Journal of Pathology Informatics 13 (2022) 100152

Contents lists available at ScienceDirect

Journal of Pathology Informatics

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

Development of a method for digital assessment of tumor regression grade in patients with rectal cancer following neoadjuvant therapy

Dea Natalie Munch Jepsen [a,b,*], Henrik Høeg [c], Jeppe Thagaard [c], Julie Sparholt Walbech [b], Ismail Gögenur [b,d], Anne-Marie Kanstrup Fiehn [a,b,d]

[a] Department of Pathology, Zealand University Hospital, Roskilde, Denmark
[b] Center for Surgical Science, Department of Surgery, Zealand University Hospital, Køge, Denmark
[c] Visiopharm A/S, Hørsholm, Denmark
[d] Department of Clinical Medicine, University of Copenhagen, Denmark

ARTICLE INFO

Keywords:
Rectal cancer
Neoadjuvant treatment
Tumor regression grade
Digital pathology
Deep learning

ABSTRACT

Neoadjuvant chemo-radiotherapy (nCRT) followed by surgical resection is the standard treatment strategy in patients with locally advanced rectal cancer (RC). The pathological effect of nCRT is assessed by determining the tumor regression grade (TRG) of the resected tumor. Various methods exist for assessing TRG and all are performed manually by the pathologist with an accompanying risk of interobserver variation. Automated digital image analysis could be a more objective and reproducible approach to evaluate TRG. This study aimed at developing a digital method to assess TRG in RC following nCRT, and correlate the results to the currently used Mandard method. A deep learning-based semi-automatic Epithelium-Tumor area Percentage (ETP) algorithm enabling quantification of tumor regression by determining the percentage of residual tumor epithelium out of the total tumor area was developed. The ETP was quantified in 50 cases treated with nCRT and 25 cases with no prior nCRT served as controls. Median ETP was 39.25% in untreated compared with 6.64% in patients who received nCRT ($P < .001$). The ETP of the resected tumors treated with nCRT increased along with increasing Mandard grade ($P < .001$). As new treatment strategies in RC are emerging, performing an accurate and reproducible evaluation of TRG is important in the assessment of treatment response and prognosis. TRG is often used as an outcome point in clinical trials. The ETP algorithm is capable of performing a precise and objective value of tumor regression.

Introduction

Colorectal cancer (CRC) is the third most common cancer type and is ranking second according to mortality only surpassed by lung cancer.[1] The incidence of rectal cancer (RC) has been increasing among young adults and around 1/3 of patients with RC are younger than 55 years.[2] Currently, surgery is the main curative treatment in patients with RC and the standard treatment strategy for patients with locally advanced (T3/4, N+) disease is neoadjuvant chemo-radiotherapy (nCRT) followed by surgical resection.[3]

The pathological effect of nCRT is assessed by determining the tumor regression grade (TRG) of the resected specimen. The Mandard TRG method was introduced in 1994 for the evaluation of tumor regression after nCRT in patients with carcinoma of the esophagus.[4] Since then, various similar methods have been used in tumor regression evaluation including Dworak, American Joint Committee on Cancer (AJCC), and Ryan.[5–8] Different rates of agreement between pathologists assessing TRG have been reported in studies evaluating interobserver agreement.[7,9,10] The Mandard method classify the regression changes into 5 grades, based on an estimate of the distribution of carcinoma cells and fibrosis in the assumed tumor area. A lower TRG score represents a better response to nCRT (Table 1).[4] The different methods used in the clinic to assess TRG in RC are all based on a subjective estimate made by the individual pathologist. In contrast to this, automated digital image analysis is a standardized and objective method.[11] The latest development in the field, machine learning with neural networks and deep learning, increases the robustness of the analyses.

This study aimed to develop a digital deep learning-based method to assess TRG in RC following nCRT and correlate the results to the manually assessed Mandard TRG.

* Corresponding author at: Department of Pathology, Zealand University Hospital, Sygehusvej 9, 4000 Roskilde, Denmark.
E-mail addresses: dea.natalie.jepsen@gmail.com (D.N.M. Jepsen), hho@visiopharm.com (H. Høeg), jth@visiopharm.com (J. Thagaard), julwa@regionsjaelland.dk (J.S. Walbech), igo@regionsjaelland.dk (I. Gögenur), ankf@regionsjaelland.dk (A.-M.K. Fiehn).

http://dx.doi.org/10.1016/j.jpi.2022.100152
Received 2 September 2022; Received in revised form 7 October 2022; Accepted 2 November 2022
Available online 8 November 2022

Table 1
Mandard tumor regression grading.

TRG1	No residual cancer cells. Complete regression.
TRG2	Rare residual cancer cells.
TRG3	Fibrosis outgrowing residual cancer.
TRG4	Residual cancer outgrowing fibrosis.
TRG5	Absence of regressive changes.

Subjects and methods

Inclusion of patients and clinicopathological data

We included patients who received surgical resection due to a RC at the Department of Surgery at Slagelse or Køge, Zealand University Hospital, Denmark, followed by histological analysis of the surgical specimen at the Department of Pathology, Zealand University Hospital, Denmark. In daily routine, tumors are categorized according to the 8th edition of the International Union Against Cancer´s TNM Classification and the tumor histology according to the WHO classification.[12] Specimens are coded according to the SNOMED system and data is entered in an electronic database.[13] Patients for the study were identified by performing a search in the database, and included based on date of operation from June 1, 2020 and going chronologically backwards. Fifty patients treated with nCRT, and a control group of 25 patients receiving surgical resection without prior nCRT were included. Only patients with a diagnosis of adenocarcinoma (AC) in the pre-operative diagnostic biopsy combined with either a diagnosis of AC in the resected specimen or complete tumor regression after nCRT were included. Cases with mucinous AC and cases where the slides were unavailable at the archive were not included. Also, patients initially treated with local resection of the tumor without lymphadenectomy were excluded. For each patient, the pathology reports from the diagnostic biopsy and resected specimen were reviewed and clinicopathological data including sex, age, (y)pT-category, and (y)pN-category were extracted.

