European Journal of Radiology Open 11 (2023) 100505

Contents lists available at ScienceDirect

European Journal of Radiology Open

journal homepage: www.elsevier.com/locate/ejro

A mutation-based radiomics signature predicts response to imatinib in Gastrointestinal Stromal Tumors (GIST)

Giovanni Cappello [a,*], Valentina Giannini [a,b], Roberto Cannella [c,d], Emanuele Tabone [a], Ilaria Ambrosini [e,l], Francesca Molea [a,b], Nicolò Damiani [a,b], Ilenia Landolfi [a,b], Giovanni Serra [b], Giorgia Porrello [c], Cecilia Gozzo [f], Lorena Incorvaia [g], Giuseppe Badalamenti [g], Giovanni Grignani [h], Alessandra Merlini [h,i], Lorenzo D'Ambrosio [i,j,l], Tommaso Vincenzo Bartolotta [c,k], Daniele Regge [a,b]

[a] Department of Radiology, Candiolo Cancer Institute, FPO-IRCCS, Strada Provinciale 142 Km 3.95, Candiolo, Turin 10060, Italy
[b] Department of Surgical Sciences, University of Turin, Turin 10124, Italy
[c] Section of Radiology - Department of Biomedicine, Neuroscience and Advanced Diagnostics (BiND), University of Palermo, Via del Vespro 129, Palermo 90127, Italy
[d] Department of Health Promotion, Mother and Child Care, Internal Medicine and Medical Specialties (PROMISE), University of Palermo, Italy
[e] Department of Translational Research, Academic Radiology, University of Pisa, Italy
[f] Department of Radiology, Humanitas, Istituto Clinico Catanese, Catania, Italy
[g] Department of Surgical, Oncological, and Oral Sciences, Section of Medical Oncology, University of Palermo, Palermo, Italy
[h] Division of Medical Oncology, Candiolo Cancer Institute, FPO-IRCCS, Strada Provinciale 142 Km 3.95, Candiolo, Turin 10060, Italy
[i] Department of Oncology, University of Turin, Regione Gonzole 10, Orbassano, Turin 10043, Italy
[j] Medical Oncology, University Hospital S. Luigi Gonzaga, regione Gonzole 10, Orbassano, Turin 10043, Italy
[k] Department of Radiology, Fondazione Istituto Giuseppe Giglio, Ct.da Pietrapollastra, Via Pisciotto, Cefalù, Palermo 90015, Italy
[l] Previously at Division of Medical Oncology, Candiolo Cancer Institute, FPO-IRCCS, Strada Provinciale 142 Km 3.95, Candiolo, Turin 10060, Italy

HIGHLIGHTS

- Only GIST with specific genotypes are sensitive to imatinib.
- Radiomics helps to correlate pre-treatment CT features to GIST mutational status, predicting response to imatinib treatment.
- The developed radiomics signature improve treatment allocation, accuracy and appropriateness.

ARTICLE INFO

Keywords:
Artificial intelligence
Radiomics
GIST
Mutational status
Response to therapy

ABSTRACT

Objectives: To develop a mutation-based radiomics signature to predict response to imatinib in Gastrointestinal Stromal Tumors (GISTs).
Methods: Eighty-two patients with GIST were enrolled in this retrospective study, including 52 patients from one center that were used to develop the model, and 30 patients from a second center to validate it. Reference standard was the mutational status of tyrosine-protein kinase (KIT) and platelet-derived growth factor α (PDGFRA). Patients were dichotomized in imatinib sensitive (group 0 - mutation in KIT or PDGFRA, different from exon 18-D842V), and imatinib non-responsive (group 1 - PDGFRA exon 18-D842V mutation or absence of mutation in KIT/PDGFRA). Initially, 107 texture features were extracted from the tumor masks of baseline computed tomography scans. Different machine learning methods were then implemented to select the best combination of features for the development of the radiomics signature.

Abbreviations: ANOVA, Analysis of Variances; CT, Computed Tomography; FS, Feature selection; GISTs, Gastrointestinal Stromal Tumors; GIT, Gastrointestinal tract; GLCM, Gray level co-occurrence matrices; GLDM, Gray level dependence matrix; GLRLM, Gray level run length matrices; GLSZM, Gray level size zone matrix; IBSI, Image biomarker standardization initiative; KIT, Tyrosine-protein kinase; kNN, k-Nearest Neighbor; LASSO, Least Absolute Shrinkage and Selection Operator; NGTDM, Neighboring gray tone difference matrix; PDGFRA, Platelet-derived growth factor-α; RF, Random Forest; SVM, Support-Vector Machine; TKI, Tyrosine kinase inhibitor; WT, Wild-Type.

* Corresponding author.
E-mail address: giovanni.cappello@ircc.it (G. Cappello).

https://doi.org/10.1016/j.ejro.2023.100505

Results: The best performance was obtained with the 5 features selected by the ANOVA model and the Bayes classifier, using a threshold of 0.36. With this setting the radiomics signature had an accuracy and precision for sensitive patients of 82 % (95 % CI:60–95) and 90 % (95 % CI:73–97), respectively. Conversely, a precision of 80 % (95 % CI:34–97) was obtained in non-responsive patients using a threshold of 0.9. Indeed, with the latter setting 4 patients out of 5 were correctly predicted as non-responders.
Conclusions: The results are a first step towards using radiomics to improve the management of patients with GIST, especially when tumor tissue is unavailable for molecular analysis or when molecular profiling is inconclusive.

1. Introduction

Gastrointestinal stromal tumors (GISTs) are the most common sarcomas of the gastrointestinal tract (GIT), despite representing less than 1 % of all GIT malignant tumors [1]. They originate from the interstitial cells of Cajal [2], being the most frequent sites the stomach (50 %), small intestine (25 %), and rectum (5 %) [1]. About 80 % of GISTs harbor mutations in the gene encoding the transmembrane tyrosine-protein kinase (KIT) [3], in either exon 11 or 9. In 5–10 % of cases, mostly in GIST arising from the stomach, mutations involve the platelet-derived growth factor α (PDGFRA) gene [4], while in another 10–15 % of cases no mutations are detected in either KIT or PDGFRA (historically called KIT/PDGFRA Wild-Type – WT GISTs) [5].

