pubmed.ncbi.nlm.nih.gov/35312376/)]
44. Liu H, Yin H, Li J, et al. A deep learning model based on MRI and clinical factors facilitates noninvasive evaluation of KRAS mutation in rectal cancer. *J Magn Reson Imaging*. Dec 2022;56(6):1659-1668. [doi: [10.1002/jmri.28237](https://doi.org/10.1002/jmri.28237)] [Medline: [35587946](https://pubmed.ncbi.nlm.nih.gov/35587946/)]
45. Ghareeb WM, Draz E, Madbouly K, et al. Deep neural network for the prediction of KRAS genotype in rectal cancer. *J Am Coll Surg*. Sep 1, 2022;235(3):482-493. [doi: [10.1097/XCS.000000000000277](https://doi.org/10.1097/XCS.000000000000277)] [Medline: [35972169](https://pubmed.ncbi.nlm.nih.gov/35972169/)]
46. Crimi F, Zanon C, Cabrelle G, et al. Contrast-enhanced CT texture analysis in colon cancer: correlation with genetic markers. *Tomography*. Aug 31, 2022;8(5):2193-2201. [doi: [10.3390/tomography8050184](https://doi.org/10.3390/tomography8050184)] [Medline: [36136880](https://pubmed.ncbi.nlm.nih.gov/36136880/)]
47. Ma Y, Wang J, Song K, Qiang Y, Jiao X, Zhao J. Spatial-frequency dual-branch attention model for determining KRAS mutation status in colorectal cancer with T2-weighted MRI. *Comput Methods Programs Biomed*. Sep 2021;209:106311. [doi: [10.1016/j.cmpb.2021.106311](https://doi.org/10.1016/j.cmpb.2021.106311)] [Medline: [34352652](https://pubmed.ncbi.nlm.nih.gov/34352652/)]
48. Bian C, Wang Y, Lu Z, et al. ImmunoAizer: a deep learning-based computational framework to characterize cell distribution and gene mutation in tumor microenvironment. *Cancers (Basel)*. Apr 1, 2021;13(7):1659. [doi: [10.3390/cancers13071659](https://doi.org/10.3390/cancers13071659)] [Medline: [33916145](https://pubmed.ncbi.nlm.nih.gov/33916145/)]
49. Song K, Zhao Z, Ma Y, et al. A multitask dual-stream attention network for the identification of KRAS mutation in colorectal cancer. *Med Phys Mex Symp Med Phys*. Jan 2022;49(1):254-270. [doi: [10.1002/mp.15361](https://doi.org/10.1002/mp.15361)]
50. Zhang Z, Shen L, Wang Y, et al. MRI radiomics signature as a potential biomarker for predicting KRAS status in locally advanced rectal cancer patients. *Front Oncol*. 2021;11:614052. [doi: [10.3389/fonc.2021.614052](https://doi.org/10.3389/fonc.2021.614052)] [Medline: [34026605](https://pubmed.ncbi.nlm.nih.gov/34026605/)]
51. Zhang G, Chen L, Liu A, et al. Comparable performance of deep learning-based to manual-based tumor segmentation in KRAS/NRAS/BRAF mutation prediction with MR-based radiomics in rectal cancer. *Front Oncol*. 2021;11:696706. [doi: [10.3389/fonc.2021.696706](https://doi.org/10.3389/fonc.2021.696706)] [Medline: [34395262](https://pubmed.ncbi.nlm.nih.gov/34395262/)]
52. Jang BS, Song C, Kang SB, Kim JS. Radiogenomic and deep learning network approaches to predict KRAS mutation from radiotherapy plan CT. *Anticancer Res*. Aug 2021;41(8):3969-3976. [doi: [10.21873/anticancer.15193](https://doi.org/10.21873/anticancer.15193)] [Medline: [34281860](https://pubmed.ncbi.nlm.nih.gov/34281860/)]
53. Sanchez-Ibarra HE, Jiang X, Gallegos-Gonzalez EY, et al. KRAS, NRAS, and BRAF mutation prevalence, clinicopathological association, and their application in a predictive model in Mexican patients with metastatic colorectal cancer: a retrospective cohort study. *PLoS ONE*. 2020;15(7):e0235490. [doi: [10.1371/journal.pone.0235490](https://doi.org/10.1371/journal.pone.0235490)] [Medline: [32628708](https://pubmed.ncbi.nlm.nih.gov/32628708/)]
54. Oh JE, Kim MJ, Lee J, et al. Magnetic resonance-based texture analysis differentiating KRAS mutation status in rectal cancer. *Cancer Res Treat*. Jan 2020;52(1):51-59. [doi: [10.4143/crt.2019.050](https://doi.org/10.4143/crt.2019.050)] [Medline: [31096736](https://pubmed.ncbi.nlm.nih.gov/31096736/)]
55. Jang HJ, Lee A, Kang J, Song IH, Lee SH. Prediction of clinically actionable genetic alterations from colorectal cancer histopathology images using deep learning. *World J Gastroenterol*. Oct 28, 2020;26(40):6207-6223. [doi: [10.3748/wjg.v26.i40.6207](https://doi.org/10.3748/wjg.v26.i40.6207)] [Medline: [33177794](https://pubmed.ncbi.nlm.nih.gov/33177794/)]
56. Guo XF, Yang WQ, Yang Q, et al. Feasibility of MRI radiomics for predicting KRAS mutation in rectal cancer. *Curr Med Sci*. Dec 2020;40(6):1156-1160. [doi: [10.1007/s11596-020-2298-6](https://doi.org/10.1007/s11596-020-2298-6)] [Medline: [33428144](https://pubmed.ncbi.nlm.nih.gov/33428144/)]
57. Cui Y, Liu H, Ren J, et al. Development and validation of a MRI-based radiomics signature for prediction of KRAS mutation in rectal cancer. *Eur Radiol*. Apr 2020;30(4):1948-1958. [doi: [10.1007/s00330-019-06572-3](https://doi.org/10.1007/s00330-019-06572-3)]
58. Li Y, Eresen A, Shangguan J, et al. Preoperative prediction of perineural invasion and KRAS mutation in colon cancer using machine learning. *J Cancer Res Clin Oncol*. Dec 2020;146(12):3165-3174. [doi: [10.1007/s00432-020-03354-z](https://doi.org/10.1007/s00432-020-03354-z)] [Medline: [32779023](https://pubmed.ncbi.nlm.nih.gov/32779023/)]
59. He K, Liu XM, Li MY, Li XY, Yang HL, Zhang HM. Noninvasive KRAS mutation estimation in colorectal cancer using a deep learning method based on CT imaging. *BMC Med Imaging*. Jun 1, 2020;20(1):59. [doi: [10.1186/s12880-020-00457-4](https://doi.org/10.1186/s12880-020-00457-4)] [Medline: [32487083](https://pubmed.ncbi.nlm.nih.gov/32487083/)]
60. Yu Z, Yu H, Zou Q, et al. Nomograms for prediction of molecular phenotypes in colorectal cancer. *Onco Targets Ther*. 2020;13:309-321. [doi: [10.2147/OTT.S234495](https://doi.org/10.2147/OTT.S234495)] [Medline: [32021277](https://pubmed.ncbi.nlm.nih.gov/32021277/)]