Inclusion of slides and manual histological evaluation

All hematoxylin and eosin (H&E)-stained slides with tissue from the diagnostic biopsies and resected specimens were retrieved from the archive. Diagnostic biopsy slides were reviewed to confirm an initial diagnosis of AC. All slides from the resected specimens were reviewed. Only slides with invasive tumor or areas presumed as representing regression changes were included, except in cases with complete regression of the primary tumor (ypT0), in which only one representative slide from the resected specimen was included. Areas with regression in the patients who had received nCRT were defined as fibrosis potentially combined with acellular mucin pools, calcifications, foreign body giant cells, or cholesterol crystals. Additionally, the quality of the H&E-staining was evaluated. In cases with very poor staining quality, a new slide was cut from the corresponding formalin-fixed paraffin-embedded (FFPE) tissue block and stained with H&E. Each of the 50 cases treated with nCRT were manually evaluated according to Mandard TRG by two pathologists (one specialized GI pathologist and one trainee) reaching consensus using a two-armed microscope.

Developing a digital algorithm

The included slides were scanned using the Hamamatsu NanoZoomer S60 (Hamamatsu, Japan), at magnification x20, and uploaded to the database in the Visiopharm AI software platform, version 2021.02.0.9368 (Visiopharm A/S, Hoersholm, Denmark). A digital algorithm calculating the percentage of epithelium of the total tumor area was developed with an already existing baseline algorithm from Visiopharms archive used as a starting point. One slide from 39 patients with colon cancer from The Cancer Genome Atlas (TCGA) and one slide from 35 patients from the present study population was selected and used as a training set. Thus, the selected slides were representative of the morphological spectrum of changes seen in patients with colorectal cancer. Representative regions were labeled and segmented into one of the following morphological classes; 1. tumor epithelium, 2. stroma incl. areas of regression, and 3. ulcer/necrosis/intraglandular debris. The algorithm was fully supervised trained with the AI module using deep learning-based neural networks (DeepLabv3 + , 200 000 iterations). Visual controls of the segmentation on specific field of views (FOVs) were executed intermittently. Relabeling and retraining was performed until the results were assessed as satisfying by a GI pathologist. Finally, a hold-out validation set of 25 additional full slides from the study population were processed and visually evaluated to ensure that the algorithm was able to classify the tissue into the defined morphological classes. Rule-based post-processing steps further improved the segmentation into one of the above-mentioned three tissue classes, and areas of white background were excluded (e.g., holes in the tissue). Areas of ulcer/necrosis/intraglandular debris were also excluded from the analysis. Output variables were set up to quantify the area of tumor epithelium and the total tumor area (tumor epithelium + stroma incl. areas of regression). Based on the obtained values, a calculation of the area of tumor epithelium out of the total tumor area was performed and the algorithm was named the Epithelium-Tumor area Percentage (ETP). A diagram of the workflow and utilization of the deep learning algorithm are presented in Figs. 1 and 2.

Selection of tumor area for quantification by the Epithelium-Tumor area Percentage (ETP) algorithm

In all cases, the region of interest (ROI) was defined as the total tumor area/the tumor bed. The total tumor area included the tumor epithelium, associated tumor stroma, and areas of tumor regression. Fig. 3 shows an example of the ETP algorithm workflow. On each included slide, the ROI was manually drawn as one coherent block with a small margin around the most peripherally located carcinoma glands or cells.

The image analysis algorithm was applied within the ROI, and by combining the obtained data of all slides for each individual case, a single ETP was calculated for every patient.

Evaluating the performance of the ETP algorithm

All processed slides were evaluated visually to ensure satisfactory segmentation into the three morphological classes. Further, to evaluate and quantify the performance of the ETP algorithm, labels on slides from 17 randomly selected cases covering the different response groups were

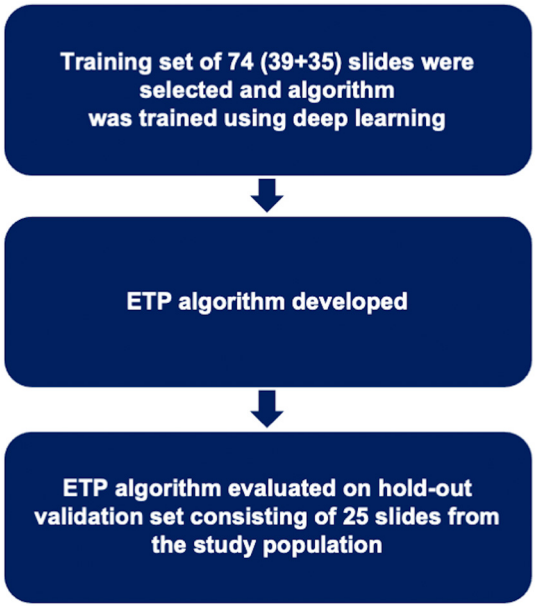

Fig. 1. Flow diagram of Epithelium-Tumor area Percentage (ETP) algorithm development.