The prognosis of GIST patients has been revolutionized by the advent of the first tyrosine kinase inhibitor (TKI), imatinib [6]. GIST mutational status represents a predictive factor of response to TKI along with prognostic relevance [7–9]. Exon 11 and exon 9 KIT mutations are imatinib-sensitive, with exon 9 mutations requiring a higher dose of the drug. PDGFRA mutations are also imatinib-sensitive, apart from exon 18 D842V mutation, which is refractory to imatinib treatment, while WT GISTs could be considered mainly non-responsive to imatinib [1].

Small and localized GISTs are usually surgically removed, with optional adjuvant treatment performed depending on risk category [1]. In case of locally-advanced or metastatic GISTs, medical therapy is usually recommended up-front [1], following determination of the mutational status by means of tissue biopsy. Unfortunately, biopsy is usually not representative of the whole lesion, which may presents as a large and heterogeneous mass [10]. Moreover, retrieving adequate tissue for accurate staging, assessment of mitotic count, and mutational analysis may not be trivial when the tumor is located in deep, difficult to access regions such as the small bowel [11].

In this context, the extraction of information from routinely acquired computed tomography (CT) scans of the whole tumor through radiomics analysis could represent a low-cost solution that might help the decision-making process [12]. In recent years, CT features [13] and radiomics signatures [14] have been successfully developed to predict the malignant potential of GIST and for risk stratification. The use of radiomics to correlate image phenotype to the different molecular profiles of the tumor may have practical implications since it could, by instance, avoid the toxicity of therapy in patients with imatinib-resistant GIST.

Therefore, the aim of this dual-center retrospective study is to develop and validate a radiomics signature to non-invasively differentiate GIST harboring imatinib-resistant mutations from imatinib-sensitive genotypes, using features extracted from pre-treatment CT images.

2. Materials and methods

2.1. Study design and patients

This retrospective study included patients with diagnosis of primary GIST between 2008 and 2021 from two centers in Italy: Candiolo Cancer Institute in Torino (center A) and University Hospital "Paolo Giaccone" in Palermo (center B). Patients from center A were used to train and tune the algorithms (construction set), while patients from center B were used to externally validate the radiomics models. The study was performed in accordance with the principles of the Declaration of Helsinki and the International Con-ference on Harmonization and Good Clinical Practice guidelines and was approved by the Institutional Review Board and Ethic Committee of both centers (Protocol CE IRCCS 257/2019 of September 25TH 2019 for Center A and protocol N° 2/2020 of February 19th, 2020, for Center B). Informed consent was obtained from patients that were alive and traceable.

Inclusion criteria were the following: 1) diagnosis of GIST confirmed by histo-pathological analysis including mutational profile; 2) availability of baseline contrast enhanced CT images in the portal venous phase; 3) lesions with maximum diameter $\geq$ 1.5 cm; 4) no previous medical treatment; 5) absence of other tumors. Exclusion criteria were the fol-lowing: 1) lack of preoperative imaging or unavailability of the portal venous phase; 2) low quality CT scans due to image artifacts or scan thickness of more than 3 mm; 3) lack of information on mutational status.

2.2. Imaging technique

Patients performed multiphasic contrast-enhanced CT imaging using the following scanners: 64-slice (Somatom Definition Flash, Siemens Healthineers) for center A and 16-slice (Bright- Speed, GE Medical Systems) or 128-slice (Somatom Definition AS, Sie-mens Healthineers) for center B. Post-contrast images were obtained after the intravenous administration of contrast medium (1.5–2 mL/kg of a nonionic contrast medium [370–400 mgI/mL]) injected at a rate of 3–5 mL/s, using an automatic power injector. Radiomics analysis was performed on the portal-phase acquisition, obtained with a delay of 60–70 s after the beginning of the contrast agent injection. CT scanning parameters were the following: X-ray tube voltage: 100–140 kV; tube current: 180–230 mA; slice thickness $\leq$ 3 mm; field of view 300–350 mm; 512 $\times$ 512 matrix.

2.3. Reference standard

Reference standard was the mutational status of KIT and PDGFRA. Patients were dichotomized according to imatinib sensitivity as follows:

1. Group 0: imatinib sensitive patients, with KIT mutation or PDGFRA mutation different from exon 18 - D842V;
2. Group 1: imatinib non-responsive patients, with PDGFRA exon 18 - D842V mutation and WT GISTs.

2.4. Image segmentation and features extraction

Three resident radiologists with at least 3 years of clinical experience in oncologic imaging, manually contoured the lesions using ITK-Snap [15] (http://www.itksnap.org). Segmentations were performed on all CT slices including the tumor and, if present, intralesional necrotic areas were removed. All segmentation masks were then reviewed by an experienced radiologist (> 20 years of experience in reporting CT scans) and revised if needed. Approximately 70 % of baseline segmentations were slightly revised (minor revisions of lesion boundaries), while 10 % of them were strongly revised (when other anatomic structures adjacent

to the lesions or necrotic areas were erroneously segmented).

Before performing feature extraction, all GISTs masks were re-segmented between the 1st and the 99th percentile of the region-of-interest to remove outliers and noisy voxels that might bias the subsequent radiomics analysis, as previously reported [16]. Additionally, all images were discretized using a fixed number of bins (n = 32) to introduce a normalizing effect which may be beneficial when intensity units are arbitrary or differ among centers, and when image contrast is considered important [17]. We then computed the following 107 features: (1) 14 shape-based; (2) 18 first-order intensity-based statistics, such as mean, 25th, 50th, 75th percentiles, skewness, kurtosis, intensity kurtosis, and intensity variance; (3) 24 derived from the gray level co-occurrence matrices (GLCM); (4) 16 derived from the gray level run length matrices (GLRLM); (5) 16 derived from the gray level size zone matrix (GLSZM); (6) 5 derived from the neighboring gray tone difference matrix (NGTDM); (7) 14 derived from the gray level dependence matrix (GLDM). All features were computed symmetrically for each of the 4 directions of a 2D image, and then averaged. Texture features were computed using PyRadiomics [18] that was compliant to the image biomarker standardization initiative (IBSI) [19].