61. Wu X, Li Y, Chen X, et al. Deep learning features improve the performance of a radiomics signature for predicting KRAS status in patients with colorectal cancer. *Acad Radiol*. Nov 2020;27(11):e254-e262. [doi: [10.1016/j.acra.2019.12.007](https://doi.org/10.1016/j.acra.2019.12.007)] [Medline: [31982342](https://pubmed.ncbi.nlm.nih.gov/31982342/)]
62. Wang J, Cui Y, Shi G, et al. Multi-branch cross attention model for prediction of KRAS mutation in rectal cancer with t2-weighted MRI. *Appl Intell*. Aug 2020;50(8):2352-2369. [doi: [10.1007/s10489-020-01658-8](https://doi.org/10.1007/s10489-020-01658-8)] [Medline: [34764567](https://pubmed.ncbi.nlm.nih.gov/34764567/)]
63. Shi R, Chen W, Yang B, et al. Prediction of KRAS, NRAS and BRAF status in colorectal cancer patients with liver metastasis using a deep artificial neural network based on radiomics and semantic features. *Am J Cancer Res*. 2020;10(12):4513-4526. [Medline: [33415015](https://pubmed.ncbi.nlm.nih.gov/33415015/)]
64. Taguchi N, Oda S, Yokota Y, et al. CT texture analysis for the prediction of KRAS mutation status in colorectal cancer via a machine learning approach. *Eur J Radiol*. Sep 2019;118:38-43. [doi: [10.1016/j.ejrad.2019.06.028](https://doi.org/10.1016/j.ejrad.2019.06.028)] [Medline: [31439256](https://pubmed.ncbi.nlm.nih.gov/31439256/)]
65. Meng X, Xia W, Xie P, et al. Preoperative radiomic signature based on multiparametric magnetic resonance imaging for noninvasive evaluation of biological characteristics in rectal cancer. *Eur Radiol*. Jun 2019;29(6):3200-3209. [doi: [10.1007/s00330-018-5763-x](https://doi.org/10.1007/s00330-018-5763-x)] [Medline: [30413959](https://pubmed.ncbi.nlm.nih.gov/30413959/)]
66. Yang L, Dong D, Fang M, et al. Can CT-based radiomics signature predict KRAS/NRAS/BRAF mutations in colorectal cancer? *Eur Radiol*. May 2018;28(5):2058-2067. [doi: [10.1007/s00330-017-5146-8](https://doi.org/10.1007/s00330-017-5146-8)] [Medline: [29335867](https://pubmed.ncbi.nlm.nih.gov/29335867/)]
67. Chen SW, Shen WC, Chen WTL, et al. Metabolic imaging phenotype using radiomics of [¹⁸F]FDG PET/CT associated with genetic alterations of colorectal cancer. *Mol Imaging Biol*. Feb 2019;21(1):183-190. [doi: [10.1007/s11307-018-1225-8](https://doi.org/10.1007/s11307-018-1225-8)] [Medline: [29948642](https://pubmed.ncbi.nlm.nih.gov/29948642/)]
68. Pershad Y, Govindan S, Hara AK, et al. Using Naïve Bayesian analysis to determine imaging characteristics of KRAS mutations in metastatic colon cancer. *Diagnostics (Basel)*. Sep 2, 2017;7(3):50. [doi: [10.3390/diagnostics7030050](https://doi.org/10.3390/diagnostics7030050)] [Medline: [28869500](https://pubmed.ncbi.nlm.nih.gov/28869500/)]
69. Guitton T, Allaume P, Rabilloud N, et al. Artificial Intelligence in predicting microsatellite instability and KRAS, BRAF mutations from whole-slide images in colorectal cancer: a systematic review. *Diagnostics (Basel)*. Dec 31, 2023;14(1):38201408. [doi: [10.3390/diagnostics14010099](https://doi.org/10.3390/diagnostics14010099)] [Medline: [38201408](https://pubmed.ncbi.nlm.nih.gov/38201408/)]
70. Jia LL, Zhao JX, Zhao LP, Tian JH, Huang G. Current status and quality of radiomic studies for predicting KRAS mutations in colorectal cancer patients: a systematic review and meta-analysis. *Eur J Radiol*. Jan 2023;158(110640):110640. [doi: [10.1016/j.ejrad.2022.110640](https://doi.org/10.1016/j.ejrad.2022.110640)] [Medline: [36525703](https://pubmed.ncbi.nlm.nih.gov/36525703/)]