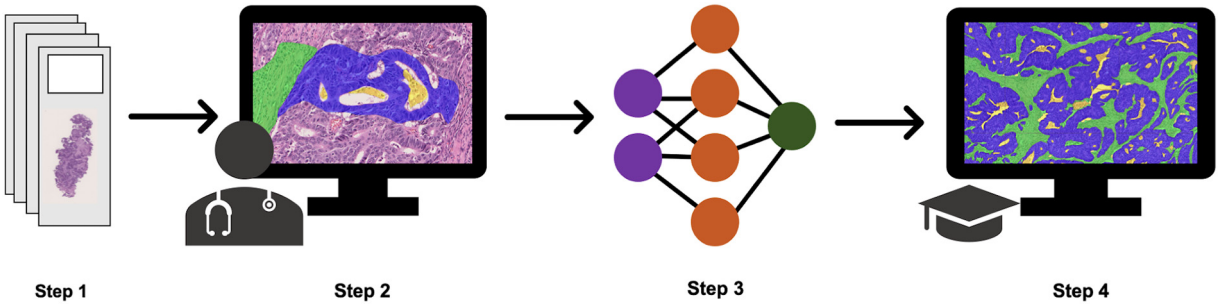

Fig. 2. Flow diagram of the deep learning-based algorithm development. Step 1: H&E-stained slides were scanned and uploaded to the database. Step 2: Images were labeled by pathologists. Blue: tumor epithelium, green: stroma, yellow: intraglandular debris. Step 3: Training of the algorithm with neural networks. Step 4: Deep learning algorithm developed. Rule-based post-processing of AI inference and output variables added to automated ETP algorithm.

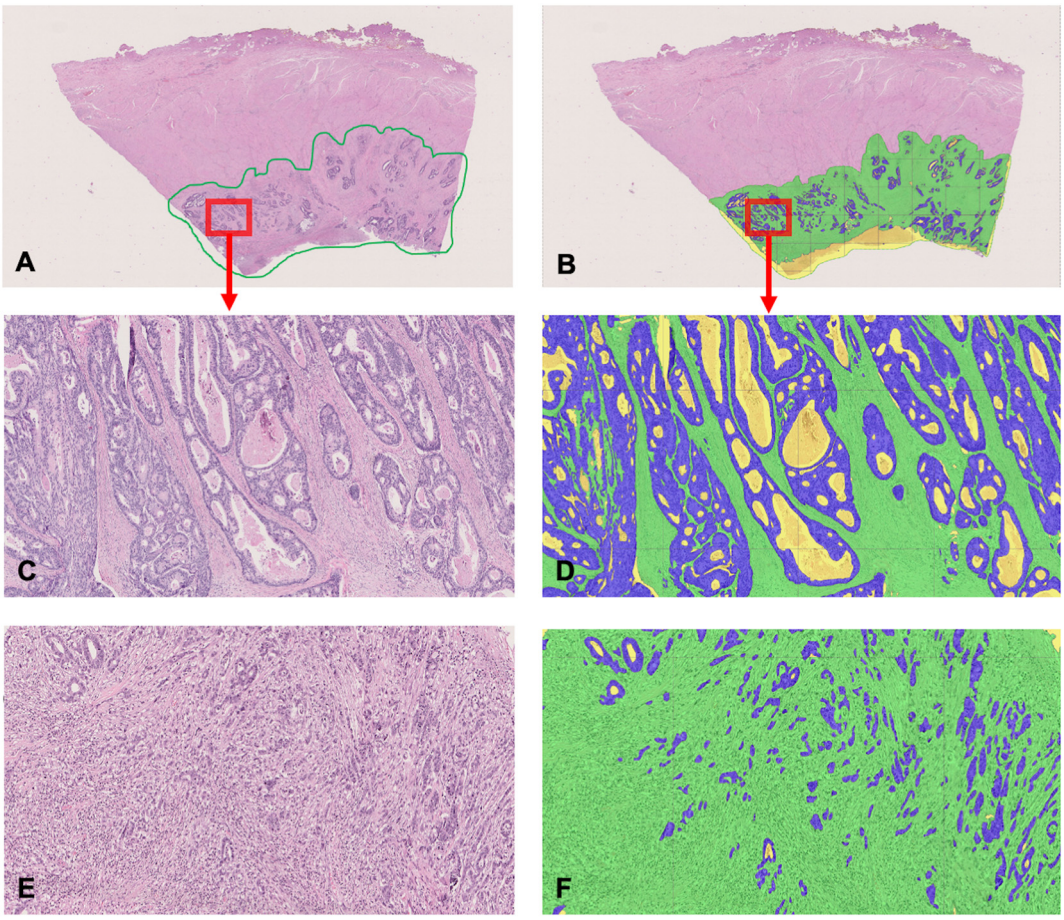

Fig. 3. The ETP algorithm workflow. A–D: A case of RC treated with nCRT. A (x0.44): The total tumor area is manually outlined on the image (green line). B (x0.44): The ETP algorithm segments the total tumor area into classes to perform quantification. Blue: tumor epithelium, green: stroma (incl. areas of regression), yellow: ulcer/necrosis/intraglandular debris, and white background (the yellow class was excluded from the analysis). On this slide, the determined ETP was 14.5%. C–D (x5): Higher magnification of tumor area in image A–B. E–F (x10): Another case with a less well-differentiated tumor.

visually evaluated by the pathologists. All areas with tumor epithelium misclassifications were corrected, and a new corrected ETP for each case was calculated and compared with the original ETP.

Statistics

Median ETP and interquartile ranges (IQR) are provided according to each TRG group as well as for the separate group of patients not receiving nCRT. The Mann–Whitney U test was used to compare the ETP of the patients who received nCRT with the patients who did not receive nCRT, and a *P*-value of <.05 was considered statistically significant. Mandard TRG groups with less than five patients were excluded from further statistical analysis comparing the ETP with manual TRG as recommended for the statistical test. The Kruskal–Wallis H test was used to assess if the ETP in one Mandard TRG group were different compared to the other Mandard groups, and a *P*-value of <.05 was considered statistically significant. The ETP of Mandard TRG groups were then compared in adjacent pairs (TRG1 vs. TRG2, TRG2 vs. TRG3, and TRG3 vs. TRG4) using the Mann–Whitney U

test. When comparing the ETP in adjacent pairs, the Bonferroni correction was used to adjust for multiple comparisons, and by dividing 0.05 with the number of comparisons performed (0.05/3), a *P*-value of <.017 was considered statistically significant.