2.5. Features selection and radiomics model development

Feature selection (FS) was carried-out to select the most performing radiomics features among all extracted ones, to remove irrelevant data, reduce dimensionality, increase learning accuracy, improve results comprehensibility and reduce the risk of over-fitting [20,21]. The following methods for FS were tested and compared: Pearson Correlation, One Way ANOVA (Analysis of Variances) test, Least Absolute Shrinkage and Selection Operator (LASSO), ANOVA + LASSO and Importance computed through Random Forest (RF). Each FS algorithm was then combined with different statistical classifier, including Bayesian, Support-Vector Machine (SVM), k-Nearest Neighbor (kNN), and RF. Since KIT mutation is much more common than PDGFR-α mutation and WT status, we decided to use a balanced dataset to train the algorithms, thus avoiding biasing the output towards the most represented group. To this scope, from the whole construction set (i.e., center A) we created a balanced training set using all non-responsive patients (less numerous group) and an equal number of sensitive patients (most numerous group). The following 16 models were finally trained by exploring all combinations of FS method and classifiers: Anova+SVM, Anova+Bayes, Anova+KNN, Pearson+SVM, Pearson+Bayes, Pearson+KNN, Lasso+SVM, Lasso+Bayes, Lasso+KNN, Joined+SVM, Joined+Bayes, Joined+KNN, Importance+SVM, Importance+Bayes, Im-portance+KNN, RF. All models were fine-tuned, i.e., hyper-parameters were selected, by using a k-fold stratified cross-validation (k = 4) approach on the balanced training set.

2.6. Statistical analysis

Performances of each model were computed on the whole construction set (i.e., the whole center A) using the following metrics: 1) accuracy for imatinib sensitive: defined as the number of correctly classified imatinib sensitive patients over the total number of imatinib sensitive patients; 2) accuracy for imatinib non-responsive: defined as the number of correctly classified imatinib non-responsive patients over the total number of imatinib non-responsive patients; 3) precision for imatinib sensitive: defined as the number of correctly classified imatinib sensitive patients over the total number of patients classified as imatinib sensitive; 4) precision for imatinib non-responsive: defined as the number of correctly classified imatinib non-responsive patients over the total number of patients classified as imatinib non-responsive; 5) balanced accuracy: defined as the average between the accuracy in detecting imatinib sensitive and imatinib non-responsive patients. To evaluate these metrics, the radiomics score produced by each model was dichotomized using the Youden Index, a commonly used measure of overall diagnostic effectiveness that optimizes accuracy on both groups [22].

After training and fine-tuning all combinations, we discarded all models that did not reach at least 60 % on all metrics and/or that showed overfitting (i.e., accuracy on the less represented group equal to 100 %), and/or that selected a number of features at least 5 times larger than the number of patients in the training, as previously suggested [23]. Finally, the best performing model was chosen as the one with the highest balanced accuracy on the construction set (the whole center A). Once the best model was selected, its performances were also evaluated on the external validation set (center B) using the same metrics, computed using, as threshold of the radiomics score, both the Youden Index and the threshold that maximizes accuracy on sensitive patients. Differences in performances between the construction and the validation set and between accuracy and precision between the two groups were assessed using the "N-1" Chi-squared test. A p-value < 0.05 was considered statistically significant.

3. Results

The study flowchart is shown in Fig. 1. The final dataset was composed of 82 patients, of whom 52 from center A and 30 from center B. Patients from center A (63.4 % of the dataset) were imaged with a 64-slice scanner (Somatom Definition Flash, Siemens Healthineers); patients from center B were imaged with a 16-slice scanner (Bright- Speed, GE Medical Systems. 12 patients, 14.6 % of the dataset) or a 128-slice scanner (Somatom Definition AS, Siemens Healthineers. 18 Patients, 22 % of the dataset).

Overall, there were 57 patients in Group 0 (sensitive) and 25 in Group 1 (non-responsive). Demographic and clinical characteristics of the two datasets were compared in Table 1.

Fig. 2 shows the balanced accuracy for all combinations (FS methods + statistical classifiers) and the difference between accuracy in detecting imatinib sensitive and imatinib nonresponsive (i.e., delta) obtained on the construction set. The highest balanced accuracies (> 80 %) were obtained using the KNN classifier. However, this classifier was the one that was also characterized by a strong difference in accuracy between the two groups (delta > 20 %), meaning that the KNN is over-specialized towards one group, as it classifies most patients as non-responsive. This behaviour might be a symptom that overfitting on the less represented group has occurred since all cases of group 1 were included in the training set.

Supplementary Table 1 reports the results on both the construction and the validation set of all models reaching at least 60 % on all metrics, and that did not show a clear overfitting (i.e., accuracy on the less represented group equal to 100 %).

The model that produced the best radiomics signature discriminating imatinib sensitive and non-responsive patients was obtained by combining the FS method based on the ANOVA, with 5 features selected, and the Bayes classifier using a threshold on the radiomics score = 0.36. Performances of this model are reported in Table 2, while Fig. 3 shows the waterfall plot of the classification on both the construction and the validation set. No significant differences were reported between the two sets. The model was able to predict patients sensitive to imatinib with an accuracy of 82 % and precision of 90 % in the validation set. Slightly lower performances were obtained in predicting non-responsive patients, with an accuracy of 75 % and a precision of 60 %.

The radiomics score that maximizes accuracy on sensitive patients (0.9) increased the accuracy for sensitive patients to 95 % in the validation set, with a precision of 84 %. Moreover, it produced an increase of the precision of 20 %, i.e. from 60 % to 80 %, in non-responsive patients, balanced by a reduction of its accuracy from 75 % to 50 %. (Table 2 and Fig. 3).

Fig. 4 shows the weight of each of the selected feature in the signature. The three most important features were related to entropy, i.e., a physical property that is mostly associated with a state of disorder.