71. De Mattia E, Polesel J, Mezzalana S, et al. Predictive and prognostic value of oncogene mutations and microsatellite instability in locally-advanced rectal cancer treated with neoadjuvant radiation-based therapy: a systematic review and meta-analysis. *Cancers (Basel)*. Feb 25, 2023;15(5):1469. [doi: [10.3390/cancers15051469](https://doi.org/10.3390/cancers15051469)] [Medline: [36900260](https://pubmed.ncbi.nlm.nih.gov/36900260/)]
72. Obaro AE, Plumb AA, Fanshawe TR, et al. Post-imaging colorectal cancer or interval cancer rates after CT colonography: a systematic review and meta-analysis. *Lancet Gastroenterol Hepatol*. May 2018;3(5):326-336. [doi: [10.1016/S2468-1253\(18\)30032-3](https://doi.org/10.1016/S2468-1253(18)30032-3)] [Medline: [29472116](https://pubmed.ncbi.nlm.nih.gov/29472116/)]
73. Kijima S, Sasaki T, Nagata K, Utano K, Lefor AT, Sugimoto H. Preoperative evaluation of colorectal cancer using CT colonography, MRI, and PET/CT. *World J Gastroenterol*. Dec 7, 2014;20(45):16964-16975. [doi: [10.3748/wjg.v20.i45.16964](https://doi.org/10.3748/wjg.v20.i45.16964)] [Medline: [25493009](https://pubmed.ncbi.nlm.nih.gov/25493009/)]
74. Liu SC, Zhang H. Early diagnostic strategies for colorectal cancer. *World J Gastroenterol*. Sep 7, 2024;30(33):3818-3822. [doi: [10.3748/wjg.v30.i33.3818](https://doi.org/10.3748/wjg.v30.i33.3818)] [Medline: [39351429](https://pubmed.ncbi.nlm.nih.gov/39351429/)]
75. Lai Y. A comparison of traditional machine learning and deep learning in image recognition. Presented at: 3rd International Conference on Electrical, Mechanical and Computer Engineering; Aug 9-11, 2019; Guizhou, China. [doi: [10.1088/1742-6596/1314/1/012148](https://doi.org/10.1088/1742-6596/1314/1/012148)]
76. Janiesch C, Zschech P, Heinrich K. Machine learning and deep learning. *Electron Mark*. Sep 2021;31(3):685-695. [doi: [10.1007/s12525-021-00475-2](https://doi.org/10.1007/s12525-021-00475-2)]
77. Shlezinger N, Eldar YC. Model-based deep learning. *FNT Signal Proc*. 2023;17(4):291-416. [doi: [10.1561/20000000113](https://doi.org/10.1561/20000000113)]
78. Prieto-Potin I, Montagut C, Bellosillo B, et al. Multicenter evaluation of the Idylla NRAS-BRAF mutation test in metastatic colorectal cancer. *J Mol Diagn*. Sep 2018;20(5):664-676. [doi: [10.1016/j.jmolx.2018.05.008](https://doi.org/10.1016/j.jmolx.2018.05.008)] [Medline: [29959022](https://pubmed.ncbi.nlm.nih.gov/29959022/)]
79. Modest DP, Ricard I, Heinemann V, et al. Outcome according to KRAS-, NRAS- and BRAF-mutation as well as KRAS mutation variants: pooled analysis of five randomized trials in metastatic colorectal cancer by the AIO colorectal cancer study group. *Ann Oncol*. Sep 2016;27(9):1746-1753. [doi: [10.1093/annonc/mdw261](https://doi.org/10.1093/annonc/mdw261)] [Medline: [27358379](https://pubmed.ncbi.nlm.nih.gov/27358379/)]
80. Samaga D, Hornung R, Braselmann H, et al. Single-center versus multi-center data sets for molecular prognostic modeling: a simulation study. *Radiat Oncol*. May 14, 2020;15(1):109. [doi: [10.1186/s13014-020-01543-1](https://doi.org/10.1186/s13014-020-01543-1)] [Medline: [32410693](https://pubmed.ncbi.nlm.nih.gov/32410693/)]
81. Carvalho T, Moniz N, Faria P, Antunes L. Towards a data privacy-predictive performance trade-off. *Expert Syst Appl*. Aug 2023;223:119785. [doi: [10.1016/j.eswa.2023.119785](https://doi.org/10.1016/j.eswa.2023.119785)]