All statistical analyses were performed using R, version 4.0.0, and the box and whisker plots using Excel, version 16.55.

Approvals

The study has been approved by The National Committee on Health Research Ethics (SJ-835), and was performed in accordance with the Declaration of Helsinki.

Results

Clinicopathological features of the included patients

The clinicopathological features of the included patients are presented in Table 2. Of the 50 patients with RC who received nCRT, the mean age was 65.5 years and 68% were male. Of the 25 patients with RC, who did not receive nCRT, the mean age was 71.9 years and 56% of the patients were male.

From the 75 resected RC specimens, 539 slides were included and analyzed with the ETP algorithm (369 slides from the 50 RC treated with nCRT and 170 slides from 25 RC not treated with nCRT). Of the 539 slides, only one slide was of very poor staining quality (too light pink) and a new section had to be cut from the corresponding FFPE block and stained with H&E. Additionally, in one slide, a small area of the tumor was manually excluded from the ETP analysis due to scanning artifact (blurry image).

Manual evaluation of Mandard tumor regression grades

All cases were manually evaluated by two pathologists prior to the digital assessment according to the Mandard TRG and the results are presented in Table 2. Of the 50 included tumors treated with nCRT, 7 (14%) were classified as TRG1, 5 (10%) as TRG2, 15 (30%) as TRG3, 21 (42%) as TRG4, and 2 (4%) as TRG5.

Table 2
Clinicopathological characteristics of the included patients (N = 75).

	Patients treated with nCRT	Patients not treated with nCRT
	n = 50	n = 25
Sex, n (%)		
Male	34 (68)	14 (56)
Female	16 (32)	11 (44)
Age at time of surgery in years, mean (range)	65.5 (37–86)	71.9 (54–88)
(y)pT-category, n (%)		
(y)pT0	7 (14)	0
(y)pT1	1 (2)	1 (4)
(y)pT2	8 (16)	11 (44)
(y)pT3	30 (60)	11 (44)
(y)pT4	4 (8)	2 (8)
(y)pN-category, n (%)		
(y)pN0	30 (60)	15 (60)
(y)pN1	17 (34)	6 (24)
(y)pN2	3 (6)	4 (16)
Manual Mandard TRG, n (%)		
TRG1	7 (14)	Na
TRG2	5 (10)	Na
TRG3	15 (30)	Na
TRG4	21 (42)	Na
TRG5	2 (4)	Na

Na: Not applicable.

Epithelium-Tumor area Percentage (ETP)

A box and whisker plot of the ETP distributions of the patients treated with nCRT compared to the patients not treated with nCRT are presented in Fig. 4A. The median ETP was significantly higher in the resected tumors of the patients who did not receive nCRT 39.25% (33.53–44.23) compared to the group who received nCRT 6.64% (2.07–16.97), ($P < .001$). In patients treated with nCRT, the median ETP increased along with increasing Mandard TRG; 0% in TRG1, 1.23% in TRG2, 3.82% in TRG3, 17.31% in TRG4, and 35.26% in TRG5 (Fig. 4B). Only two tumors were classified as Mandard TRG5 (no regression), and this group was excluded from further analysis. Comparing TRG group 1–4 using the Kruskal–Wallis H test, the distributions of ETP in at least one of the Mandard TRG groups was statistically different compared to the other groups ($P < .001$). Table 3 presents the difference in ETP when comparing the Mandard TRG groups in adjacent pairs, and Fig. 4B show the ETP distributions according to the Mandard TRG groups. The ETP was higher in TRG2 compared to TRG1 ($P = .069$) and in TRG3 compared to TRG2 ($P = .042$) but statistical significance was only reached in TRG4 compared to TRG3 ($P < .001$). The ETP, Mandard grade, number of evaluated slides, and size of ROI in mm^2 according to each case is presented in Table S1 of the supplementary material.

Evaluation of misclassifications and performance of the ETP algorithm

Manual check of all the digitally processed slides showed minor misclassifications by the algorithm. In three of the seven cases manually categorized as complete tumor regression (ypT0, TRG1) small areas (0.023 mm^2, 0.003 mm^2, and 0.035 mm^2 with a corresponding ETP of