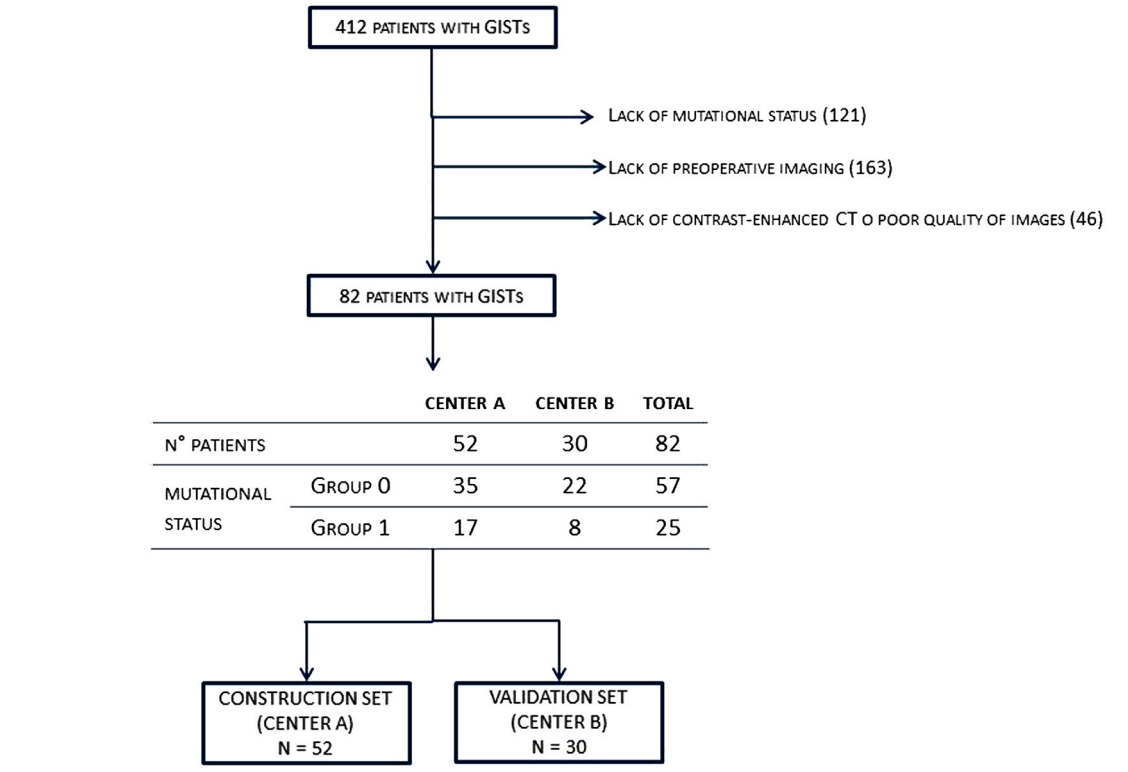

Fig. 1. Flowchart of the study.

Table 1
Demographic and clinical characteristics of the two dataset.

	Group 0 (n = 57)	Group 1 (n = 25)
Gender		
Males	34 (59.6)	15 (60.0)
Females	23 (40.4)	10 (40.0)
Age	63.5 ± 11.8	61.0 ± 10.8
Location / Size (mm)		
Stomach	35 (61.5) / 77.4 ± 52.7	22 (88) / 66.6 ± 38.9
Small bowel		
Duodenum	4 (7.0) / 45.5 ± 20.1	2 (8) / 58.5 ± 7.7
Jejunum	3 (5.2) / 65.6 ± 44.8	–
Ileum	11 (19.3) / 107.5 ± 66.5	1 (4) / 22.0 ± 0.0
Rectum	4 (7.0) / 65.7 ± 33.7	–

Group 0: imatinib sensitive patients, with KIT mutation or PDGFRA mutation different from exon 18 - D842V.
Group 1: imatinib non-responsive patients, with PDGFRA exon 18 - D842V mutation and non-KIT/non-PDGFR mutant GISTs (WT).
Continuous variables (age and size) are expressed as mean ± SD; categorical variables (gender and location) are expressed as numbers and percentages in parenthesis.

Predicted non-responsive patients showed a median higher value of these variables, meaning that they have a less homogenous appearance (Fig. 5).

4. Discussion

Today the management of patients affected by localized GISTs is mainly driven by risk stratification and mutational status. Risk of GIST recurrence depends on tumor location, tumor size and mitotic index [24, 25]; tumor capsule rupture is an additional major determinant of relapse risk [26]. The above reported criteria classify GISTs in different relapse risk groups, spanning from "none" to "high", and substantiate treatment decisions and follow-up management. Conversely, mutational status has not yet been formally included in any risk classification criteria, and its correlation with risk relapse is currently a matter of debate [1]. Nonetheless, mutational status should be included in the decision-making workflow because of its predictive value for sensitivity to TKIs therapy. Indeed, it is well-known that some genotypes are less sensitive to imatinib (e.g, WT GISTs) while others are completely resistant/refractory (e.g. PDGFRα exon 18 D842V).

This study developed a radiomics signature able to differentiate GISTs with genotypes that are usually sensitive to imatinib (KIT mutations and PDGFR-α mutations different from exon 18 - D842V) from genotypes with lower sensitivity or that are insensitive to imatinib (PDGFR-α exon 18 - D842V mutation and WT GISTs) using 5 features extracted from images of the whole tumor at the baseline CT scan.

Using a threshold of 0.36, the model was able to predict with high accuracy and precision patients sensitive to imatinib. Indeed, in the validation set, 18 out of 22 (accuracy = 82 %) sensitive patients were correctly predicted by the model, while of the 20 patients predicted as sensitive, 18 were in fact sensitive to imatinib (precision = 90 %). The performance of our radiomics signature was slightly lower in predicting non-responsive patients, since 2 out of 8 were incorrectly classified as sensitive (accuracy of non-response = 75 %), while among the 10 patients classified as non-responsive only 6 were actually non-responsive (precision of non-responsive = 60 %).