82. Avanzo M, Wei L, Stancanella J, et al. Machine and deep learning methods for radiomics. *Med Phys*. Jun 2020;47(5):e185-e202. [doi: [10.1002/mp.13678](https://doi.org/10.1002/mp.13678)] [Medline: [32418336](https://pubmed.ncbi.nlm.nih.gov/32418336/)]
83. Zhang X, Zhang Y, Zhang G, et al. Deep learning with radiomics for disease diagnosis and treatment: challenges and potential. *Front Oncol*. 2022;12(773840):35251962. [doi: [10.3389/fonc.2022.773840](https://doi.org/10.3389/fonc.2022.773840)]
84. Castiglioni I, Rundo L, Codari M, et al. AI applications to medical images: from machine learning to deep learning. *Phys Med*. Mar 2021;83(9-24):9-24. [doi: [10.1016/j.ejmp.2021.02.006](https://doi.org/10.1016/j.ejmp.2021.02.006)] [Medline: [33662856](https://pubmed.ncbi.nlm.nih.gov/33662856/)]
85. Meyer M, Ronald J, Vernuccio F, et al. Reproducibility of CT radiomic features within the same patient: influence of radiation dose and CT reconstruction settings. *Radiology*. Dec 2019;293(3):583-591. [doi: [10.1148/radiol.2019190928](https://doi.org/10.1148/radiol.2019190928)] [Medline: [31573400](https://pubmed.ncbi.nlm.nih.gov/31573400/)]
86. Wichtmann BD, Harder FN, Weiss K, et al. Influence of image processing on radiomic features from magnetic resonance imaging. *Invest Radiol*. Mar 1, 2023;58(3):199-208. [doi: [10.1097/RLI.0000000000000921](https://doi.org/10.1097/RLI.0000000000000921)] [Medline: [36070524](https://pubmed.ncbi.nlm.nih.gov/36070524/)]
87. Chen Y, Xia R, Yang K, Zou K. MFMAM: Image inpainting via multi-scale feature module with attention module. *Comput Vis Image Underst*. Jan 2024;238:103883. [doi: [10.1016/j.cviu.2023.103883](https://doi.org/10.1016/j.cviu.2023.103883)]
88. Maleki F, Ovens K, Gupta R, Reinhold C, Spatz A, Forghani R. Generalizability of machine learning models: quantitative evaluation of three methodological pitfalls. *Radiol Artif Intell*. Jan 2023;5(1):e220028. [doi: [10.1148/ryai.220028](https://doi.org/10.1148/ryai.220028)] [Medline: [36721408](https://pubmed.ncbi.nlm.nih.gov/36721408/)]
89. Moons KGM, Wolff RF, Riley RD, et al. PROBAST: a tool to assess risk of bias and applicability of prediction model studies: explanation and elaboration. *Ann Intern Med*. Jan 1, 2019;170(1):W1-W33. [doi: [10.7326/M18-1377](https://doi.org/10.7326/M18-1377)] [Medline: [30596876](https://pubmed.ncbi.nlm.nih.gov/30596876/)]
90. Langenhuijsen LFS, Janse RJ, Venema E, et al. Systematic metareview of prediction studies demonstrates stable trends in bias and low PROBAST inter-rater agreement. *J Clin Epidemiol*. Jul 2023;159(159-73):159-173. [doi: [10.1016/j.jclinepi.2023.04.012](https://doi.org/10.1016/j.jclinepi.2023.04.012)] [Medline: [37142166](https://pubmed.ncbi.nlm.nih.gov/37142166/)]
91. Abbaspour E, Karimzadgh S, Monsef A, Joukar F, Mansour-Ghanaei F, Hassanipour S. Application of radiomics for preoperative prediction of lymph node metastasis in colorectal cancer: a systematic review and meta-analysis. *Int J Surg*. Jun 1, 2024;110(6):3795-3813. [doi: [10.1097/JS9.0000000000001239](https://doi.org/10.1097/JS9.0000000000001239)] [Medline: [38935817](https://pubmed.ncbi.nlm.nih.gov/38935817/)]
92. Russo L, Bottazzi S, Kocak B, et al. Evaluating the quality of radiomics-based studies for endometrial cancer using RQS and METRICS tools. *Eur Radiol*. Jan 2025;35(1):202-214. [doi: [10.1007/s00330-024-10947-6](https://doi.org/10.1007/s00330-024-10947-6)] [Medline: [39014086](https://pubmed.ncbi.nlm.nih.gov/39014086/)]