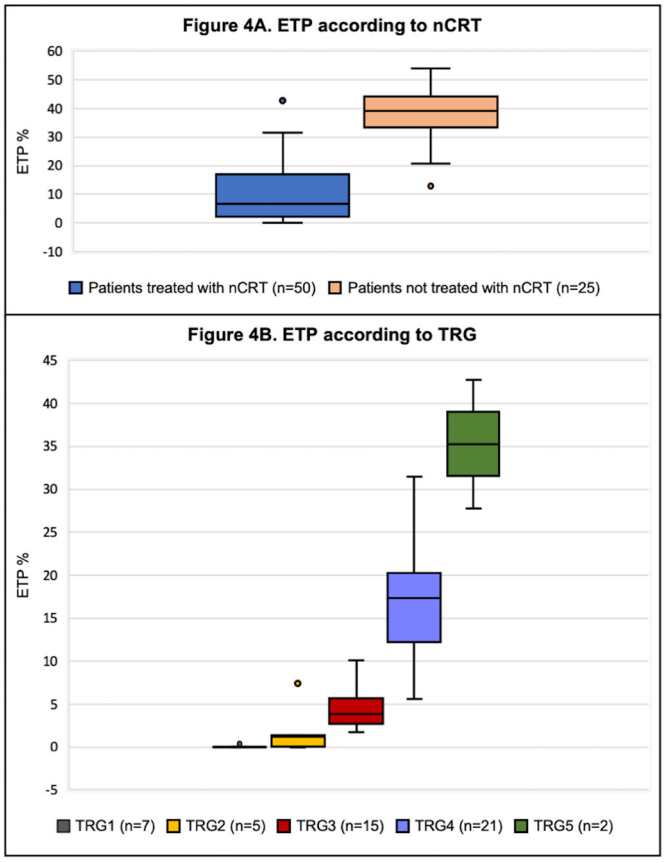

Fig. 4. A: Box and whisker plot of ETP comparing the patients treated with nCRT to the patients not treated with nCRT. Small circles represent outliers in the dataset. B: Box and whisker plot of ETP according to Mandard TRG of the patients treated with nCRT. Small circles represent outliers in the dataset.

Table 3
ETP according to manual Mandard TRG of patients receiving nCRT.

	Mandard TRG1 (N = 7)	Mandard TRG2 (N = 5)	Mandard TRG3 (N = 15)	Mandard TRG4 (N = 21)	Mandard TRG5 (N = 2)
ETP, % Median (Q1, Q3)	0.00 (0.00, 0.03)	1.23 (0.02, 1.36)	3.82 (2.68, 5.69)	17.31 (12.26, 20.27)	35.26 (31.52, 39.01)
ETP, % Range	0.00–0.34	0.003–7.41	1.72–10.11	5.63–31.45	27.78–42.75

Difference in ETP when comparing adjacent Mandard TRG1-TRG4 in pairs

	P-value
ETP of TRG1 vs. TRG2	.069
ETP of TRG2 vs. TRG3	.042
ETP of TRG3 vs. TRG4	<.001*

* Statistically significant *P*-values after adjusting for multiple comparisons.

0.03%, 0.02%, and 0.34%, respectively) were falsely classified as tumor epithelium by the digital algorithm.

The ETP algorithm performance was evaluated by manually correcting labels on 139 slides from 17 cases; three cases with TRG1, three cases with TRG2, three cases with TRG3, three cases with TRG4, two cases with TRG5 (only two patients included in this group), and three cases from the group of patients who did not receive nCRT. The differences in the ETP before and after manual correction of the slides were assessed, resulting in changes in ETP ranging from -0.03% to 0.59% in 16 of the cases, and 2.50% in a single case. The corrected ETP and change in ETP for each case are presented in Table S2.

Discussion

In this study, a digital method to assess the tumor regression in patients with RC following nCRT based on archived H&E-stained slides used in a routine diagnostic setting was developed, and the results were correlated with the manual Mandard method. To our knowledge, this is the first study using deep learning to assess TRG in RC after nCRT. The method was named the ETP to reflect the manual classification system based on the amount of vital tumor epithelium in relation to stroma. A semi-automatic algorithm capable of determining the ETP after manually delineating the tumor area was developed providing an accurate value on a continuous scale of the remaining amount of vital malignant epithelium in the tumor area. Correlation of the ETP with the Mandard method revealed that the median ETP increased with increasing Mandard TRG.

Interestingly, the estimated median ETP did not reflect the Mandard definitions, e.g., the Mandard TRG4 is defined as "Residual cancer outgrowing fibrosis" and the median ETP of the cases classified as TRG4 was 17.31%, meaning that the tumor epithelium only made up 17.31% of the total tumor area incl. areas of regression. This can be the result of the different approach of the digital algorithm. The digital algorithm did not include a separate class for smooth muscle cells which were classified as stroma, resulting in a lower ETP compared with an alternative exclusion of smooth muscle. However, most often the tumor epithelium was surrounded by desmoplastic stroma and only smaller amount of muscle was included in the tumor area. More importantly, luminal parts of the glands including intraglandular debris and white background were excluded from the analysis also resulting in a reduction of ETP. In comparison when manually assessed by the pathologist this area might be considered as part of the remaining tumor area. In attempt to reach a histological homogenous study population, all mucinous AC were excluded, but not AC with a mucinous component or AC including signet ring cells as long as this component was not predominant. Areas of mucin without epithelial cells were included in the ROI of the training set and labeled as part of the tumor stroma. If epithelial cells were present located in the mucin, they were labeled as tumor epithelium. Acellular mucin is often present in cases treated with nCRT, and were thus classified as part of the stroma area. This also applied to cases without complete response and might contribute to a lower ETP compared to a manual estimation.

A control group of patients who did not receive nCRT were included and compared to the patients who did receive nCRT. The wide variation in the ETP of the patients who did not receive nCRT exemplifies the natural variation in tumor morphology and composition experienced between patients. Comparing the two groups, the ETP was significantly lower in tumors treated with nCRT compared to the tumors not treated with nCRT. This is to be expected as the majority of tumors treated with nCRT showed varying degrees of tumor regression; 48 out of the 50 tumors were classified as Mandard TRG1-TRG4. Only two patients showed no sign of regression (Mandard TRG5) and had an ETP comparable to that of the non-treated patients.