However, using the threshold that maximizes accuracy on sensitive patients (from 0.36 to 0.9, Fig. 3), the model reached a precision of 80 % for non-responsive patients. Indeed, in the validation set, 4 patients were correctly classified as non-responsive to imatinib among the 5 patients predicted as non-responsive. Altogether, these findings are potentially relevant, particularly in cases where tumor tissue specimens are insufficient for molecular analysis and/or when tumor is located in hardly accessible regions for biopsy. Indeed, in these cases, applying our radiomics signature we could have correctly avoided a useless and toxic

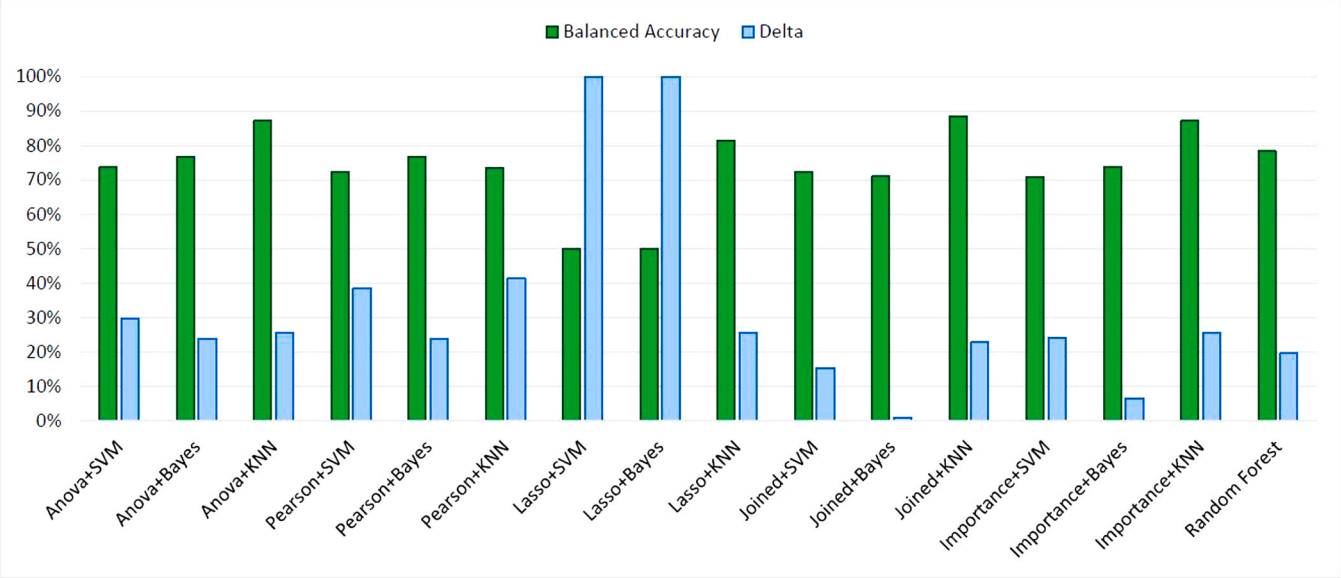

Fig. 2. Balanced accuracy and delta obtained on the construction set for all combinations of features selection and classifier. Delta represents the difference between the accuracy in detecting imatinib sensitive and imatinib non-responsive patients.

Table 2
Performances of the best model (ANOVA + Bayes) on the construction and the external validation set.

	Accuracy on non-responsive patients (%) [95 %CI] (rate)	Accuracy on sensitive patients (%) [95%CI] (rate)	p-value	Precision for non-responsive patients (%) [95 %CI] (rate)	Precision for sensitive patients (%) [95 %CI] (rate)	p-value
Youden Index (0.36)						
CONSTRUCTION SET	65 [38–86] (11/17)	89 [70–95] (31/35)	0.08	73 [48–84] (11/15)	84 [72–91] (31/37)	0.26
VALIDATION SET	75 [34–97] (6/8)	82 [60–95] (18/22)	0.50	60 [36–80] (6/10)	90 [73–97] (18/20)	0.06
p-value	0.62	0.69		0.64	0.48	
Threshold that maximizes accuracy on sensitive patients (0.9)						
CONSTRUCTION SET	29 [10–56] (5/17)	94 [81–99] (33/35)	<0.001	71 [35–92] (5/7)	73 [67–79] (33/45)	0.91
VALIDATION SET	50 [16–84] (4/8)	95 [77–100] (21/22)	0.004	80 [34–97] (4/5)	84 [72–91] (21/25)	0.83
p-value	0.31	0.87		0.73	0.30	

therapy in 80 % of non-responsive patients. Finally, we found that the three most important features of our model were related to entropy, a physical property that is mostly associated with a state of disorder, with predicted non-responsive patients having a median higher value of the variables and consequently a less homogeneous appearance (Fig. 5).

Only one recent single center study performed radiomics analysis to differentiate responder versus non-responder genotypes to imatinib treatment. In their study, Palatresi et al. [27] divided patients affected by GISTs in a heterogeneous cohort (intended as images acquired from 3 different CT scans) and homogeneous cohort (from one CT scan). In the heterogeneous cohort (27 responder patients and 21 non-responder patients) they found a statistically significant relationship between the two groups and two radiomics features for non-contrast enhanced CT (NCECT) images, but none for the contrast enhanced CT (CECT). In the homogeneous cohort (9 responder patients and 6 non-responder patients) they found a statistically significant relationship between the two groups and one radiomics feature for NCECT images and four radiomics features for CECT. Other four studies have also proposed the use of a radiomics algorithm, or nomogram including clinical features, to evaluate the mutational status of patients affected by GISTs. However, in the study of Starmans et al. [28] no significant result were achieved (best AUC: 0.56 to identify KIT exon 11 mutation), while in the other studies the main purpose was to identify patients with KIT exon 11 mutation (i.e., higher sensibility to imatinib) from the all other genotypes [29–31].