Abbreviations

AI: artificial intelligence
BRAF: v-raf murine sarcoma viral oncogene homolog B
c-index: concordance index
CRC: colorectal cancer
CT: computed tomography
DL: deep learning
EPV: event per variable
KRAS: Kirsten rat sarcoma
ML: machine learning
MRI: magnetic resonance imaging
NRAS: neuroblastoma ras viral oncogene homolog
PET/CT: positron emission tomography/computed tomography
PRISMA: Preferred Reporting Items for Systematic Reviews and Meta-Analyses
PROBAST: Prediction Model Risk of Bias Assessment Tool
PROSPERO: International Prospective Register of Systematic Reviews
RC: rectal cancer
RQS: radiomics quality score
XGBoost: Extreme Gradient Boosting

Edited by Javad Sarvestan; peer-reviewed by Chibuzo Onah, Leela Prasad Gorrepati, Qiang Wang, Yuvaranesh Vedaraju; submitted 06.03.2025; final revised version received 30.04.2025; accepted 15.05.2025; published 18.07.2025

Please cite as:

Chen K, Qu Y, Han Y, Li Y, Gao H, Zheng D
 Performance of Machine Learning in Diagnosing KRAS (Kirsten Rat Sarcoma) Mutations in Colorectal Cancer: Systematic Review and Meta-Analysis
J Med Internet Res 2025;27:e73528

URL: <https://www.jmir.org/2025/1/e73528>
doi: [10.2196/73528](https://doi.org/10.2196/73528)

© Kaixin Chen, Yin Qu, Ye Han, Yan Li, Huiyan Gao, De Zheng. Originally published in the Journal of Medical Internet Research (<https://www.jmir.org>), 18.07.2025. This is an open-access article distributed under the terms of the Creative Commons Attribution License (<https://creativecommons.org/licenses/by/4.0/>), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in the Journal of Medical Internet Research (ISSN 1438-8871), is properly cited. The complete bibliographic information, a link to the original publication on <https://www.jmir.org/>, as well as this copyright and license information must be included.