Digital assessment has the potential of being more reproducible compared to manual estimations of TRG, thereby minimizing interobserver variation when evaluating TRG in RC. Furthermore, exact values allow for a more specific evaluation of treatment response compared to the categorical data obtained by manual assessment. However, the developed method is not fully automatic as the area for ETP quantification was manually delineated. Determining areas of regression and fibrosis resulting from nCRT can be challenging and thus, interobserver variation is not completely eliminated. In a recently published work studying the practice of tumor regression grading of gastrointestinal carcinomas, the estimation of therapy-induced fibrosis after nCRT was considered difficult or very difficult by more than half of the participating pathologists.[8] Another study revealed that 32% of RC cases were categorized as showing signs of tumor regression (30% with TRG4 and 2% with TRG2) by the pathologist, even though they only received surgery and no nCRT.[14] Some of the morphological changes observed in the tissue after nCRT are not specific to regression changes. Chetty *et al.* provided two alternative reason for fibrosis in rectal cancer specimens after nCRT, other than tumor regression; tumor desmoplasia or fibrosis of the normal tissue around the tumor due to nCRT, and further suggested only considering fibrosis part of regression when intimately associated with other features, e.g., pools of mucin, necrosis, and foamy macrophages, etc.[9] In the present study, areas of regression were defined as fibrosis potentially combined with acellular mucin pools, calcifications, foreign body giant cells, or cholesterol crystals. These features were used as guidance in the selection of areas of regression when manually delineating the ROIs.

Additionally, the size of the ROI analyzed in each case varied. In the cases with complete regression, only one slide was included, resulting in analysis of smaller regions. In the cases of Mandard grade 2–5 and cases not receiving nCRT, the difference in size of the ROIs and number of included slides for each case, reflects the tumor heterogeneity specifically in size, and response to treatment etc.

The prognostic role of TRG in RC treated with nCRT has been debated due to conflicting results, varying methodology, and the use of different TRG methods. However, despite these variations the degree of pathological response is considered associated with prognostic outcomes, and patients with complete pathological response have better prognostic outcomes; e.g., long-term survival and lower rates of local recurrence.[7,15–19]

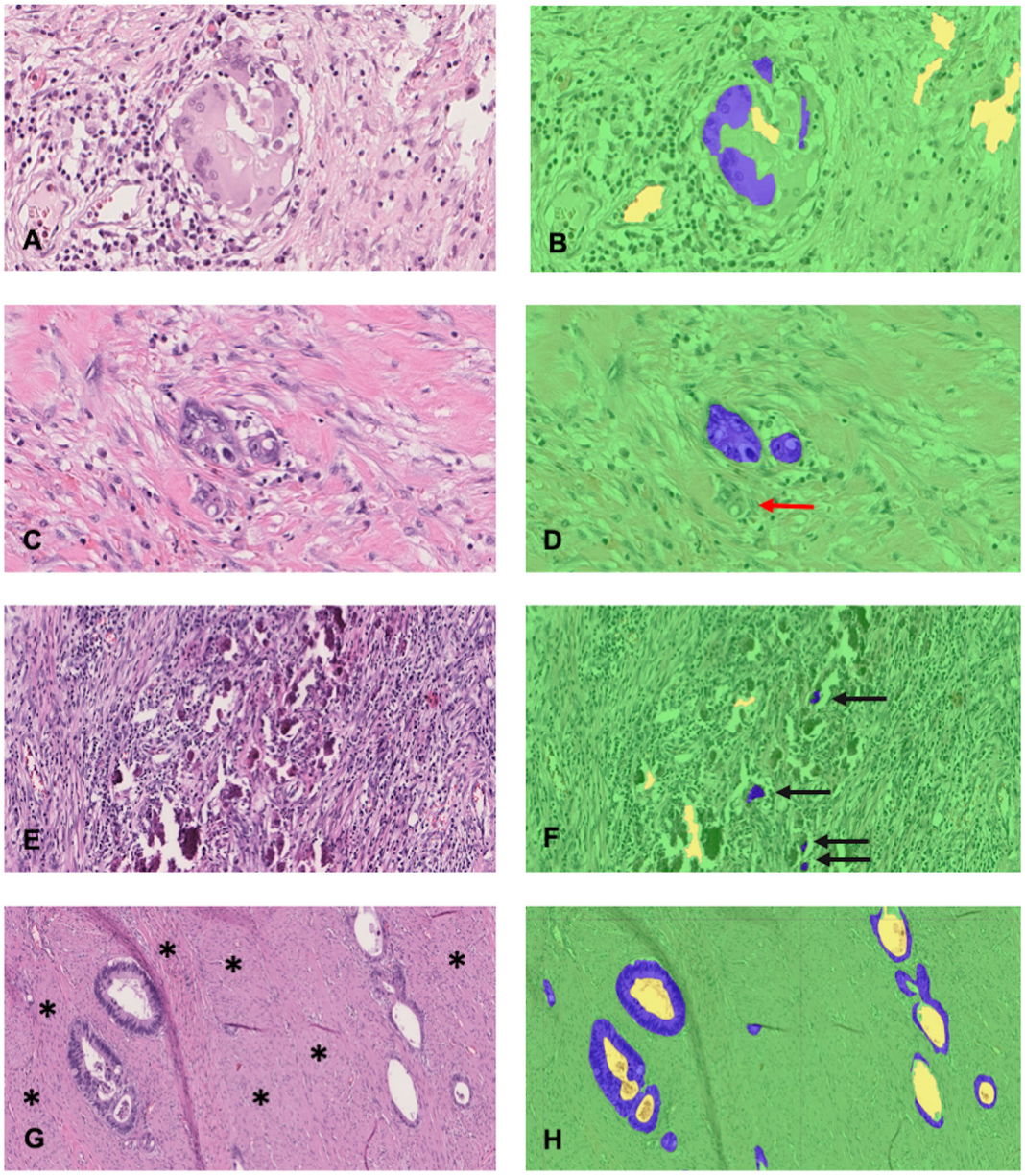

Fig. 5. Examples of limitations and misclassification of the ETP algorithm. A–B (x40): Foreign body giant cell misclassified as tumor epithelium (blue). C–D (x40): A few tumor cells were missed and misclassified as stroma (red arrow). E–F (x20): A few calcifications were misclassified as tumor epithelium (black arrows). G–H (x20): Tumor islands within the muscularis propria, and smooth muscle cells (*) classified as stroma.