Our study has some limitations. First, it is a retrospective study and the sample size is relatively small and unbalanced due to the rarity of the tumor and, in particular, of the non-responsive group. The small size may explain why we did not find any significant difference between accuracy and precision of the two groups of sensitive and non-responsive patients. Hopefully, validation studies on larger, multi-institutional databases might improve our promising preliminary findings. Second, GISTs are often large tumors, and the manual segmentation of the entire volume is a time-consuming process and a possible source of errors. Automatic segmentation with deep learning algorithms will hopefully

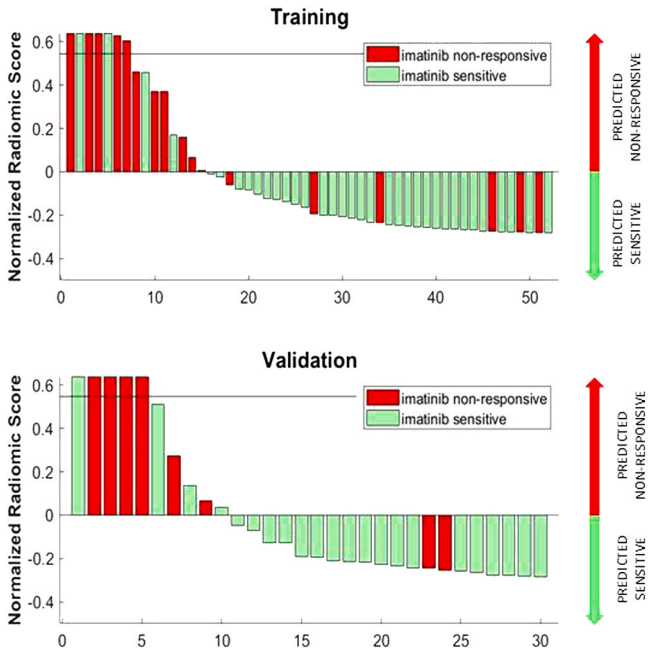

Fig. 3. Waterfall plot of the prediction probabilities of the Bayes Classifier. The radiomics score of 0.36 is normalized to 0. Negative radiomics score represent the probability of being imatinib sensitive while positive radiomics score represent the probability of being non-responsive. Therefore, green bars having a positive section represent patients incorrectly classified as non-responsive, while red bars having a negative radiomics score are patients incorrectly classified as sensitive. The black line represents the threshold that maximizes accuracy on sensitive patients resulting in a lower number of patients correctly classified as non-responsive (reduced accuracy), but an higher number of correctly classified imatinib non-responsive patients over the total number of patients classified as non-responsive (increase of precision for imatinib non-responsive patients).

help to overcome this limitation in the future. Notwithstanding, to the best of our knowledge, this is the first study that has developed a radiomics model to differentiate imatinib sensitive from non-responsive genotypes in GIST patients with an external validation cohort, potentially providing an effective biomarker in support of clinical decision. The possibility to correlate radiomics features with other known risk factors deserves further investigation aiming to improve our ability to predict GIST behaviour.

5. Conclusion

In this study we have developed a radiomics signature correlating pre-treatment CT features to GIST mutational status with the aim of predicting response to imatinib treatment. This novel biomarker will be of help especially when tumor tissue is unavailable/inadequate for molecular analysis. In perspective, radiomics along with demographic, pathological, laboratory, and clinical variables might improve the management of patients with GIST, paving the way to a more personalized, "patient-tailored", and accurate treatment algorithm.

Funding

This research did not receive any specific grant from funding agencies in the public, commercial, or not-for-profit sectors.

CRediT authorship contribution statement

Landolfi Ilenia: Data curation. **Serra Giovanni:** Validation, Software. **Molea Francesca:** Data curation. **Damiani Nicolò:** Data curation. **Regge Daniele:** Writing – review & editing, Supervision, Project administration, Methodology, Conceptualization. **Incorvaia Lorena:** Data curation. **Cappello Giovanni:** Writing – original draft, Visualization, Methodology, Formal analysis, Data curation, Conceptualization. **Badalamenti Giuseppe:** Data curation. **Giannini Valentina:** Writing – original draft, Validation, Software, Methodology, Formal analysis, Conceptualization. **Porrello Giorgia:** Data curation. **Gozzo Cecilia:** Investigation. **D'Ambrosio Lorenzo:** Writing – review & editing, Conceptualization. **Ambrosini Ilaria:** Investigation. **Bartolotta Tommaso Vincenzo:** Writing – review & editing, Supervision. **Grignani Giovanni:** Writing – review & editing, Conceptualization. **Cannella Roberto:** Methodology, Investigation, Conceptualization. **Merlini Alessandra:** Writing – review & editing, Conceptualization. **Tabone Emanuele:** Investigation.

Declaration of Competing Interest

The authors declare the following financial interests/personal relationships which may be considered as potential competing interests: Roberto Cannella: - support for attending meetings from Bracco and Bayer; co-funding by the European Union - FESR or FSE, PON Research

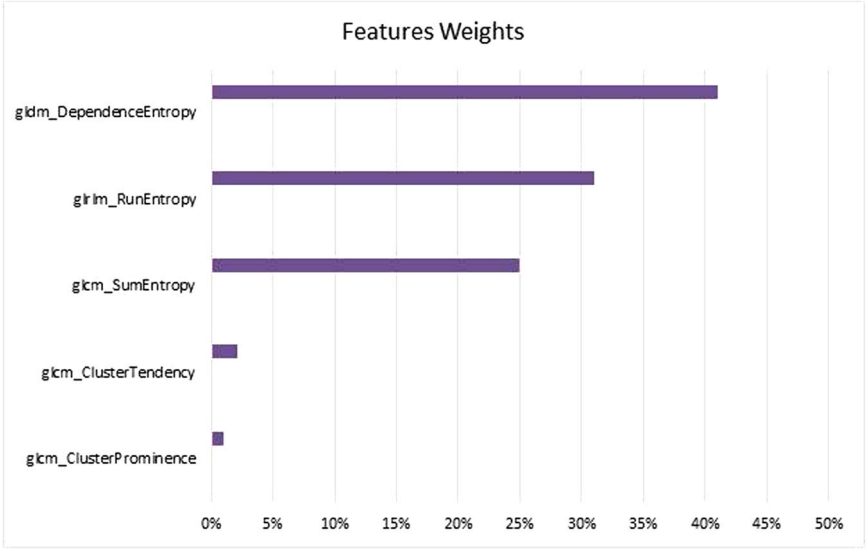

Fig. 4. Weights of each selected feature on the Bayes Classifier.