Additionally, some studies have shown superiority in predicting recurrence and survival by some of the TRG methods compared with others.[20,21] A 3-tiered TRG method compared to 5-tiered TRG method has also proven to be more reproducible.[7] Digital algorithms estimating tumor regression have potential to replace or at least be an aid to the assessment performed by the pathologists, but needs validation in larger prognostic studies before considering implementation in routine practice.

The developed ETP algorithm has limitations, and examples are presented in Fig. 5. In some cases, small areas were misclassified as tumor epithelium, typically areas of granulation tissue or reactive fibroblasts, foreign body giant cells, ganglion cells, or calcification. In other cases, small areas of tumor epithelium were missed, most frequently in cases with single tumor cells, dedifferentiated tumors, or squeezed and damaged tissue. As exemplified by manual correction in a subset of cases, the tumor epithelium misclassifications only resulted in minor changes in the ETP. Additionally, artifacts in the tissue (folded tissue, air bubbles, etc.) could also result in misclassifications.

In recent years, increasing attention has been paid to image-based machine learning, including neural networks and deep learning, evaluating different aspects of CRC.[22] Studies have used deep learning to examine prognosis, identify prognostic biomarkers, and in predicting molecular profiles, based on standard H&E-stained slides in CRC.[23–26] As new treatment strategies with total neoadjuvant treatment and immunotherapy emerge,[27–29] an accurate tumor regression grade is highly important as the pathological response often is used as endpoint in clinical trials.

Conclusion

In this study, a semi-automatic digital deep learning-based ETP algorithm was developed to assess TRG in RC treated with nCRT. The ETP algorithm is an objective and precise method resulting in continuous data. However, it is important to keep in mind the existing natural variation in the composition of a tumor between patients which was demonstrated by the variance of the ETP in the control group. At the moment, this variation

is not taken into account when evaluating TRG neither in manual assessment nor in the developed algorithm. A biopsy-adapted ETP would probably have added value to obtain an even further standardized assessment of treatment response.

Supplementary data to this article can be found online at https://doi.org/10.1016/j.jpi.2022.100152.

Funding

This work was supported by grants from the Folketingsmand Jens Christensen and wife Korna Christensens Foundation, the Harboe Foundation, the Buschard Foundation, the Købmand Kristjan Kjær and wife Margrethe Kjærs Foundation, the Region Zealand Health Research Foundation, the Carl and Ellen Hertz' Scholarship for Danish Medical and Natural Sciences, the Inge and Jørgen Larsens Memorial Scholarship, the Manufacturer Einar Willumsens Memorial Scholarship, the A. V. Lykfeldts and wifes Scholarship. These funders had no role in the design of the study, collection of data, analysis of data, interpretation of data, the writing of the report or the decision to submit the article for publication.

Author contribution

Study design and concept: DNMJ, AKF
Execution: DNMJ, HH, JT, JSW, AKF
Analysis and interpretation of data: DNMJ, JSW, IG, AKF
Writing – first draft: DNMJ
Writing – review and editing: DNMJ, HH, JT, JSW, IG, AKF
All authors read and approved the final manuscript.

Conflicts of interest

DNMJ, JSW, IG, AKF: No conflicts of interest related to the work described.

HH, JT: Are employed at Visiopharm A/S, Hoersholm, Denmark.

Acknowledgments

None.

References

1. Sung H, Ferlay J, Siegel RL, et al. Global cancer statistics 2020: GLOBOCAN estimates of incidence and mortality worldwide for 36 cancers in 185 countries. CA Cancer J Clin 2021;71:209–249. https://doi.org/10.3322/caac.21660.
2. Siegel RL, Fedewa SA, Anderson WF, et al. Colorectal cancer incidence patterns in the United States, 1974-2013. J Natl Cancer Inst 2017:109. https://doi.org/10.1093/jnci/djw322.
3. Glynne-Jones R, Wyrwicz L, Tiret E, et al. Rectal cancer: ESMO clinical practice guidelines for diagnosis, treatment and follow-up. Ann Oncol Off J Eur Soc Med Oncol 2017;28:iv22–iv40. https://doi.org/10.1093/annonc/mdx224.
4. Mandard AM, Dalibard F, Mandard JC, et al. Pathologic assessment of tumor regression after preoperative chemoradiotherapy of esophageal carcinoma. Clinicopathologic correlations. Cancer 1994;73:2680–2686. https://doi.org/10.1002/1097-0142(19940601)73:11<2680::aid-cncr2820731105>3.0.co;2-c.
5. Dworak O, Keilholz L, Hoffmann A. Pathological features of rectal cancer after preoperative radiochemotherapy. Int J Colorectal Dis 1997;12:19–23. https://doi.org/10.1007/s003840050072.
6. Amin MB, Edge S, Greene F, et al. *AJCC Cancer Staging Manual. 8th ed.* Chicago: Springer International Publishing. 2017.
7. Ryan R, Gibbons D, Hyland JMP, et al. Pathological response following long-course neoadjuvant chemoradiotherapy for locally advanced rectal cancer. Histopathology 2005;47:141–146. https://doi.org/10.1111/j.1365-2559.2005.02176.x.
8. Westerhoff M, Osecky M, Langer R. Varying practices in tumor regression grading of gastrointestinal carcinomas after neoadjuvant therapy: results of an international survey. Mod Pathol an Off J United States Can Acad Pathol Inc 2020;33:676–689. https://doi.org/10.1038/s41379-019-0393-7.
9. Chetty R, Gill P, Govender D, et al. International study group on rectal cancer regression grading: interobserver variability with commonly used regression grading systems. Hum Pathol 2012;43:1917–1923. https://doi.org/10.1016/j.humpath.2012.01.020.
10. Agarwal A, Chang GJ, Hu C-Y, et al. Quantified pathologic response assessed as residual tumor burden is a predictor of recurrence-free survival in patients with rectal cancer who undergo resection after neoadjuvant chemoradiotherapy. Cancer 2013;119:4231–4241. https://doi.org/10.1002/cncr.28331.
11. Niazi MKK, Parwani AV, Gurcan MN. Digital pathology and artificial intelligence. Lancet Oncol 2019;20:e253–e261. https://doi.org/10.1016/S1470-2045(19)30154-8.
12. WHO Classification of Tumours. *Digestive system tumours. 5th ed.* Lyon (France): International Agency for Research on Cancer. 2019.
13. Erichsen R, Lash TL, Hamilton-Dutoit SJ, Bjerregaard B, Vyberg M, Pedersen L. Existing data sources for clinical epidemiology: the Danish National Pathology Registry and Data Bank. Clin Epidemiol 2010;2:51–56. https://doi.org/10.2147/clep.s9908.
14. Vironen J, Juhola M, Kairaluoma M, Jantunen I, Kellokumpu I. Tumour regression grading in the evaluation of tumour response after different preoperative radiotherapy treatments for rectal carcinoma. Int J Colorectal Dis 2005;20:440–445. https://doi.org/10.1007/s00384-004-0733-y.
15. Martin ST, Heneghan HM, Winter DC. Systematic review of outcomes after intersphincteric resection for low rectal cancer. Br J Surg 2012;99:603–612. https://doi.org/10.1002/bjs.8677.
16. Huh JW, Kim HC, Kim SH, et al. Tumor regression grade as a clinically useful outcome predictor in patients with rectal cancer after preoperative chemoradiotherapy. Surgery 2019;165:579–585. https://doi.org/10.1016/j.surg.2018.08.026.
17. Kong JC, Guerra GR, Warrier SK, et al. Prognostic value of tumour regression grade in locally advanced rectal cancer: a systematic review and meta-analysis. Color Dis Off J Assoc Coloproctology Gt Britain Irel 2018;20:574–585. https://doi.org/10.1111/codi.14106.
18. Fokas E, Liersch T, Fietkau R, et al. Tumor regression grading after preoperative chemoradiotherapy for locally advanced rectal carcinoma revisited: updated results of the CAO/ARO/AIO-94 trial. J Clin Oncol Off J Am Soc Clin Oncol 2014;32:1554–1562. https://doi.org/10.1200/JCO.2013.54.3769.
19. On J, Shim J, Mackay C, et al. Pathological response post neoadjuvant therapy for locally advanced rectal cancer is an independent predictor of survival. Color Dis Off J Assoc Coloproctology Gt Britain Irel 2021;23:1326–1333. https://doi.org/10.1111/codi.15512.
20. Trakarnsanga A, Gönen M, Shia J, et al. Comparison of tumor regression grade systems for locally advanced rectal cancer after multimodality treatment. J Natl Cancer Inst 2014:106. https://doi.org/10.1093/jnci/dju248.
21. Kim SH, Chang HJ, Kim DY, et al. What is the ideal tumor regression grading system in rectal cancer patients after preoperative chemoradiotherapy? Cancer Res Treat 2016;48:998-1009. https://doi.org/10.4143/crt.2015.254.
22. Pacal I, Karaboga D, Basturk A, Akay B, Nalbantoglu U. A comprehensive review of deep learning in colon cancer. Comput Biol Med 2020;126, 104003. https://doi.org/10.1016/j.compbiomed.2020.104003.
23. Kather JN, Krisam J, Charoentong P, et al. Predicting survival from colorectal cancer histology slides using deep learning: a retrospective multicenter study. PLoS Med 2019;16, e1002730. https://doi.org/10.1371/journal.pmed.1002730.
24. Skrede O-J, De Raedt S, Kleppe A, et al. Deep learning for prediction of colorectal cancer outcome: a discovery and validation study. Lancet (London, England) 2020;395:350–360. https://doi.org/10.1016/S0140-6736(19)32998-8.
25. Bychkov D, Linder N, Turkki R, et al. Deep learning based tissue analysis predicts outcome in colorectal cancer. Sci Rep 2018;8:3395. https://doi.org/10.1038/s41598-018-21758-3.
26. Yamashita R, Long J, Longacre T, et al. Deep learning model for the prediction of microsatellite instability in colorectal cancer: a diagnostic study. Lancet Oncol 2021;22:132–141. https://doi.org/10.1016/S1470-2045(20)30535-0.
27. Jimenez-Rodriguez RM, Quezada-Diaz F, Hameed I, et al. Organ preservation in patients with rectal cancer treated with total neoadjuvant therapy. Dis Colon Rectum 2021;64:1463–1470. https://doi.org/10.1097/DCR.0000000000002122.
28. Petrelli F, Trevisan F, Cabiddu M, et al. Total neoadjuvant therapy in rectal cancer: a systematic review and meta-analysis of treatment outcomes. Ann Surg 2020;271:440–448. https://doi.org/10.1097/SLA.0000000000003471.
29. André T, Shiu K-K, Kim TW, et al. Pembrolizumab in microsatellite-instability-high advanced colorectal cancer. N Engl J Med 2020;383:2207–2218. https://doi.org/10.1056/NEJMoa2017699.