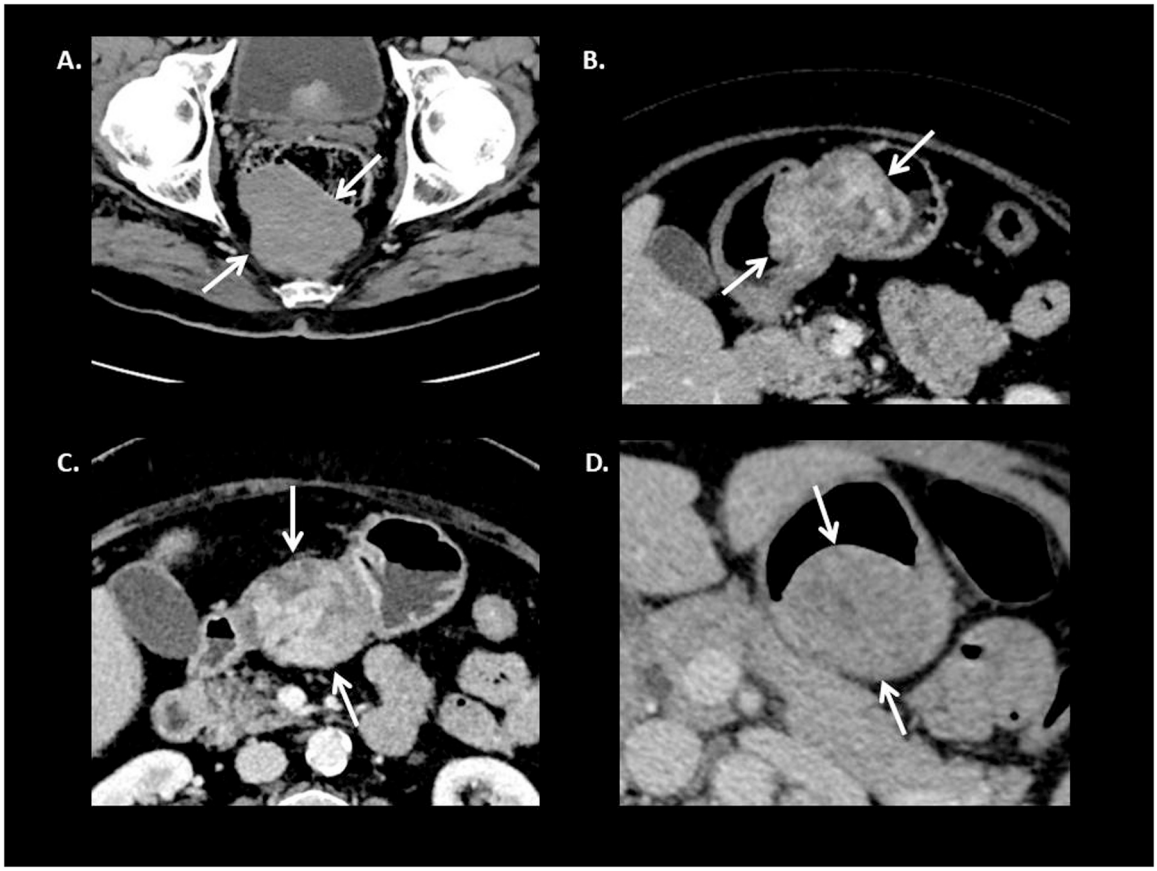

Fig. 5. four portal phase CT scans of different patients. GIST with KIT exon 11 mutation (imatinib sensitive) shows a homogeneous and hypodense appearance (Fig. 5A, white arrows), while GIST with WT status (non-responsive) shows an heterogeneous appearance with different regions of higher contrast enhancement (Fig. 5B, white arrows). Both GISTs were correctly predicted by our radiomics model. Fig. 5C (white arrows) shows a KIT exon 11 GIST (imatinib sensitive) and it represents one of the eight imatinib-sensitive GIST erroneously predicted as non-responsive. At CT scan it shows an heterogeneous enhancement, differently than expected. Conversely, Fig. 5D (white arrows) shows a PDGFRA D842V GIST (non-responsive) where we expected an heterogeneous enhancement, while CT images showed an homogeneous appearance: this example represent one of the eight non-responsive GIST erroneously predicted as imatinib sensitive.

and Innovation 2014–2020 - DM 1062/2021.

Acknowledgments

GIST-CRC 5 per 1000 2014.

Appendix A. Supporting information

Supplementary data associated with this article can be found in the online version at doi:10.1016/j.ejro.2023.100505.

References

[1] P.G. Casali, J.Y. Blay, N. Abecassis, et al., Gastrointestinal stromal tumours: ESMO-EURACAN-GENTURIS clinical practice guidelines for diagnosis, treatment and follow-up, Ann. Oncol. J. Eur. Soc. Med Oncol. (2022), https://doi.org/10.1016/j.annonc.2021.09.005.

[2] L.G. Kindblom, H.E. Remotti, F. Aldenborg, J.M. Meis-Kindblom, Gastrointestinal pacemaker cell tumor (GIPACT): gastrointestinal stromal tumors show phenotypic characteristics of the interstitial cells of Cajal, Am. J. Pathol. 152 (5) (1998) 1259–1269.

[3] H. Zhang, Q. Liu, Prognostic indicators for gastrointestinal stromal tumors: a review, Transl. Oncol. (2020), https://doi.org/10.1016/j.tranon.2020.100812.

[4] N. Machairiotis, I. Kougioumtzi, P. Zarogoulidis, A. Stylianaki, K. Tsimogiannis, N. Katsikogiannis, Gastrointestinal stromal tumor mesenchymal neoplasms: the offspring that choose the wrong path, J. Multidiscip. Health (2013). DOI: 10.2147/JMDH.S43703.

[5] C.L. Corless, C.M. Barnett, M.C. Heinrich, Gastrointestinal stromal tumours: origin and molecular oncology, Nat. Rev. Cancer (2011), https://doi.org/10.1038/nrc3143.

[6] T.M. Parab, M.J. DeRogatis, A.M. Boaz, et al., Gastrointestinal stromal tumors: a comprehensive review, J. Gastrointest. Oncol. (2019), https://doi.org/10.21037/jgo.2018.08.20.

[7] A. Wozniak, P. Rutkowski, P. Schöffski, et al., Tumor genotype is an independent prognostic factor in primary gastrointestinal stromal tumors of gastric origin: a european multicenter analysis based on ConticaGIST, Clin. Cancer Res. (2014), https://doi.org/10.1158/1078-0432.CCR-14-1677.

[8] S. Rossi, D. Gasparotto, R. Miceli, et al., KIT, PDGFRA, and BRAF mutational spectrum impacts on the natural history of imatinib-naive localized GIST: a population-based study, Am. J. Surg. Pathol. (2015), https://doi.org/10.1097/PAS.0000000000000418.

[9] L. Incorvaia, D. Fanale, B. Vincenzi, et al., Type and gene location of KIT mutations predict progression-free survival to first-line imatinib in gastrointestinal stromal tumors: a look into the exon, Cancers (2021), https://doi.org/10.3390/cancers13050993.

[10] A. Inoue, S. Ota, M. Yamasaki, et al., Gastrointestinal stromal tumors: a comprehensive radiological review, Jpn J. Radio. (2022), https://doi.org/10.1007/s11604-022-01305-x.

[11] T. Nishida, S. Yoshinaga, T. Takahashi, et al., Recent progress and challenges in the diagnosis and treatment of gastrointestinal stromal tumors, Cancers (Basel) (2021), https://doi.org/10.3390/cancers13133158.

[12] R.J. Gillies, P.E. Kinahan, H. Hricak, Radiomics: images are more than pictures, they are data, Radiology (2016), https://doi.org/10.1148/radiol.2015151169.

[13] R. Cannella, E. Tabone, G. Porrello, Assessment of morphological CT imaging features for the prediction of risk stratification, mutations, and prognosis of gastrointestinal stromal tumors, Eur. Radio. (2021), https://doi.org/10.1007/s00330-021-07961-3.

[14] Y. Song, J. Li, H. Wang, et al., Radiomics nomogram based on contrast-enhanced CT to predict the malignant potential of gastrointestinal stromal tumor: a two-center study, Acad. Radio. (2022), https://doi.org/10.1016/j.acra.2021.05.005.

[15] P.A. Yushkevich, J. Piven, H.C. Hazlett, et al., User-guided 3D active contour segmentation of anatomical structures: significantly improved efficiency and reliability, NeuroImage (2006), https://doi.org/10.1016/j.neuroimage.2006.01.015.

[16] S.S.F. Yip, H.J.W.L. Aerts, Applications and limitations of radiomics, Phys. Med. Biol. (2016), https://doi.org/10.1088/0031-9155/61/13/R150.
[17] A. Zwanenburg, S. Leger, L. Agolli, et al., Assessing robustness of radiomic features by image perturbation, Sci. Rep. (2019), https://doi.org/10.1038/s41598-018-36938-4.
[18] J.J.M. van Griethuysen, A. Fedorov, C. Parmar, et al., Computational radiomics system to decode the radiographic phenotype, Cancer Res (2017), https://doi.org/10.1158/0008-5472.CAN-17-0339.
[19] M. Hatt, M. Vallieres, D. Visvikis, A. Zwanenburg, IBSI: an international community radiomics standardization initiative, J. Nucl. Med. 59 (2018) 287-287.
[20] Rosati S., Gianfreda C.M., Balestra G., Martincich L., Giannini V., Regge D. (2018) Correlation based Feature Selection impact on the classification of breast cancer patients response to neoadjuvant chemotherapy. IEEE International Symposium on Medical Measurements and Applications (MeMeA) DOI: 10.1109/MeMeA.2018.8438698.
[21] Yu L., Liu H. (2022) Feature Selection for High-Dimensional Data. Proceedings, Twentieth International Conference on Machine Learning. Available via ⟨https://researchr.org/publication/YuL03%3A1⟩.
[22] E.F. Schisterman, N.J. Perkins, A. Liu, H. Bondell, Optimal cut-point and its corresponding Youden index to discriminate individuals using pooled blood samples, Epidemiology 16 (1) (2005) 73–81.
[23] E. Vittinghoff, C.E. McCulloch, Relaxing the rule of ten events per variable in logistic and Cox regression, Am. J. Epidemiol. (2007), https://doi.org/10.1093/aje/kwk052.
[24] M. Miettinen, J. Lasota, Gastrointestinal stromal tumors: pathology and prognosis at different sites, Semin Diagn. Pathol. (2006), https://doi.org/10.1053/j.semdp.2006.09.001.
[25] M. Miettinen, J. Lasota, Gastrointestinal stromal tumors: review on morphology, molecular pathology, prognosis, and differential diagnosis, Arch. Pathol. Lab Med. (2006), https://doi.org/10.5858/2006-130-1466-GSTROM.
[26] H. Joensuu, A. Vehtari, J. Riihimäki, et al., Risk of recurrence of gastrointestinal stromal tumour after surgery: an analysis of pooled population-based cohorts, Lancet Oncol. (2012), https://doi.org/10.1016/S1470-2045(11)70299-6.
[27] D. Palatresi, F. Fedeli, G. Danti, Correlation of CT radiomic features for GISTs with pathological classification and molecular subtypes: preliminary and monocentric experience, Radio. Med. (2022), https://doi.org/10.1007/s11547-021-01446-5.
[28] M.P.A. Starmans, M.J.M. Timbergen, M. Vos, et al., Differential diagnosis and molecular stratification of gastrointestinal stromal tumors on CT images using a radiomics approach, J. Digit Imaging (2022), https://doi.org/10.1007/s10278-022-00590-2.
[29] F. Xu, X. Ma, Y. Wang, et al., CT texture analysis can be a potential tool to differentiate gastrointestinal stromal tumors without KIT exon 11 mutation, Eur. J. Radio. (2018), https://doi.org/10.1016/j.ejrad.2018.07.025.
[30] X. Liu, Y. Yin, X. Wang, et al., Gastrointestinal stromal tumors: associations between contrast-enhanced CT images and KIT exon 11 gene mutation, Ann. Transl. Med (2021), https://doi.org/10.21037/atm-21-3811.
[31] B. Liu, H. Liu, L. Zhang, et al., Value of contrast-enhanced CT based radiomic machine learning algorithm in differentiating gastrointestinal stromal tumors with KIT exon 11 mutation: a two-center study, Diagn. Inter. Radio. (2022), https://doi.org/10.5152/dir.2021.21600.